# COMPARISON OF FIVE ACCELEROMETER METRICS FOR ASSESSING THE TEMPORAL PATTERNS OF CHILDREN'S FREE-PLAY PHYSICAL ACTIVITY

By

Katherine Louise McKee

# A THESIS

Submitted to Michigan State University in partial fulfillment of the requirements for the degree of

Department of Kinesiology - Master of Science

#### ABSTRACT

Accelerometers are frequently used to measure physical activity (PA) in children, which is important for overall health and development. Lack of uniformity in data processing methods, such as the metric used to summarize accelerometer data, limits comparability between studies. The objective was to determine the convergent validity of five accelerometer metrics for characterizing the intensity and temporal patterns of first and second graders' (n=88) recess PA. At a 5-s epoch level, Pearson's correlations between various metrics ranged from 0.69 to 0.98. When each epoch was classified into one of four activity levels based on quartiles, agreement between metrics as indicated by weighted kappa ranged from 0.81 to 0.96. When collapsed to time spent in each activity level, metrics were most often statistically equivalent for estimating time spent in quartile 3 or 4. Children were ranked from least to most active, and agreement between metrics was strong with Spearman's correlation coefficients of over r=0.86. Temporal patterns were characterized using five fragmentation indices calculated using each of the five metrics. Pearson's correlations between metrics ranged from r=0.53 to 0.99, with the strongest associations for number of high activity bouts. Most fragmentation indices were not statistically equivalent between metrics. While metrics captured similar trends in activity intensity and temporal patterns, caution is warranted when making comparisons of point estimates derived from different metrics. However, all metrics were able to similarly capture higher intensity activity (i.e., quartile 3 or 4), the most common outcome of interest in intervention studies

INTRODUCTION	1
LITERATURE REVIEW	5
Nature of Children's Physical Activity	5
Measurement of Children's Physical Activity	7
Accelerometry	7
Count-Based Metrics	9
Vertical Axis Counts	10
Triaxial Vector Magnitude Counts	10
Acceleration-Based Metrics	10
Mean amplitude deviation	12
Euclidean norm minus one	13
Activity index	14
Comparability of Metrics	15
Temporal Patterns of Physical Activity	17
Measurement of Temporal Patterns	17
Conclusions	19
METHODS	20
Setting and Participants	20
Data Collection	20
Data Processing	21
Physical Activity Intensity	21
Temporal Patterns	22
Statistical Analyses	22
DECLUTE	24
Destiniant characteristics	24
A jm 1	24
AIIII 1	24
AIIII 2	32
DISCUSSION	46
Strengths & Limitations	50
Future Directions	50
REFERENCES	52

# **TABLE OF CONTENTS**

#### **INTRODUCTION**

Regular physical activity (PA) participation during childhood is associated with improved cognition, increased bone density and muscular strength, and decreased risk of developing noncommunicable disease in later life (Loprinzi et al., 2012). Despite these and other benefits, many children fail to meet the recommendations set forth in the National Physical Activity Guidelines for Americans, which state that children aged 6- through 17-years should engage in at least 60-minutes day<sup>-1</sup> of moderate-to-vigorous physical activity (MVPA) (Health and Human Services, 2018). Unfortunately, only about 25% of school-aged children meet PA recommendations according to both parent-report and device-based measures of PA (Katzmarzyk et al., 2016; Troiano et al., 2008). This is a major public health concern, thus prompting investigation into the underlying causes of insufficient PA levels among children.

Interventions aimed at promoting and increasing children's PA often occur in the school setting due to the bulk of time children spend in school. There are few opportunities for children to engage in PA during the school day; however, some opportunities for PA include physical education (Meyer et al., 2013), classroom activity breaks (Carlson et al., 2015), or after-school programs (Arundell et al., 2015). The outdoor recess period, however, is often the only opportunity for children to engage in a developmentally important type of play called unstructured free-play PA, in which children are able to determine the structure and intensity of play without a high degree of adult influence (Ginsberg, 2007). This type of play, along with the benefits from regular PA engagement, contributes to children's social, emotional, and physical development (Ginsberg, 2007). Therefore, efforts to optimize the unstructured free play PA that the outdoor recess period provides is salient in improving the overall well-being of children.

Research on children's PA during recess has primarily focused on overall PA levels (Ridgers et al., 2005; Mota et al., 2005; Holmes et al., 2012) or differences amongst groups (e.g., boys vs girls; Trost et al., 2002; Mota et al., 2005; Pate et al., 2013). Interventions to promote PA during recess have included strategies like modifications to the physical environment (e.g., painted markings; Ridgers et al., 2007; Blaes et al., 2013) and addition of portable equipment (e.g., playground balls, jump ropes; Ridgers, Fairclough, & Stratton, 2010). However, little research has considered the chronological succession of children's activity over time, or the temporal patterns of PA. It is known that children do not maintain the same level of PA over the recess period (McKenzie et al., 1997); therefore, better understanding of the temporality of PA can highlight critical periods when children are most and/or least active during recess. Further research into how these temporal patterns of recess PA differ by participant demographic or environmental factors is warranted as this information can be used to create activity-promoting interventions to optimize PA participation during the recess period.

Accelerometry is an ideal method to capture children's PA due to its ability to capture the timing, frequency, intensity, and duration of movement, as well as the temporality of children's PA. However, there are a multitude of decisions required to process the accelerometer data, and lack of consistency amongst processing methods limits data harmonization and comparability across studies. Accelerometers measure acceleration in gravitational units (g; 1 g = 9.8 ms<sup>-2</sup>), which has historically been filtered, rectified, and summed over a specific period, or epoch (e.g., 5-s) as activity counts (Farrahi et al., 2019; Strath et al., 2012). Each of the epochs can then be assigned an activity intensity of light, moderate, or vigorous depending on the count value using population-specific thresholds called cut points, ultimately allowing researchers to determine time spent in each PA intensity.

Over the years, advances in accelerometry have resulted in the shift from vertical axis (VA) counts to triaxial counts measured across three planes (i.e., anteroposterior, mediolateral, vertical), aggregated as vector magnitude (VM) counts. However, the proprietary nature of the manufacturer-specific algorithms used to calculate the count-based metrics has greatly reduced generalizability of results across studies, particularly when studies use different device manufacturers (Duncan et al., 2016; Rowlands et al., 2014). A proposed solution to this issue was the development of open-source metrics based upon the raw acceleration – not activity counts – to determine activity intensities (Farrahi et al., 2019; Clevenger et al., 2020). Since researchers gained access to raw acceleration data around a decade ago (John & Freedson, 2012), several acceleration-based metrics have emerged. Specifically, three commonly used acceleration-based metrics are mean amplitude deviation (MAD) (Vähä-Ypyä et al., 2015; Aittasalo et al., 2015), the Euclidean norm minus one (ENMO) (van Hees et al., 2013; Hildebrand et al., 2017), and the activity index (AI) (Bai et al., 2016).

Previous research in adults has investigated the comparability of these metrics for assessing overall PA participation (Sasaki et al., 2013; Bakrania et al., 2016; Karas et al., 2022) but there is a paucity of information regarding the comparability of these acceleration-based metrics compared to count-based metrics in populations of children, who participate in more sporadic and variable PA compared to adults (Bailey et al., 1995; Welk, Corbin, & Dale, 2000). Furthermore, little research has investigated whether these metrics similarly capture temporal patterns of activity. Therefore, the overall purpose of this study is to determine the convergent validity of count- and acceleration-based metrics to measure the overall levels and temporality of children's PA during recess. Two specific aims and their respective hypotheses will be examined.

Aim 1: To determine the convergent validity of VM, VA, MAD, ENMO and AI when characterizing children's PA intensity during recess, at both the epoch and participant level.

Hypothesis 1A: At the epoch-level, acceleration-based metrics will show moderate to high correlations with each other (r 0.60) and count-based metrics will show moderate to high correlation with each other (r 0.60) but not with acceleration-based metrics (r 0.60).

Hypothesis 1B: Time spent in each quartile of activity intensity will be equivalent between all metrics and agreement at the epoch-level will be substantial (k = 0.61 to 0.80).

Hypothesis 1C: At the participant-level, rank-order of children's overall PA will be similar amongst all metrics (0.80)

Aim 2: To determine comparability of VM, VA, MAD, ENMO and AI when characterizing temporal patterns of children's PA during recess.

Hypothesis 2A: Temporal patterns, as assessed using fragmentation indices (transition probabilities, mean fragment duration, and number of fragments), will be moderately to highly correlated amongst all metrics (r 0.60).

Hypothesis 2B: Fragmentation indices will be statistically equivalent among accelerometer metrics.

Investigation into these aims will help characterize the comparability of accelerometer metrics for measuring both overall activity levels and the temporal pattern of PA, thus adding to the growing body of literature utilizing acceleration-based metrics instead of traditional countbased approaches to capturing PA. This information can inform future study design and enhance PA researchers understanding of comparability amongst studies.

#### LITERATURE REVIEW

It has been thoroughly substantiated that regular participation in PA during childhood is important in reducing risk of noncommunicable disease (Lee et al., 2012). The school setting is an attractive environment to analyze children's PA because children spend about 30-hours per week in school (Nettlefold et al., 2011). One of the most commonly used tools for measurement of PA is the accelerometer (Trost & O'Neil et al., 2014). However, there is little uniformity when it comes to processing and analyzing the acceleration data produced from these devices, leading to misinterpretation of study results and inability to generalize findings across studies. Following an overview of the nature, benefits, and context of children's physical activity with a focus on the school recess setting, this literature review will discuss analysis of accelerometer data for capturing PA intensity and characterizing temporal patterns of children's PA.

#### Nature of Children's Physical Activity

In their seminal paper from 1995, Bailey and colleagues highlighted the highly transitory nature of children's PA which occurs in bouts lasting no more than 15-s at a time. Currently, the precise reasoning behind this pattern of PA innate to children is unknown. One prominent theory posits that the spontaneous PA movements demonstrated by children provides the central nervous systems with adequate stimulation, information about one's environment, and helps maintain homeostasis via energy balance and metabolic regulation (Hills, King, & Armstrong, 2007). Because children's PA is inherently different than adult PA (Bailey et al., 1995), special attention must be made when selecting an appropriate method of PA measurement in children to ensure it is being accurately captured.

While PA is largely responsible for children's physical development, regular participation in PA during childhood is also positively associated with children's cognitive and

socio-emotional development, as problem-solving and social skills are developed through play interaction with other children (Bjorklund & Brown, 1998; Pellegrini & Smith, 1998). Therefore, encouragement of PA is important to facilitate typical growth and development of children overall (Rowland, 1998). While the developmental benefits of PA participation are abundantly clear, more than half of children in the United States struggle to meet the National Physical Activity Guidelines for Americans (Katzmarzyk et al., 2016; Troiano et al., 2008; HHS, 2018). This is a major public health concern and has led to the development and implementation of PA interventions aimed at increasing children's PA levels.

The context in which PA occurs is an important factor to consider when promoting children's PA. Many PA-promoting interventions are focused on the school setting because most children spend the bulk of their time in school. While there are other opportunities for PA during the school day like physical education or active classroom breaks, recess is of particular interest because it is often the only opportunity for unstructured PA during the otherwise structured school day (Mota et al., 2005). Recess contributes about 20% of children's daily PA, although children spend approximately half of recess participating in MVPA (Rooney, 2018). Common interventions to promote PA during recess include addition of painted playground markings (Stratton & Mullan, 2005; Ridgers et al., 2007) or portable equipment (Ridgers, Fairclough, & Stratton, 2010; Yu, Kulinna, & Mulhearn, 2021). However, a more thorough understanding of the temporal patterns of recess PA could provide useful information to inform novel activity-promoting interventions, since children's PA does not occur at a continuous rate over the course of recess (McKenzie et al., 1997; Holmes, 2012). Understanding the temporality of children's PA can highlight periods and or patterns of high and low PA and ultimately provide PA

researchers with ways to maintain PA levels over the recess period and increase children's overall PA.

#### **Measurement of Children's Physical Activity**

There are a multitude of ways to measure children's PA. Some of the most common methods are direct observation, self- or proxy-reported measures like questionnaires or daily activity logs, and device-based measures like pedometry or accelerometry. Direct observation is considered a gold standard method of PA measurement, particularly when observers are thoroughly trained. However, direct observation inherently has a high degree of researcher burden (i.e., time-consuming, high-energy cost) making it an unattractive option in large-scale studies (Gardner, 2000; Sylvia et al., 2013; Rachele et al., 2013). Self-reported measures like questionnaires in young populations are often obscured by children's attention spans and tendency to inaccurately recount their daily activities, particularly if the recall period is longer than one week, while proxy-reports are often not feasible since children are not always with their parent or guardian (Sylvia et al., 2013; Corder et al., 2013; Biddle et al., 2011). While pedometers are relatively inexpensive, they are unable to capture horizontal or upper limb movements, as well as the frequency, intensity, timing, and duration of PA, all of which are key contributors to overall PA and can ultimately result in inaccurate quantification of PA (Rowlands & Eston, 2007; Sylvia et al., 2013; Butte et al., 2012). While all methods of PA measurement have their respective strengths, weaknesses, and uses, accelerometry is a device-based method of measurement with specific traits that make it ideal when capturing children's PA patterns.

#### Accelerometry

Accelerometry is frequently used to capture PA of children due to its high sampling rates and ability to collect and store large amounts of data. One of the most commonly used

manufacturers of research-grade accelerometers in PA research is ActiGraph (Actigraph LLC, Pensacola, FL). ActiGraph has produced several generations of activity monitors with each iteration receiving updates in physical appearance, storage capacity, hardware and firmware, Bluetooth capability, and number of axes (e.g., uniaxial versus triaxial) (Grydeland et al., 2014). ActiGraph's hallmark monitor, the wGT3x-BT, has been validated in age groups spanning from preschool to older adulthood, in populations with physical disabilities, and across multiple wearlocations (e.g., hip, wrist, ankle) (Albaum et al., 2019; Hänggi et al., 2014; Johansson et al., 2015; Johansson et al., 2016; Karaca et al., 2021; McGarty, Penpraze, & Melville, 2016; Nyström et al., 2017; Pate et al., 2013; Peterson et al., 2015; Sylvia et al., 2013; Trost & O'Neil, 2014). Accelerometry can detect information about the frequency, intensity, timing, and duration of PA across a variety of participants and settings. Lastly, this high-resolution data can be collected for weeks at a time, stored, and analyzed at a later date, making accelerometers a particularly advantageous tool in large scale epidemiological studies.

There are several decisions to be made when collecting and analyzing accelerometer data, and each will impact the data collected and, ultimately, the harmonization of data and comparability of results across studies. One decision that must be made is the 'accelerometer metric,' or the way in which raw acceleration data are summarized, which can subsequently be used to identify activity intensity. For example, acceleration data can be summarized as variability in acceleration or total magnitude of the acceleration. While the accelerometer metric will be the focus of this literature review, other decisions include (but are not limited to) device brand, wear location, and epoch length.

Accelerometer metrics are often summarized over a time interval, usually ranging from 1 to 60-s, called an epoch. From there, "intensity bins" called cut points can be used to classify

each epoch as a certain activity intensity. Many PA measurement researchers have created their own cut points to classify activity intensities, and selection of the "right" or "best" cut point greatly depends on researcher judgement in relation to other published work in the same population (Kim et al., 2012). For example, Freedson et al. (2005), Evenson et al. (2008), and Puyau et al. (2002) have all developed their own cut points for moderate-to-vigorous PA (MVPA) intensity for children aged 6- through 10-years. Not only do these cut points often yield considerably different estimates of activity intensities, they are often validated under dissimilar conditions, such as different age groups (e.g., 5-8 y vs. 6-18 y) (Trost et al., 2012; Evenson et al., 2008; Pfeiffer et al., 2006; Clevenger et al., 2020), different wear locations (e.g., hip vs waist) (Crotti et al., 2020; Hangii et al., 2013; Montoye et al., 2018), and different epoch lengths (Hilsop et al., 2012) making it nearly impossible to equate findings across studies.

#### **Count-Based Metrics**

Activity counts or count-based metrics have been the standard since accelerometers were first used to capture movement. Counts are derived from the raw acceleration and summed over the three axes. From there, the counts are aggregated to a user-specific time interval ranging from 1 to 60-s, called an epoch. Next, upper and lower thresholds are applied to the count data called cut points to separate the data into light, moderate, and vigorous intensity activity. Cut points are created by PA researchers and often times validated against criterion measures such as oxygen consumption or heart rate to determine intensity level. However, since the count data originate from the software-specific proprietary algorithms that for many years were not opensourced, transparency and comparability of estimates of time spent in activity intensities from various device brands, and various applied cut points, is not possible.

#### **Vertical Axis Counts**

Vertical axis (VA) counts are the uniaxial count values produced exclusively from the vertical axis (i.e., y-axis). In the early years of development, VA were widely used and have their own validated cut points (Freedson, Melanson, & Sirard, 1998). However, typical human movement does not occur exclusively in one plane of motion, and utilization of VA disregards important acceleration produced from the other planes of motion (Howe, Staudenmayer, & Freedson, 2009), urging the use of triaxial vector magnitude counts. Despite their potential shortcomings, VA are still the most used metric in children (Migueles et al., 2017).

### **Triaxial Vector Magnitude Counts**

Triaxial vector magnitude (VM) counts are defined as the square root of the sum of the squared count values across each axis. The equation to calculate VM is:

$$\sqrt{a_1^2 + a_2^2 + a_3^2}$$

Where a<sub>1</sub> equals the axis 1 (vertical) counts, a<sub>2</sub> equals axis 2 (mediolateral) counts, and a<sub>3</sub> equals axis 3 (anteroposterior) counts. Compared to uniaxial VA, triaxial VM is less dependent on monitor orientation as the metric encompasses all three axes, making it advantageous to use in children.

#### **Acceleration-Based Metrics**

In 2009, a council of PA measurement scholars at the Objective Measurement of Physical Activity Expert Consensus Meeting expressed the need for a more transparent and open-sourced method to process accelerometer data (Troiano, 2005; Welk et al., 2012; Rowlands, 2018). From this meeting emerged the shift to raw acceleration-based metrics, which are calculated using the data obtained by the raw signal before any other processing has occurred (e.g., reintegration into

activity counts) (Rowlands, 2018). The use of raw acceleration data is attractive for a variety of reasons. Until recently, the algorithms used to calculate activity counts were proprietary to each device manufacturer, which greatly limited transparency of processing methods and generalizability of data sets. Using the raw acceleration theoretically allows findings to be equated across studies regardless of device manufacturer (Aittasalo et a., 2015) and, ultimately, create more meaningful ways to interpret PA data obtained via surveillance and/or intervention research (Rowlands, 2018). Assuming the accelerometer has been appropriately calibrated, has similar dynamic range, and sampling rate, among other factors, the acceleration output should, theoretically, be comparable across brands (e.g., ActiGraph, GENEActiv, Hookie) due to the fact that data are collected in the same unit (i.e., g), and that there is no opaque filtration or processing (Rowlands, 2018; Aittasalo et a., 2015; Rowlands et al., 2018). While acceleration-based metrics may still rely on cut points to determine activity intensities, the cut points themselves are no longer derived from the arbitrary activity counts.

A variety of acceleration-based metrics have gained popularity throughout the years as the call to switch to metrics utilizing the raw acceleration has grown. While machine learning approaches (e.g., random forests, artificial neural networks) for raw acceleration have also been developed, they will not be included in this literature review. A non-exhaustive list of metrics that have gained traction among PA researchers in recent years are as follows: Euclidean norm of the high-pass filtered signals (HFEN), HFEN plus Euclidean norm of the low-pass filtered signals minus 1 g (HFEN+), Euclidean norm minus one with negative values set to zero (ENMO) (van Hees et al., 2013), mean amplitude deviation (MAD) (Aittasalo et al., 2015; Vähä-Ypyä et al., 2015), and the activity index (AI) (Bai et al., 2016). In its conception paper, HFEN+ outperformed metrics HFEN and ENMO when detecting total PA energy expenditure (van Hees

et al., 2013). However, HFEN+ proved to be too computationally complex for understanding and, ultimately reproduction, compared to ENMO, thus, halting its development (van Hees et al., 2013). Therefore, MAD, ENMO, and the AI have been selected for this proposed master's thesis and will be the only acceleration-based metrics discussed in the remainder of the literature review.

#### Mean amplitude deviation

Mean amplitude deviation (MAD), measured in milli-gravitational units (mg), is a raw acceleration metric that captures the variation around the mean of the raw acceleration signals (Aittasalo et al., 2015; Vähä-Ypyä et al., 2015). The equation to calculate MAD is:



Where n is the number of samples in each epoch, is the resultant acceleration in the *i*th timepoint, and is the mean resultant acceleration value of the entire epoch. The resultant acceleration is defined as the vector magnitude of acceleration in the three axes (Vähä-Ypyä et al., 2015).

MAD has been developed and validated in a variety of populations, but the most applicable to the current thesis was in a population of 20 healthy adolescents aged 13- to 15years (female=10; age=14.2) (Aittasalo et al., 2015). Participants wore a heart rate monitor at the chest and two accelerometers, an ActiGraph GT3X at the left hip and Hookie AM13 at the right hip, on an elastic belt while performing a variety of free-living activities (e.g., sitting while working on computer, lying supine) as well as walking and jogging around a 100-m indoor track. The MAD values obtained via the ActiGraph and Hookie monitors exhibited a strong linear

correlation with heart rate (0.97 and 0.96, respectively). For the ActiGraph, three intensityspecific cut points were determined to separate sedentary, light, moderate, and vigorous activity (26.9, 332.0, 558.3 mg).

Despite using two different manufacturers (e.g., ActiGraph and Hookie), the MAD cut points obtained from both devices were mostly consistent and differed, at most, by 45.5 mg (558.3 mg vs 603.8 mg for ActiGraph and Hookie, respectively) at the vigorous intensity. Aittasalo et al. (2018) state that this dissonance could be due in part to Hookies' higher sampling frequency of 100Hz, compared to ActiGraph's 30Hz sampling rate at the time of the study. Since this study, ActiGraph monitors have the ability to sample at 100Hz, which may result in even more precise MAD performance between brands.

#### Euclidean norm minus one

Euclidean norm minus one (ENMO), reported in milli-gravitational units (mg), is a second raw acceleration metric that is calculated as the square root of the sum of the squared accelerations values in each axis, minus 1, as an adjustment for gravity, and with the negative values rounded to zero (van Hees et al., 2013; Hildebrand et al., 2014). Some researchers calculate ENMO after conducting an auto-calibration procedure (van Hees et al., 2014).

$$\sqrt{a_1^2 + a_2^2 + a_3^2}$$
 - 1

where negative values are rounded to zero after subtraction

Where a<sub>1</sub> equals axis 1 acceleration, a<sub>2</sub> equals axis 2 acceleration, and a<sub>3</sub> equals axis 3 acceleration. ENMO was developed and validated in children and adults aged 7- through 11- years and 18- through 65-years, respectively (Hildebrand et al., 2014). Participants completed eight activities of varying intensities (e.g., lying, sitting, running) while wearing an ActiGraph

and GENEActiv accelerometer at the right hip and non-dominant wrist. Oxygen consumption (VO<sub>2</sub>) was measured via portable metabolic unit and metabolic equivalents (METs) were used as the criterion measure for development of ENMO thresholds. MET values were classified as light (<3 METs), moderate (>3 but <6 METs), or vigorous (6 METs) intensity activity for both children and adults (Ainsworth et al., 2011).

Overall, the ENMO metric exhibited strong correlations with VO<sub>2</sub> (Hildebrand et al., 2014). The intensity classification accuracy of the developed ENMO cut points was highest when detecting sedentary/light intensity activities, correctly identifying 93-96% of values, and lowest at moderate intensity activities, correctly identifying 54-59% of values. Between brands, ENMO performed well, demonstrating no main differences in adults or children.

#### Activity index

The activity index (AI) is the final acceleration-based metric of interest that will be used in the current investigation. AI is a unitless metric and is calculated as the variance of the acceleration value along the three axes (Bai et al., 2016). The equation to calculate the AI is:

$$AI_{i}^{new}(t;H) = \sqrt{\max\left(\frac{1}{3}\left\{\sum_{n=1}^{3}\frac{\sigma_{im}^{2}(t;H) - \bar{\sigma}_{i}^{2}}{\bar{\sigma}_{i}^{2}}\right\}, 0\right)}$$

Where is the variance of the participant *i*'s acceleration signals along each axis, m (m=1, 2, 3), in the window of length *H* starting at *t*. The value , called sigma, is the systematic noise variance which is calculated when the device is not moving.

The AI was developed and validated by Bai et al. (2016) in a sample of healthy adult women between 60- and 91-years of age. Participants wore an ActiGraph accelerometer at the right hip and performed activities of daily living (e.g., sitting, doing laundry, washing dishes) and more intense activities like brisk walking. Oxygen consumption (VO<sub>2</sub>) was measured via portable metabolic unit and METs were used as the criterion measure to which both the AI and VM were compared. MET values <1.5 were classified as sedentary,  $\geq$ 1.5 but less than 3 were classified as light PA, and values  $\geq$ 3 were classified as MVPA. Unlike metrics MAD and ENMO, the AI was not validated in children, and findings from the study by Bai et al. (2016) cannot be applied to younger populations. Therefore, a study utilizing the AI in populations of children is warranted to understand normative values of the AI metric.

#### **Comparability of Metrics**

Despite the shifted interest in developing and validating acceleration-based metrics, little work has been done on the comparability of these metrics to each other or to established countbased metrics. This is crucial in overall metric development, as this supports the basis of data harmonization. For example, surveillance PA data from the United States has primarily been collected in the activity counts metric (Luke et al., 2011) while surveillance data from Finland has used the acceleration-based metric MAD (Husu et al., 2016) and large-scale epidemiological studies in the United Kingdom use ENMO (Doherty et al., 2017). This makes harmonization problematic, and comparisons cannot be made between PA data sets on the country-wide level.

Some work that has been done regarding comparability of metrics has demonstrated inconclusive results. A study done in 2016 by Bai and colleagues comparing acceleration-based metric ENMO and count-based VM against their newly developed acceleration-based metric, the AI, added to the literature demonstrating the ability of acceleration-based metrics to outperform

count-based metrics. Bai and colleagues (2016) found that the AI was a better classifier of sedentary/light activity, but not MVPA, when compared to ENMO. Furthermore, the AI metric outperformed VM when classifying all activity intensities as demonstrated by receiver operating curve (ROC) analyses (Bai et al., 2016). Bai et al. (2016) also found that ENMO outperformed traditional activity counts when detecting MVPA, but activity counts demonstrated better classification of sedentary/light activities compared to ENMO.

Other studies have compared count- and acceleration-based metrics amongst themselves in free living and without a criterion measure. Migueles and colleagues (2019) reported significantly higher estimations of time spent in MVPA produced by VM when compared to estimations of MVPA derived from VA and the acceleration-based metric ENMO. A recent paper from Karas and colleagues (2022) compared VM to MAD, ENMO, AI, and a fourth acceleration summary metric, the Movement-Independent Movement Summary (MIMS-unit) (John et al., 2019). Using data from the Baltimore Study of Longitudinal Aging, Karas and colleagues found strong minute-level Pearson's correlations across all metric comparisons, ranging from 0.87 (MIMS vs ENMO) to 0.99 (VM vs MIMS). Both the acceleration- and countbased metrics yielded similar graphical curves when estimating minute-level patterns of daily PA (Karas et al., 2022). However, drawing conclusions on the performance of acceleration- versus count-based metrics is difficult due to the lack of studies directly comparing estimates of PA as produced by each metric.

In addition to the limited number of studies in this area, existing literature often does not comprehensively compare multiple metrics. This presents a considerable problem as researchers continue to develop and utilize new metrics without understanding the strengths and limitations of other acceleration-based metrics. Furthermore, there are no studies that provide comparability

data on these metrics in populations of children. Performance of different metrics is critical when attempting to look at patterns of PA overtime, as different data reduction techniques via different metrics may alter overall PA.

#### **Temporal Patterns of Physical Activity**

Another aspect of PA that should be considered but for which we have limited information are temporal patterns, which are defined as the chronological succession of PA over time (De Baere et al., 2015). The context and temporality of PA, particularly during the recess period, can provide useful insight into what drives or deters unstructured free play PA in children. Previous research regarding the temporality of recess PA indicate that PA is higher during the transition to and start of recess, and gradually declines as the recess period continues (McKenzie et al., 1997; Holmes, 2012). However, more specific, recent information about temporal patterns of PA is not available due to limited work in this area.

#### **Measurement of Temporal Patterns**

Accelerometry is an ideal method for capturing temporal patterns of movement behaviors because data are time-stamped and collected at a high resolution. These data can be processed using machine learning approaches like clustering or by quantifying the frequency and intensity of bouts. Activity fragmentation is an approach that is used to identify patterns of fragmented activity, and or continuous or discontinuous PA patterns via accelerometry with a focus on length and duration of active or inactive bouts (Wanigatunga et al., 2019; Palmberg et al., 2020; Tian et al., 2021). The activity fragmentation approach can be operationalized or quantified in a number of ways with some of the more common methods being identification of transition probabilities (e.g., active-to-sedentary, sedentary-to-active) (Schrack et al., 2019), number of activity fragments, and mean duration of activity fragments (Chastin et al., 2012).

Previously, activity fragmentation has been used primarily in populations of older adults to illustrate the inverse relationship between more fragmented PA and worse functional health (Wanigatunga et al., 2019; Palmberg et al., 2020; Tian et al., 2021). While activity fragmentation is a relatively new approach, and no prior research has used activity fragmentation metrics in children, activity fragmentation may manifest differently in children and prove be a useful tool for quantifying the temporal patterns of children's PA by providing a numerical representation of continous or discontinous PA. Investigation into the temporal patterns of children's recess activity, for example by using activity fragmentation metrics, can be used to highlight beneficial or detrimental patterns, as well as group differences, that can further inform PA interventions. However, because children's PA occurs in transient bouts lasting approximately 20-s (Bailey et al. 1995), important fluctuations of children's PA would be lost if PA was condensed over a 60-s epoch, as is standard practice in adults. As shorter epochs capture more accurate estimates of PA when reducing accelerometry data in children (Aadland et al., 2020), use of a 5-s window length for calculating activity fragmentation metrics (i.e., classifying bouts and transitions) may be more appropriate in children.

In populations of adolescents and adults, activity fragmentation has been calculated using traditional count-based metrics (Schrak et al., 2019; Del Pozo Cruz & Del Pozo-Cruz, 2021), as well as acceleration-based metrics ENMO (Osborn et al., 2018) and MAD (Palmberg et al., 2020), but no research on comparability of activity fragmentation calculated from different accelerometer metrics has been done. Similar to the issue with accelerometer metrics and quantification of PA, it is unknown if different metrics similarly capture the temporal nature of activity fragmentation metrics and should be investigated as popularity of acceleration-based metrics grow.

# Conclusions

In summary, the use of count-based metrics was an important contribution to the field of PA measurement via accelerometry but is outdated as between-manufacturer and between-study comparisons cannot be made. Using metrics generated from raw acceleration facilitates uniformity, transparency, and comparability between devices and studies, and is a critical step for the development of PA measurement research. Further investigation into the comparability of the acceleration metrics of interest, MAD, ENMO, and the AI, for capturing both overall intensity and temporality of PA is warranted. Accelerometer metrics in combination with activity fragmentation can provide crucial information regarding children's recess PA levels in order to develop more effective strategies to optimize the PA obtained during outdoor recess, further contributing to children's overall health and development.

#### **METHODS**

#### **Setting and Participants**

Three elementary schools in East Lansing, Michigan agreed to participate in this study. Children (N=88; age= $7.8\pm0.7$  yrs) from five classrooms per school, including seven 1<sup>st</sup> grade (n=50) and eight 2<sup>nd</sup> grade (n=38) classrooms, participated. All children enrolled in 1<sup>st</sup> or 2<sup>nd</sup> grade at the time of data collection were eligible to participate. One parent/guardian provided written informed consent and each child provided verbal and written assent on the first day of data collection.

Data collection occurred during May and June 2019, when the average temperature was 72 degrees Fahrenheit. Each school provided two, 20-minute recess periods per day with one recess in the morning and one in the afternoon for a total of 40-minutes per day of scheduled outdoor recess. Recess took place on the schoolyard, which consisted of an open grassy field, fixed equipment (e.g., slides, swings), and asphalt areas used for games like basketball or foursquare.

# **Data Collection**

Each child participated in up to four days of data collection. However, to reduce inconsistencies in the amount of data per child due to absences and to limit the hierarchical nature of the data, one recess period per child was selected using a random number generator for inclusion in the analyses.

Children wore a triaxial accelerometer (ActiGraph, LLC, Pensacola, FL) on an elastic belt at the right hip, the most commonly used accelerometer wear location in this age group (Migueles et al. 2017). The accelerometer was an ActiGraph wGT3X-BT (firmware v1.9.2; n = 50), GT3X+ (firmware v3.2.1; n = 24), wGT3X+ (firmware v3.2.1; n = 9), or a GT9X Link

(v1.7.2; n = 5). Multiple generations of ActiGraphs were used due to limited availability of accelerometers at the time of the study. The GT3X+ and wGT3X+ devices have a slightly different dynamic range ( $\pm 6g$ ) and internal processing steps compared to the wGT3X-BT and GT9X (dynamic range of  $\pm 8g$ ). Clevenger et al. (2020) demonstrated that, despite these small differences, data are comparable across multiple generations of ActiGraph devices. Furthermore, the goal of the current study was to compare metric outcomes within a participant from the same device, so differences in accelerometer models should not affect the findings. Accelerometers were initialized to record raw acceleration data at 30 Hz with the same start time using ActiLife software (version 2.0.0). After each day, accelerometers were returned to a study member and data were downloaded using the ActiLife software and stored on a computer in a protected location on Michigan State University campus as raw acceleration and activity counts per 5-s

# **Data Processing**

Only data from the selected recess periods were included in the present analysis. Recess start/end times were identified using the schedule provided by each school and video recordings of the schoolyard. The start of each recess period was determined by the first child in the camera angle that set foot outside the school building and onto the schoolyard. Similarly, the end of recess was determined by the last child in the camera angle that set foot inside the school building. Attendance logs completed by research staff were used to determine each child's presence/absence during each recess period.

#### **Physical Activity Intensity**

Count and raw data from the accelerometers were loaded into RStudio (version 1.3.1056) as ".csv" files using the "AGread" package (version 1.1.1) (Hibbing, 2018), which was also used to calculate ENMO (Hildebrand et al., 2014). MAD was calculated as the variability in the

triaxial acceleration about the mean (Aittasalo et al., 2015). AI (Bai et al., 2016) was calculated the variance of the acceleration value along the three axes (Bai et al., 2016) using the package "SummarizedActigraphy" (version 0.5.0). All five accelerometer metrics (VM, VA, MAD, ENMO, and AI) were calculated over a 5-s epoch and are continuous variables wherein higher values indicate higher intensity activity. For each metric, all 5-s epoch data across all participants was used to identify quartiles. Each 5-s epoch was then assigned a value of 1 through 4 based on these quartile thresholds as a proxy for activity intensity. For each participant, the average of each metric (e.g., mean VM) and time spent in each quartile according to each metric was calculated over the participant's selected recess period. Finally, each child was ranked from most to least active using the mean of each of the five metrics and these rankings were used to further classify children as least active, below average, above average, or most active.

#### **Temporal Patterns**

For each participant, five fragmentation indices were calculated for each metric using modified code from the "GGIR" package (version 2.6) (van Hees et al., 2022). The number of active fragments were calculated by segmenting epoch-level data into active (quartiles 2 to 4) or inactive (quartile 1). Number of high activity fragments were calculated similarly but were defined as epochs classified as quartiles 3 to 4. Mean duration of high activity fragments were calculated in seconds. Inactivity-to-physical-activity transition probability, which represents the likelihood of switching from inactivity to activity, was calculated as 1 divided by the mean duration of inactive fragments. Physical-activity transition probability was calculated as the reciprocal of the mean duration of activity fragments.

#### **Statistical Analyses**

Pearson's (r) correlations were used to assess the strength of the association between

metrics at the epoch-level and were interpreted as poor (r=0.20 to 0.30), fair (r=0.30 to 0.50), moderate (r=0.60 to 0.70), strong (r=0.80 to 0.90), and perfect (r=1.0) (Chan, 2003). Weighted kappa was () was calculated using the "irr" package (version 0.84.1) (Gamer, Lemon, & Singh, 2012) to assess agreement between metrics in the assigned quartile of each epoch (i.e., 1-4) and interpreted as no agreement (=0 to 0.20), minimal (=0.21 to 0.39), weak (=0.40 to 0.59), moderate (=0.69 to 0.79), strong (=0.80 to 0.90), and almost perfect (=>0.90) (McHugh, 2012). Weighted kappa is appropriate because it accounts for ordering of the categories (Cohen, 1968). For example, a misclassification of a quartile 4 epoch as quartile 1 is more greatly penalized than a quartile 4 epoch misclassified as quartile 3.

Once collapsed to the participant level, Pearson's *r* correlation coefficients were used to compare the mean of each metric, time spent in each quartile, and the five fragmentation indices between metrics. Two one-sided tests of equivalence were performed using the "TOSTER" package (version 0.4.0) (Lakens, 2017) and assessed equivalence in percent of time spent in each quartile and the fragmentation indices between the five metrics. Equivalence testing is more appropriate than traditional hypothesis testing when a meaningful difference is not expected between comparison group averages (Dixon et al., 2018). Equivalence bounds were determined as five percent of the mean of each metric. Spearman's rho was used to the assess associations of each child's ranking based on the mean of each metric. Confusion matrices were created using the package "caret" (version 6.0-86) (Kuhn, 2008) to assess agreement between metrics when categorizing children as least active, below average, above average, or most active.

#### RESULTS

#### **Participant characteristics**

The final study sample consisted of 50 first graders and 38 second graders across all 3 schools (N=88) with a mean age of 7.8 years (SD=0.7). School 1 had 52 children participate; School 2 had 16 participants; and School 3 had 20 children participate. Three children were excluded from the present analysis (one child dropped out due to behavioral issues, two children did not wear the accelerometer belt). Overall, there were more female participants (74%; n=65) than males (26%; n=23). Children included in the final analyses had an average of 25.9 3.0 minutes of data for the randomly selected recess period.

# Aim 1

Physical activity intensity

Mean values per 5-s epoch and per participant for the five accelerometer metrics are reported in Table 1. At the epoch-level, correlation coefficients for the association between metrics ranged from moderately high (r=0.69, VM vs ENMO) to high (r=0.98, AI vs ENMO, AI vs MAD, MAD vs ENMO) (Table 2). Weighted kappa indicated that agreement between each 5-s epoch ranged from strong to almost perfect, with the weakest agreement between VM and ENMO (= 0.81) and the strongest agreement between AI and MAD (= 0.96) (see Table 2).

Metric	Epoch-level	Participant-level
VM (counts/5s)	305.3 327.6	308.4 146.9
VA (counts/5s)	154.4 224.9	156.5 89.7
ENMO (mg)	135.8 188.4	137.8 77.9
MAD (mg)	192.7 241.9	195.4 104.1
AI	1.1 1.2	1.1 0.5

Table 1. Means SD for metrics at the epoch- and participant-level.

VM = vector magnitude; VA = vertical axis; ENMO = Euclidean norm minus one; MAD = mean amplitude deviation; AI = activity index Note: AI is a unit less metric.

Table 2. Pearson's (r) correlations and weighted kappa () for all metrics when compared at a 5-s epoch.

Comparison	Pearson's r	Kappa
AI vs VM	0.70	0.87
AI vs ENMO	0.98	0.93
AI vs MAD	0.98	0.96
AI vs VA	0.70	0.87
MAD vs ENMO	0.98	0.94
MAD vs VM	0.72	0.84
MAD vs VA	0.75	0.86
VM vs ENMO	0.69	0.81
VM vs VA	0.94	0.90
ENMO vs VA	0.72	0.83

AI = activity index; VM = vector magnitude; ENMO = Euclidean norm minus one; MAD = mean amplitude deviation; VA = vertical axis

Participant-level correlations between mean metrics ranged from moderately high (*r*=0.77, AI vs VM) to high (*r*=0.99, ENMO vs MAD, *r*=0.99, MAD vs AI). Participant-level

correlations between time spent in each quartile ranged from small (*r*=0.29, ENMO vs VA quartile 2) to high (*r*=0.99, MAD vs AI quartiles 1 and 4). Percent of time spent in quartile 1 was equivalent for five of the ten pairwise metric comparisons (AI vs VM, AI vs MAD, AI vs VA, VM vs MAD, VM vs VA) (Table 3). Similarly, percent of time spent in quartile 2 was equivalent for five metric comparisons (AI vs VM, AI vs MAD, AI vs VA, VA vs VM, VA vs MAD) (Table 4). All metrics were equivalent when estimating percent of time spent in quartile 3 (Table 5) while percent of time spent in quartile 4 was equivalent for all but three comparisons (VM vs AI, VM vs MAD, and VM vs ENMO) (Table 6).

Bias		<u>Equivalence Test</u>		
Comparison	Mean	SE	<b>Confidence Interval</b>	Equivalent
AI vs. VM	0.11	0.34	-0.45, 0.67	Yes
MAD vs. ENMO	-0.15	1.51	-2.67, 2.36	No
VM vs. MAD	0.15	0.50	-0.68, 0.99	Yes
ENMO vs. VM	< 0.00	1.52	-2.53, 2.54	No
VA vs. ENMO	0.15	1.55	-2.44, 2.73	No
AI vs. ENMO	0.11	1.49	-2.37, 2.58	No
MAD vs. AI	-0.04	0.22	-0.41, 0.33	Yes
VA vs. AI	0.25	0.51	-0.60, 1.10	Yes
VM vs. VA	-0.14	0.41	-0.82, 0.53	Yes
MAD vs. VA	0.30	0.62	-0.73, 1.32	No

Table 3. Equivalence between metrics in percent of time spent in quartile 1 activity level.

AI = activity index; VM = vector magnitude; MAD = mean amplitude deviation; ENMO = Euclidean norm minus one; VA = vertical axis

Confidence intervals were compared to an equivalence bound of -1.230, 1.230 (5% of the mean) to determine equivalence.

Bias			<u>Equivalence Test</u>		
Comparison	Mean	SE	Confidence Interval	Equivalent	
AI vs. VM	-0.10	0.66	-1.14, 1.04	Yes	
MAD vs. ENMO	0.24	1.39	-2.10, 2.54	No	
VM vs. MAD	-0.09	0.76	-1.35, 1.17	No	
ENMO vs. VM	0.15	1.48	-2.32, 2.62	No	
VA vs. ENMO	0.13	1.43	-2.25, 2.51	No	
AI vs. ENMO	-0.20	1.36	-2.46, 2.10	No	
MAD vs. AI	0.04	0.28	-0.42, 0.50	Yes	
VA vs. AI	-0.10	0.59	-1.05, 0.92	Yes	
VM vs. VA	0.02	0.54	-0.88, 0.91	Yes	
MAD vs. VA	-0.11	0.65	-1.19, 0.98	Yes	

Table 4. Equivalence between metrics in percent of time spent in quartile 2 activity level.

Confidence intervals were compared to an equivalence bound of -1.248, 1.248 (5% of the mean) to determine equivalence.

Bias			<u>Equivalence Test</u>		
Comparison	Mean	SE	Confidence Interval	Equivalent	
AI vs. VM	-0.03	0.52	-0.88, 0.84	Yes	
MAD vs. ENMO	-0.08	0.42	-0.78, 0.61	Yes	
VM vs. MAD	0.02	0.57	-0.95, 0.95	Yes	
ENMO vs. VM	-0.09	0.70	-1.25, 1.08	Yes	
VA vs. ENMO	-0.14	0.64	-1.20, 0.92	Yes	
AI vs. ENMO	0.07	0.47	-0.71, 0.85	Yes	
MAD vs. AI	-0.02	0.25	-0.44, 0.40	Yes	
VA vs. AI	-0.08	0.45	-0.83, 0.68	Yes	
VM vs. VA	0.06	0.52	-0.81, 0.93	Yes	
MAD vs. VA	-0.06	0.48	-0.86, 0.74	Yes	

Table 5. Equivalence between metrics in percent of time spent in quartile 3 activity level.

Confidence intervals were compared to an equivalence bound of -1.254, 1.254 (5% of the mean) to determine equivalence.

Bias			<u>Equivalence Test</u>		
Comparison	Mean	SE	Confidence Interval	Equivalent	
AI vs. VM	-0.04	0.77	-1.32, 1.24	No	
MAD vs. ENMO	< -0.00	0.26	-0.44, 0.43	Yes	
VM vs. MAD	-0.06	0.80	-1.40, 1.27	No	
ENMO vs. VM	-0.07	0.80	-1.40, 1.27	No	
VA vs. ENMO	-0.13	0.62	-1.16, 0.90	Yes	
AI vs. ENMO	0.02	0.30	-0.48, 0.53	Yes	
MAD vs. AI	0.02	0.26	-0.41, 0.45	Yes	
VA vs. AI	-0.11	0.64	-1.18. 0.96	Yes	
VM vs. VA	0.07	0.46	-0.70, 0.83	Yes	
MAD vs. VA	-0.13	0.60	-1.13, 0.87	Yes	

Table 6. Equivalence between metrics in percent of time spent in quartile 4 activity level.

Confidence intervals were compared to an equivalence bound of -1.268, 1.268 (5% of the mean) to determine equivalence.

When children were ranked from most to least active using each of the five accelerometer metrics, agreement between metrics was very strong 0.87 (Table 7). When these rankings were used to classify children into one of four groups (least active, below average, above average, most active), the highest agreement was seen for classifying the least active children, wherein 82-95% of children were categorized as least active by both the referent and comparison metric (e.g., VA classification and VM classification). There was more variability when examining the below average (50-91% concordance) and above average (50-86% concordance) groups. Finally, 64-91% of children were categorized as most active by the referent and comparison metrics in their respective comparisons (Table 8).

	VM Rank	MAD Rank	ENMO Rank	VA Rank
VM Rank	-			
MAD Rank	0.88	-		
ENMO Rank	0.87	0.99	-	
VA Rank	0.97	0.90	0.89	-
AI Rank	0.87	0.99	0.98	0.87

Table 7. Association between the ranking of children's activity level during recess by five accelerometer metrics (Spearman's rho).

VM = vector magnitude; MAD = mean amplitude deviation; ENMO = Euclidean norm minus one; VA = vertical axis; AI = activity index

Table 8. Confusion matrix showing agreement in the classification of each child's overall activity level as least active, below average activity, above average activity, and most active according to five accelerometer metrics. The metric in the left-most column served as the referent group and numbers are represented as raw number (percent) of children within each activity group according to the referent group that were also classified as that activity group by the other metric.

	VA Classification			
VM Classification	Least	Below Average	Above Average	Most
Least	19 (86)	3 (14)	0 (0)	0 (0)
Below Average	3 (14)	15 (68)	4 (18)	0 (0)
Above Average	0 (0)	4 (18)	16 (73)	2 (9)
Most	0 (0)	0 (0)	2 (9)	20 (91)
		MAD Classificat	tion	
VM Classification	Least	Below Average	Above Average	Most
Least	18 (82)	4 (18)	0 (0)	0 (0)
Below Average	4 (18)	12 (55)	4 (18)	2 (9)
Above Average	0 (0)	5 (23)	11 (50)	6 (27)
Most	0 (0)	1 (5)	7 (32)	14 (64)
	ENMO Classification			
VM Classification	Least	Below Average	Above Average	Most
Least	18 (82)	4 (18)	0 (0)	0 (0)
Below Average	4 (18)	11 (50)	5 (23)	2 (9)
Above Average	0 (0)	6 (27)	11 (50)	5 (23)
Most	0 (0)	1 (5)	6 (27)	15 (68)
		AI Classification	on	
VM Classification	Least	Below Average	Above Average	Most
Least	19 (86)	3 (14)	0 (0)	0 (0)
Below Average	3 (14)	12 (55)	6 (27)	1 (5)
Above Average	0 (0)	5 (23)	11 (50)	6 (27)
Most	0 (0)	2 (9)	5 (23)	15 (68)

	VA Classification			
ENMO Classification	Least	Below Average	Above Average	Most
Least	18 (82)	4 (18)	0 (0)	0 (0)
Below Average	4 (18)	13 (59)	4 (18)	1 (5)
Above Average	0 (0)	5 (23)	10 (45)	7 (32)
Most	0 (0)	0 (0)	8 (36)	14 (64)
		AI Classification	on	
ENMO Classification	Least	Below Average	Above Average	Most
Least	21 (95)	1 (5)	0 (0)	0 (0)
Below Average	1 (5)	19 (86)	2 (9)	0 (0)
Above Average	0 (0)	2 (9)	19 (86)	1 (5)
Most	0 (0)	1 (5)	1 (5)	21 (95)
	MAD Classification			
ENMO Classification	Least	Below Average	Above Average	Most
Least	21 (95)	1 (5)	0 (0)	0 (0)
Below Average	1 (5)	19 (86)	1 (5)	0 (0)
Above Average	0 (0)	2 (9)	18 (82)	2 (9)
Most	0 (0)	0 (0)	2 (9)	20 (91)
	AI Classification			
MAD Classification	Least	Below Average	Above Average	Most
Least	21 (95)	1 (5)	0 (0)	0 (0)
Below Average	1 (5)	20 (91)	1 (5)	0 (0)
Above Average	0 (0)	1 (5)	19 (86)	2 (9)
Most	0 (0)	0 (0)	2 (9)	20 (91)
		VA Classificati	on	
MAD Classification	Least	Below Average	Above Average	Most
Least	19 (86)	3 (14)	0 (0)	0 (0)
Below Average	3 (14)	15 (68)	4 (18)	0 (0)
Above Average	0 (0)	3 (14)	11 (50)	8 (36)
Most	0 (0)	1 (5)	7 (32)	14 (64)
		AI Classificatio	on	
VA Classification	Least	Below Average	Above Average	Most
Least	18 (82)	4 (18)	0 (0)	0 (0)
Below Average	4 (18)	13 (59)	5 (23)	0 (0)
Above Average	0 (0)	3 (14)	11 (50)	8 (36)
Most	0 (0)	2 (9)	6 (27)	14 (64)

Table 8 (cont'd)

VM = vector magnitude; VA = vertical axis; MAD = mean amplitude deviation; ENMO = Euclidean norm minus one; AI = activity index

Least = Children whose overall activity level fell into the lowest quartile Most = Children whose overall activity level fell into the highest quartile

Aim 2

Temporal characteristics

Accelerometer metrics are plotted over the randomly selected recess periods by school (i.e., school 1, 2, and 3) in Figures 1 to 3. Time in number of 5-s epochs is shown on the x-axis and metrics are plotted on the y-axis, with the right side of the y-axis displaying the unitless metric AI. Means and standard deviations for fragmentation indices are displayed in Table 9.

Figure 1. Five accelerometer metrics plotted over randomly selected recess periods for school 1.



VM = vector magnitude; VA = vertical axis; MAD = mean amplitude deviation; ENMO = Euclidean norm minus one; AI = activity index



Figure 2. Five accelerometer metrics plotted over randomly selected recess periods for school 2.

VM = vector magnitude; VA = vertical axis; MAD = mean amplitude deviation; ENMO = Euclidean norm minus one; AI = activity index



Figure 3. Five accelerometer metrics plotted over randomly selected recess periods for school 3.

VM = vector magnitude; VA = vertical axis; MAD = mean amplitude deviation; ENMO = Euclidean norm minus one; AI = activity index

Metric	Mean SD	Transition Probability Physical Activity to Inactivity	Number of High Activity Fragments (Q3-4)	Number of Activity Fragments (Q2-4)	Mean Duration of High Activity Fragments in Seconds
VM	308.4 146.9	0.36 0.10	62.30 23.38	41.65 21.34	22.61 11.79
VA	156.5 9.7	0.35 0.09	60.82 23.06	43.05 21.64	23.08 10.39
MAD	195.4 104.5	0.31 0.09	51.36 21.24	34.85 19.80	29.25 20.18
ENMO	137.8 77.9	0.30 0.11	50.91 21.42	34.18 22.01	33.03 43.95
AI	1.1 0.5	0.32 0.10	52.75 22.14	37.84 20.58	28.58 19,93

Table 9. Fragmentation indices derived from five accelerometer metrics.

VM = vector magnitude; VA = vertical axis; MAD = mean amplitude deviation; ENMO = Euclidean norm minus one; AI = activity index

The lowest physical-activity-to-inactivity transition probabilities (Figure 4) were observed with acceleration-based metrics ENMO, MAD, and AI (30, 31, 32%, respectively). The highest physical-activity-to-inactivity transition probabilities were seen with metrics VA and VM, 35 and 36%, respectively. Similar trends were observed with the inactivity-to-physicalactivity transition probability shown in Figure 4. Number of high activity fragments (Figure 5) and activity fragments (Figure 6) was highest for VM (62.30) and lowest for ENMO (50.91). Mean duration of high activity fragments according to 5-s epochs was highest for ENMO (6.61 5-s epochs) and lowest for VM (4.52 5-s epochs) (Figure 7).





VM = vector magnitude; VA = vertical axis; MAD = mean amplitude deviation; ENMO = Euclidean norm minus one; AI = activity index

Figure 5. Probability of transitioning from activity to inactivity according to five accelerometer metrics.



VM = vector magnitude; VA = vertical axis; MAD = mean amplitude deviation; ENMO = Euclidean norm minus one; AI = activity index



Figure 6. Number of high activity fragments reported according to five accelerometer metrics.

VM = vector magnitude; VA = vertical axis; MAD = mean amplitude deviation; ENMO = Euclidean norm minus one; AI = activity index



Figure 7. Number of activity fragments according to five accelerometer metrics.

VM = vector magnitude; VA = vertical axis; MAD = mean amplitude deviation; ENMO = Euclidean norm minus one; AI = activity index

Correlations for each of the fragmentation indices are displayed in Table 10. Correlations for number of high activity fragments were all high (r=0.80). For number of activity fragments, correlations were lowest for VA vs ENMO (r=0.67, moderately high), and highest for MAD vs AI (r=0.97, high). Correlations for mean duration of high activity fragments in number of 5-s epochs ranged from moderate (r=0.53, ENMO vs VM) to high (r=0.99, MAD vs AI).

Comparison	Transition Probability Physical Activity to Inactivity	Transition Probability Inactivity to Physical Activity	Number of High Activity Fragments (Q3-4)	Number of Activity Fragments (Q2-4)	Mean Duration of High Activity Fragments in Seconds
AI vs. VM	0.91	0.92	0.92	0.96	0.87
MAD vs. ENMO	0.84	0.82	0.97	0.78	0.73
VM vs. MAD	0.90	0.89	0.91	0.92	0.84
ENMO vs. VM	0.78	0.73	0.89	0.69	0.53
VA vs. ENMO	0.77	0.78	0.89	0.67	0.54
AI vs. ENMO	0.83	0.84	0.96	0.73	0.71
MAD vs. AI	0.98	0.96	0.99	0.97	0.99
VA vs. AI	0.91	0.92	0.92	0.95	0.82
VM vs. VA	0.92	0.92	0.95	0.96	0.88
MAD vs. VA	0.90	0.89	0.92	0.92	0.82

Table 10. Pearson's correlations for fragmentation indices as reported by comparison of each of the five metrics.

Equivalence of various fragmentation indices as categorized by each metric are show in Tables 11-15. Only two comparisons (MAD vs AI, VM vs VA) were equivalent when comparing transition probability of activity to inactivity. One comparison (VM vs VA) was equivalent when comparing transition probability of inactivity to activity. For number of high activity fragments, only two comparisons were equivalent (MAD vs ENMO, MAD vs AI). None of the comparisons were equivalent for number of activity fragments. MAD vs AI and MAD vs VA were the only equivalent comparisons for mean duration of high activity fragments in number of 5-s epochs.

	<u>Bias</u>		<u>Equivalence Test</u>		
Comparison	Mean	SE	<b>Confidence Interval</b>	Equivalent	
AI vs. VM	-0.04	< 0.01	-0.04, -0.03	No	
MAD vs. ENMO	-0.01	0.01	-0.02, 0.003	No	
VM vs. MAD	-0.05	< 0.01	-0.06, -0.04	No	
ENMO vs. VM	-0.06	< 0.01	-0.07, -0.05	No	
VA vs. ENMO	0.05	< 0.01	0.04, 0.06	No	
AI vs. ENMO	0.02	< 0.01	0.006, 0.03	No	
MAD vs. AI	0.01	< 0.01	0.007, 0.01	Yes	
VA vs. AI	-0.03	< 0.01	-0.04, -0.03	No	
VM vs. VA	< 0.01	< 0.01	-0.01, 0.005	Yes	
MAD vs. VA	0.05	< 0.01	0.04, 0.05	No	

Table 11. Equivalence of five metrics for determining transition probability of activity to inactivity.

AI = activity index; VM = vector magnitude; MAD = mean amplitude deviation; ENMO = Euclidean norm minus one; VA = vertical axis

Confidence intervals were compared to an equivalence bound of -0.0163, 0.0163 (5% of the mean) to determine equivalence.

<u>Bias</u>			<u>Equivalence Test</u>		
Comparison	Mean	SE	<b>Confidence Interval</b>	Equivalent	
AI vs. VM	-0.02	< 0.01	-0.03, -0.008	No	
MAD vs. ENMO	-0.02	< 0.01	-0.03, 0.002	No	
VM vs. MAD	-0.03	< 0.01	-0.04, -0.02	No	
ENMO vs. VM	-0.02	0.01	-0.03, 0.002	No	
VA vs. ENMO	0.06	0.01	0.04, 0.07	No	
AI vs. ENMO	0.03	< 0.01	0.01, 0.04	No	
MAD vs. AI	0.02	< 0.01	0.007, 0.02	No	
VA vs. AI	-0.02	< 0.01	-0.04, -0.02	No	
VM vs. VA	< 0.01	< 0.01	-0.004, 0.02	Yes	
MAD vs. VA	0.04	< 0.01	0.03, 0.05	No	

Table 12. Equivalence of five metrics for determining transition probability of inactivity to activity.

AI = activity index; VM = vector magnitude; MAD = mean amplitude deviation; ENMO = Euclidean norm minus one; VA = vertical axis

Confidence intervals were compared to an equivalence bound of -0.0176, 0.0176 (5% of the mean) to determine equivalence.

	<u>Bias</u>	<u>Equivalence Test</u>			
Comparison	Mean	SE	Confidence Interval	Equivalent	
AI vs. VM	-9.55	1.0	-11.14, -7.94	No	
MAD vs. ENMO	-0.45	0.60	-1.44, 0.53	Yes	
VM vs. MAD	-10.93	1.04	-12.67, -9.19	No	
ENMO vs. VM	-11.39	1.15	-13.29, -9.48	No	
VA vs. ENMO	9.91	1.14	8.01, 11.81	No	
AI vs. ENMO	1.84	0.65	0.75, 2.93	No	
MAD vs. AI	1.39	0.37	0.77, 2.00	Yes	
VA vs. AI	-8.07	0.97	-9.68, -6.45	No	
VM vs. VA	-1.48	0.79	-2.79, -0.17	No	
MAD vs. VA	9.45	0.99	7.81, 11.10	No	

Table 13. Equivalence of five metrics for determining number of high activity fragments.

Confidence intervals were compared to an equivalence bound of -2.781, 2.781 (5% of the mean) to determine equivalence.

	<u>Bias</u>		<u>Equivalence Test</u>			
Comparison	Mean	SE	Confidence Interval	Equivalent		
AI vs. VM	-3.81	0.62	-4.84, -2.77	No		
MAD vs. ENMO	-0.67	1.52	-3.20, 1.86	No		
VM vs. MAD	-6.80	0.90	-8.30, -5.30	No		
ENMO vs. VM	-7.47	1.83	-10.51, -4.42	No		
VA vs. ENMO	8.86	1.89	5.71, 12.01	No		
AI vs. ENMO	3.66	1.66	0.89, 6.42	No		
MAD vs. AI	2.99	0.51	2.13, 3.84	No		
VA vs. AI	-5.20	0.73	-6.42, -3.99	No		
VM vs. VA	1.39	0.69	0.24, 2.56	No		
MAD vs. VA	8.19	0.91	6.68, 9.71	No		

Table 14. Equivalence of five metrics for determining number of activity fragments.

Confidence intervals were compared to an equivalence bound of -1.916, 1.916 (5% of the mean) to determine equivalence.

Comparison	<u>Bias</u>		<u>Equivalence Test</u>			
	Mean	SE	<b>Confidence Interval</b>	Equivalent		
AI vs. VM	1.19	0.69	0.79, 1.59	No		
MAD vs. ENMO	0.75	0.69	-0.39, 1.90	No		
VM vs. MAD	1.33	0.26	0.90, 1.76	No		
ENMO vs. VM	2.10	0.83	0.70, 3.46	No		
VA vs. ENMO	-1.99	0.84	-3.38, -0.69	No		
AI vs. ENMO	-0.89	0.70	-2.06, 0.28	No		
MAD vs. AI	-0.13	0.06	-0.24, -0.03	Yes		
VA vs. AI	1.10	0.27	0.65, 1.55	No		
VM vs. VA	0.09	0.12	-0.11, 0.30	No		
MAD vs. VA	-0.13	0.06	-0.23, -0.03	Yes		

Table 15. Equivalence of five metrics for determining mean duration of high activity fragments.

AI = activity index; VM = vector magnitude; MAD = mean amplitude deviation;

ENMO = Euclidean norm minus one; VA = vertical axis

Confidence intervals were compared to an equivalence bound of -0.273, 0.273 (5% of the mean) to determine equivalence.

#### DISCUSSION

Accelerometers can capture high resolution information about the frequency, intensity, duration, and timing of children's physical activity. Disparities in how these data are summarized (i.e., which metric is used) can impede harmonization across data sources or comparability of results from different studies or surveillance systems. The present analysis provides preliminary evidence regarding the convergent validity of five count- and acceleration-based metrics at both the epoch- and participant-level for capturing both overall activity levels and temporal patterns. While metrics were often strongly associated, they were not always statistically equivalent. Continued use and development of various acceleration-based metrics is warranted to provide an open-source and device-independent alternative to count-based metrics while maintaining comparability to past and future research, as well as providing reliable estimates of time spent in activity levels across studies.

We found strong epoch-level (r=0.69-0.98) and overall (r=0.77-0.99) correlations between metrics. While there is limited prior research on the comparability of the five accelerometer metrics used in the present study, particularly in children or utilizing a 5-s epoch, the findings of the present study are comparable to studies using samples of adults, and therefore, longer epochs. Specifically, correlations in the present study were similar to or stronger than those reported by Migueles et al (2019) who compared both acceleration- and count-based metrics at the right hip during waking hours ENMO and MAD (r=0.74 vs 0.98 in the present study), VM and ENMO (r=0.48 vs 0.69), and VM and MAD (r=0.81 vs 0.72) in a sample of free-living young adults. Furthermore, a study of older adults by Karas et al (2022) found stronger correlations between activity counts and acceleration-based metrics ENMO (r=0.87), MAD (r=0.91), and AI (r=0.97) than the present study (r=0.69, 0.82, and 0.70, respectively).

While the current and past studies support positive associations between metrics across different age groups, more research is needed to clarify the strength of this association.

While all metrics were at least moderately associated, relationships were stronger amongst acceleration-based metrics (all r=0.98) and between the two count-based metrics (r=0.94), while weaker associations were found between acceleration based and count-based metrics (r=0.69-0.75). The limited prior research in this area precludes us from making final conclusions about this finding, but this indicates that comparisons between data or research using different types of metrics (acceleration- versus count-based) should be done more cautiously than comparisons when using similar types of metrics. This finding also supports the increased use of acceleration-based metrics in future studies; in addition to supporting comparability across studies because they can be calculated using any device brand, our findings indicate that these metrics capture similar trends in activity intensity even when the acceleration data are processed differently.

In addition to correlating the five metrics to each other, we compared time spent in four quartiles of activity intensity. While we did not use existing cut-points due to lack of availability for all five metrics in this population, these quartiles may serve as a proxy for sedentary, light, moderate, and vigorous intensity. For example, the most used cut-point in this age group for classifying MVPA is the Evenson cut-point of  $\geq 1003$  VA counts  $\cdot 15$ -sec<sup>-1</sup> ( $\geq 334$  counts  $\cdot 5$ -sec<sup>-1</sup>), compared to the quartile 3 cut-off of 224 counts  $\cdot 5$ -sec<sup>-1</sup>. When extrapolated to a 20-min recess period, differences between metrics in time spent in each quartile were less than 5-minutes. Whether these differences are too large in magnitude may depend on the research question, but a review of physical activity interventions reported changes of 1.2 to 2-minutes of MVPA, on average (Parrish et al., 2020). However, the equivalence tests revealed that metrics almost or

almost always captured similar amount of time spent in quartile 3 and quartile 4. Coupled with the fact that all metrics were able to similarly rank and classify children as most or least active, the ability of all five metrics to capture the highest intensity activities is promising because these are the outcomes of interest for future intervention research.

In addition to activity intensity, accelerometers are capable of capturing rich information about temporal features of physical activity but whether different accelerometer metrics similarly capture these patterns was previously unknown. The fragmentation indices used in the present study have proven useful in samples of older adults but have been underutilized in younger samples. While not directly comparable due to the focus on recess, number of activity fragments and high activity fragments in the current study were compared to prior work from Wanigatunga et al (2019), and transition probabilities in the current study were compared to prior work from Schrak et al (2019) indicating that children's recess activity was more fragmented than adult's free-living PA, which is to be expected.

Findings from the activity fragmentation metrics in the current study also corroborate findings from Bailey et al (1995), in that children exhibit highly transient activity, particularly at higher intensities. In their study, Bailey et al (1995) used direct observation and reported that the majority of high intensity activities generally lasted no longer than 15-s. The current study found that average high activity bout duration was also quite short and lasted no longer than 30-s. This supports the idea that accelerometer-measured activity fragmentation metrics are able to detect highly fragmented bouts, similar to what has been seen in previous literature. This may be an advantageous approach for further characterizing recess PA overall and by group (e.g., by sex) with the overarching goal of informing activity-promoting interventions.

All five accelerometer metrics captured similar trends in fragmentation indices, as indicated by correlation coefficients from 0.53 to 0.99. The strongest associations were found for number of high activity bouts ( $\geq$ 0.89), indicating this metric could be used comparably in future studies employing different processing techniques. Conversely, mean duration of ENMO and VM resulted in the weakest correlations between metrics ( $\geq$ 0.53), indicating this metric should be used and compared between studies more cautiously. Despite overall associations, count-based metrics demonstrated slightly higher fragmentation than acceleration-based metric and many comparisons were not equivalent between metrics. Notably, because we did not have a criterion measure, we cannot conclude whether any particular metric (or type of metric) most accurately captured the temporal patterns of children's behavior.

Despite overall associations, many metrics were not statistically equivalent for the fragmentation indices. However, future analysis is needed to better understand what equivalence bounds would be relevant or meaningful in this sample and setting, as the equivalence bounds used in the present study (5% of the mean) may have been too stringent. For example, bias for mean duration of high activity bouts ranged from 0.1 to 2.1-s, which may be an acceptable level of difference for this outcome. Those interested in applying different equivalence bounds can do so by comparing the equivalence bound of interest to the confidence intervals reported in Tables 5-8. For instance, the confidence interval for the comparison of mean duration of high activity bouts determined using AI and VM was 0.79-1.59, which would be equivalent if using an equivalence bound of  $\pm$ 5.0-s. Associating the fragmentation indices with outcomes like weight status or cardiometabolic health would elucidate clinically relevant equivalence bounds to be used in future research.

# **Strengths & Limitations**

The use of an unstructured, free play environment is a strength of the current study, as this non-laboratory setting offers greater external validity. While all five accelerometer metrics have been previously validated and compared to criterion measures (e.g., heart rate, energy expenditure), the current study is the first of its kind to compare metrics to each other in this population and setting. However, the current study is not without limitations. First, only hipworn devices were employed in the present study. Accelerations produced from the hip, and therefore, the results of the current study, may not be generalizable to wrist-worn devices in children. While this is the most common wear location (Migueles et al., 2017), wrist-worn devices are becoming increasingly popular to improve wear compliance. Lastly, only one recess period was selected per child for analyses. Each recess period was 20-minutes long and occurred during the warmer spring months in Michigan. These are factors that contribute to ideal PA conditions and may have influenced the pattern and intensity of PA. However, the primary purpose of the study was not to describe typical recess PA, but simply to compare PA and temporal patterns recorded by the five accelerometer metrics. Lastly, the use of recess periods alone is only a snapshot of the entire PA profile and is not generalizable to 24-hour PA patterns.

#### **Future Directions**

The present study provides preliminary support for the comparability of five accelerometer metrics for capturing activity intensity and temporal characteristics. However, there may be other metrics, like Monitor-Independent Movement Summary (MIMS) units, which should be included in future analyses. Further, fragmentation indices which capture other temporal characteristics that are more relevant to children could be created, such as median bout

duration as reported by Bailey et al. (1995). Finally, further work is needed to determine meaningful equivalence bounds for these temporal metrics.

#### REFERENCES

- Aadland, E., Andersen, L. B., Anderssen, S. A., Resaland, G. K., & Kvalheim, O. M. (2020). Accelerometer epoch setting is decisive for associations between physical activity and metabolic health in children. *Journal of sports sciences*, 38(3), 256–263. https://doi.org/10.1080/02640414.2019.1693320
- Aguilar-Farías, N., Brown, W. J., & Peeters, G. M. (2014). ActiGraph GT3X+ cut points for identifying sedentary behaviour in older adults in free-living environments. *Journal of science and medicine in sport*, 17(3), 293–299. https://doi.org/10.1016/j.jsams.2013.07.002
- Aittasalo, M., Vähä-Ypyä, H., Vasankari, T., Husu, P., Jussila, A.-M., & Sievänen, H. (2015). Mean amplitude deviation calculated from raw acceleration data: a novel method for classifying the intensity of adolescents' physical activity irrespective of accelerometer brand. *BMC Sports Science, Medicine and Rehabilitation*, 7(1). https://doi.org/10.1186/s13102-015-0010-0
- Albaum, E., Quinn, E., Sedaghatkish, S., Singh, P., Watkins, A., Musselman, K., & Williams, J. (2019). Accuracy of the Actigraph wGT3x-BT for step counting during inpatient spinal cord rehabilitation. Spinal Cord, 57(7), 571–578. https://doi.org/10.1038/s41393-019-0254-8 (Albaum et al., 2019)
- Ainsworth, B. E., Haskell, W. L., Herrmann, S. D., Meckes, N., Bassett, D. R., Jr, Tudor-Locke, C., Greer, J. L., Vezina, J., Whitt-Glover, M. C., & Leon, A. S. (2011). 2011 Compendium of Physical Activities: a second update of codes and MET values. *Medicine* and science in sports and exercise, 43(8), 1575–1581. https://doi.org/10.1249/MSS.0b013e31821ece12
- Arundell, L., Hinkley, T., Veitch, J., & Salmon, J. (2015). Contribution of the After-School Period to Children's Daily Participation in Physical Activity and Sedentary Behaviours. *PloS one*, 10(10), e0140132. https://doi.org/10.1371/journal.pone.0140132
- Bai, J., Di, C., Xiao, L., Evenson, K. R., LaCroix, A. Z., Crainiceanu, C. M., & Buchner, D. M. (2016). An activity index for raw accelerometry data and its comparison with other activity metrics. *PLoS ONE*, *11*(8). https://doi.org/10.1371/journal.pone.0160644
- Bailey, RC, Olson, J., Pepper, SL, Porszasz, J., Barstow, TJ, & Cooper, DM. (1995). The level and tempo of children's physical activities: an observational study. *Medicine and science in sports and exercise*, 27(7), 1033-1041. http://dx.doi.org/10.1249/00005768-199507000-00012 Retrieved from https://escholarship.org/uc/item/03n3p8bz
- Beighle, A., Morgan, C. F., le Masurier, G., & Pangrazi, R. P. (2006). Children's Physical Activity During Recess and Outside of School. In *Journal of School Health* (Vol. 76, Issue Ó). American School Health Association.

- Berman, N., Bailey, R., Barstow, T. J., & Cooper, D. M. (1998). Spectral and bout detection analysis of physical activity patterns in healthy, prepubertal boys and girls. *American journal of human biology : the official journal of the Human Biology Council*, 10(3), 289–297. https://doi.org/10.1002/(SICI)1520-6300(1998)10:3<289::AID-AJHB4>3.0.CO;2-E
- Biddle, S. J. H., Gorely, T., Pearson, N., & Bull, F. C. (2011). An assessment of self-reported physical activity instruments in young people for population surveillance: Project ALPHA. *International Journal of Behavioral Nutrition and Physical Activity*, 8. https://doi.org/10.1186/1479-5868-8-1
- Bjorklund, D. F., & Brown, R. D. (1998). Physical play and cognitive development: integrating activity, cognition, and education. *Child development*, *69*(3), 604–606.
- Blaes, A., Ridgers, N. D., Aucouturier, J., Van Praagh, E., Berthoin, S., & Baquet, G. (2013). Effects of a playground marking intervention on school recess physical activity in French children. *Preventive medicine*, 57(5), 580–584. https://doi.org/10.1016/j.ypmed.2013.07.019
- Bornstein, D. B., Beets, M. W., Byun, W., Welk, G., Bottai, M., Dowda, M., & Pate, R. (2011). Equating accelerometer estimates of moderate-to-vigorous physical activity: In search of the Rosetta Stone. *Journal of Science and Medicine in Sport*, 14(5), 404–410. https://doi.org/10.1016/j.jsams.2011.03.013
- Brown, W. H., Pfeiffer, K. A., McIver, K. L., Dowda, M., Addy, C. L., & Pate, R. R. (2009). Social and environmental factors associated with preschoolers' nonsedentary physical activity. *Child development*, 80(1), 45–58. https://doi.org/10.1111/j.1467-8624.2008.01245.x
- Buchowski, M. S., Acra, S., Majchrzak, K. M., Sun, M., & Chen, K. Y. (2004). Patterns of physical activity in free-living adults in the Southern United States. *European journal of clinical nutrition*, 58(5), 828–837. https://doi.org/10.1038/sj.ejcn.1601928
- Butte, N. F., Ekelund, U., & Westerterp, K. R. (2012). Assessing physical activity using wearable monitors: Measures of physical activity. *Medicine and Science in Sports and Exercise*, 44(SUPPL. 1). https://doi.org/10.1249/MSS.0b013e3182399c0e
- Button, B. L. G., Clark, A. F., Martin, G., Graat, M., & Gilliland, J. A. (2020). Measuring temporal differences in rural canadian children's moderate-to-vigorous physical activity. *International Journal of Environmental Research and Public Health*, 17(23), 1–14. https://doi.org/10.3390/ijerph17238734
- Carlson, J. A., Engelberg, J. K., Cain, K. L., Conway, T. L., Mignano, A. M., Bonilla, E. A., Geremia, C., & Sallis, J. F. (2015). Implementing classroom physical activity breaks: Associations with student physical activity and classroom behavior. *Preventive medicine*, 81, 67–72. https://doi.org/10.1016/j.ypmed.2015.08.006

Chan, Y. H. (2003). Biostatistics 104: correlational analysis. Singapore Med J, 44(12), 614-619.

- Chastin, S. F., Ferriolli, E., Stephens, N. A., Fearon, K. C., & Greig, C. (2012). Relationship between sedentary behaviour, physical activity, muscle quality and body composition in healthy older adults. *Age and ageing*, *41*(1), 111–114. https://doi.org/10.1093/ageing/afr075
- Chen, K. Y., & Bassett, D. R., Jr (2005). The technology of accelerometry-based activity monitors: current and future. *Medicine and science in sports and exercise*, *37*(11 Suppl), S490–S500. https://doi.org/10.1249/01.mss.0000185571.49104.82
- Clevenger, K. A., Grady, S. C., Erickson, K., & Pfeiffer, K. A. (2020). Use of a spatiotemporal approach for understanding preschoolers' playground activity. *Spatial and Spatio-Temporal Epidemiology*, *35*. https://doi.org/10.1016/j.sste.2020.100376
- Clevenger, K. A., McKee, K. L., & Pfeiffer, K. A. (2021). Classroom Location, Activity Type, and Physical Activity During Preschool Children's Indoor Free-Play. *Early Childhood Education Journal*. https://doi.org/10.1007/s10643-021-01164-7
- Clevenger, K. A., Moore, R. W., Suton, D., Montoye, A. H. K., Trost, S. G., & Pfeiffer, K. A. (2018). Accelerometer responsiveness to change between structured and unstructured physical activity in children and adolescents. *Measurement in Physical Education and Exercise Science*, 22(3), 224–230. https://doi.org/10.1080/1091367X.2017.1419956
- Clevenger, K. A., Pfeiffer, K. A., & , A. H. K. (2020a). Cross-generational comparability of hipand wrist-worn ActiGraph GT3X+, wGT3X-BT, and GT9X accelerometers during freeliving in adults. *Journal of Sports Sciences*, 38(24), 2794–2802. https://doi.org/10.1080/02640414.2020.1801320
- Clevenger, K. A., Pfeiffer, K. A., & Montoye, A. H. K. (2020b). Cross-Generational Comparability of Raw and Count-Based Metrics from ActiGraph GT9X and wGT3X-BT Accelerometers during Free-Living in Youth. *Measurement in Physical Education and Exercise Science*, 24(3), 194–204. https://doi.org/10.1080/1091367X.2020.1773827
- Cohen J. (1968). Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit. *Psychological bulletin*, 70(4), 213–220. https://doi.org/10.1037/h0026256
- Cohen, J. (1988). Statistical Power Analysis for the Behavioral Sciences (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.
- Corder, K., van Sluijs, E. M., Wright, A., Whincup, P., Wareham, N. J., & Ekelund, U. (2009). Is it possible to assess free-living physical activity and energy expenditure in young people by self-report?. *The American journal of clinical nutrition*, 89(3), 862–870. https://doi.org/10.3945/ajcn.2008.26739

- Crotti, M., Foweather, L., Rudd, J. R., Hurter, L., Schwarz, S., & Boddy, L. M. (2020). Development of raw acceleration cut-points for wrist and hip accelerometers to assess sedentary behaviour and physical activity in 5–7-year-old children. *Journal of Sports Sciences*, 38(9), 1036–1045. https://doi.org/10.1080/02640414.2020.1740469
- de Baere, S., Lefevre, J., de Martelaer, K., Philippaerts, R., & Seghers, J. (2015). Temporal patterns of physical activity and sedentary behavior in 10-14 year-old children on weekdays. *BMC Public Health*, *15*(1). https://doi.org/10.1186/s12889-015-2093-7
- Del Pozo Cruz, B., & Del Pozo-Cruz, J. (2021). Associations between activity fragmentation and subjective memory complaints in middle-aged and older adults. *Experimental gerontology*, *148*, 111288. https://doi.org/10.1016/j.exger.2021.111288
- Dixon, P. M., Saint-Maurice, P. F., Kim, Y., Hibbing, P., Bai, Y., & Welk, G. J. (2018). A Primer on the Use of Equivalence Testing for Evaluating Measurement Agreement. *Medicine and science in sports and exercise*, 50(4), 837–845. https://doi.org/10.1249/MSS.00000000001481
- Doherty, A., Jackson, D., Hammerla, N., Plötz, T., Olivier, P., Granat, M. H., White, T., van Hees, V. T., Trenell, M. I., Owen, C. G., Preece, S. J., Gillions, R., Sheard, S., Peakman, T., Brage, S., & Wareham, N. J. (2017). Large Scale Population Assessment of Physical Activity Using Wrist Worn Accelerometers: The UK Biobank Study. *PloS one*, *12*(2), e0169649. https://doi.org/10.1371/journal.pone.0169649
- Duncan, M. J., Wilson, S., Tallis, J., & Eyre, E. (2016). Validation of the Phillips et al. GENEActiv accelerometer wrist cut-points in children aged 5–8 years old. *European Journal of Pediatrics*, 175(12), 2019–2021. https://doi.org/10.1007/s00431-016-2795-6
- 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
- Fairclough, S. J., Noonan, R., Rowlands, A. V., Van Hees, V., Knowles, Z., & Boddy, L. M. (2016). Wear Compliance and Activity in Children Wearing Wrist- and Hip-Mounted Accelerometers. *Medicine and science in sports and exercise*, 48(2), 245–253. https://doi.org/10.1249/MSS.000000000000771
- Farrahi, V., Niemelä, M., Kangas, M., Korpelainen, R., & Jämsä, T. (2019). Calibration and validation of accelerometer-based activity monitors: A systematic review of machinelearning approaches. In *Gait and Posture* (Vol. 68, pp. 285–299). Elsevier B.V. https://doi.org/10.1016/j.gaitpost.2018.12.003
- Freedson, P., Bowles, H. R., Troiano, R., & Haskell, W. (2012). Assessment of physical activity using wearable monitors: Recommendations for monitor calibration and use in the field.

*Medicine and Science in Sports and Exercise*, *44*(SUPPL. 1). https://doi.org/10.1249/MSS.0b013e3182399b7e

- Freedson, P. S., Melanson, E., & Sirard, J. (1998). Calibration of the Computer Science and Applications, Inc. accelerometer. Medicine and science in sports and exercise, 30(5), 777–781. https://doi.org/10.1097/00005768-199805000-00021
- Freedson, P., Pober, D., & Janz, K. F. (2005). Calibration of accelerometer output for children. Medicine and Science in Sports and Exercise, 37(11 SUPPL.). https://doi.org/10.1249/01.mss.0000185658.28284.ba
- Gamer, M., & Lemon, J., & Singh, I. (2010). irr: Various Coefficients of Interrater Reliability and Agreement.
- Gardner F. (2000). Methodological issues in the direct observation of parent-child interaction: do observational findings reflect the natural behavior of participants?. *Clinical child and family psychology review*, *3*(3), 185–198. https://doi.org/10.1023/a:1009503409699
- Ginsburg, K. R., American Academy of Pediatrics Committee on Communications, & American Academy of Pediatrics Committee on Psychosocial Aspects of Child and Family Health (2007). The importance of play in promoting healthy child development and maintaining strong parent-child bonds. *Pediatrics*, *119*(1), 182–191. https://doi.org/10.1542/peds.2006-2697
- Goran, M. I., Gower, B. A., Nagy, T. R., & Johnson, R. K. (1998). Developmental changes in energy expenditure and physical activity in children: evidence for a decline in physical activity in girls before puberty. *Pediatrics*, 101(5), 887–891. https://doi.org/10.1542/peds.101.5.887
- Groffik, D., Fromel, K., & Badura, P. (2020). Composition of weekly physical activity in adolescents by level of physical activity. *BMC Public Health*, 20(1). https://doi.org/10.1186/s12889-020-08711-8
- Grydeland, M., Hansen, B. H., Ried-Larsen, M., Kolle, E., & Anderssen, S. A. (2014).
   Comparison of three generations of ActiGraph activity monitors under free-living conditions: Do they provide comparable assessments of overall physical activity in 9-year old children? *BMC Sports Science, Medicine and Rehabilitation*, 6(1).
   https://doi.org/10.1186/2052-1847-6-26
- Gubbels, J. S., Kremers, S. P., van Kann, D. H., Stafleu, A., Candel, M. J., Dagnelie, P. C., Thijs, C., & de Vries, N. K. (2011). Interaction between physical environment, social environment, and child characteristics in determining physical activity at child care. *Health psychology : official journal of the Division of Health Psychology, American Psychological Association*, 30(1), 84–90. https://doi.org/10.1037/a0021586

- Hänggi, J. M., Phillips, L. R. S., & Rowlands, A. v. (2013). Validation of the GT3X ActiGraph in children and comparison with the GT1M ActiGraph. *Journal of Science and Medicine in Sport*, 16(1), 40–44. https://doi.org/10.1016/j.jsams.2012.05.012
- HHS. (2018). Physical Activity Guidelines for Americans 2nd edition.
- Hildebrand, M., Hansen, B. H., van Hees, V. T., & Ekelund, U. (2017). Evaluation of raw acceleration sedentary thresholds in children and adults. *Scandinavian Journal of Medicine and Science in Sports*, 27(12), 1814–1823. https://doi.org/10.1111/sms.12795
- Hildebrand, M., van Hees, V. T., Hansen, B. H., & Ekelund, U. (2014). Age group comparability of raw accelerometer output from wrist-and hip-worn monitors. *Medicine and Science in Sports and Exercise*, 46(9), 1816–1824. https://doi.org/10.1249/MSS.00000000000289
- Hills, A. P., King, N. A., & Armstrong, T. P. (2007). The contribution of physical activity and sedentary behaviours to the growth and development of children and adolescents: implications for overweight and obesity. *Sports medicine (Auckland, N.Z.)*, 37(6), 533– 545. https://doi.org/10.2165/00007256-200737060-00006
- Hislop, J. F., Bulley, C., Mercer, T. H., & Reilly, J. J. (2012). Comparison of epoch and uniaxial versus triaxial accelerometers in the measurement of physical activity in preschool children: a validation study. *Pediatric exercise science*, 24(3), 450–460. https://doi.org/10.1123/pes.24.3.450
- Holmes, R. (2012). The Outdoor Recess Activities of Children at an Urban School: Longitudinal and Intraperiod Patterns. *American Journal of Play*, *4*, 327-351.
- Howe, C. A., Staudenmayer, J. W., & Freedson, P. S. (2009). Accelerometer prediction of energy expenditure: vector magnitude versus vertical axis. *Medicine and science in sports* and exercise, 41(12), 2199–2206. https://doi.org/10.1249/MSS.0b013e3181aa3a0e
- Husu, P., Suni, J., Vähä-Ypyä, H., Sievänen, H., Tokola, K., Valkeinen, H., Mäki-Opas, T., & Vasankari, T. (2016). Objectively measured sedentary behavior and physical activity in a sample of Finnish adults: a cross-sectional study. *BMC public health*, 16(1), 920. https://doi.org/10.1186/s12889-016-3591-y
- Johansson, E., Ekelund, U., Nero, H., Marcus, C., & Hagströmer, M. (2015). Calibration and cross-validation of a wrist-worn Actigraph in young preschoolers. *Pediatric obesity*, *10*(1), 1–6. https://doi.org/10.1111/j.2047-6310.2013.00213.x
- Johansson, E., Larisch, L. M., Marcus, C., & Hagströmer, M. (2016). Calibration and Validation of a Wrist- and Hip-Worn Actigraph Accelerometer in 4-Year-Old Children. *PloS one*, *11*(9), e0162436. https://doi.org/10.1371/journal.pone.0162436

- 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–S89. https://doi.org/10.1249/MSS.0b013e3182399f5e
- John, D., Tang, Q., Albinali, F., & Intille, S. (2019). An Open-Source Monitor-Independent Movement Summary for Accelerometer Data Processing. *Journal for the measurement of physical behaviour*, 2(4), 268–281. https://doi.org/10.1123/jmpb.2018-0068
- Karas, M., Muschelli, J., Leroux, A., Urbanek, J. K., Wanigatunga, A. A., Bai, J., Crainiceanu, C. M., & Schrack, J. A. (2022). Comparison of accelerometry-based measures of physical activity. *MedRxiv*, 2022.03.16.22272518. https://doi.org/10.1101/2022.03.16.22272518
- Karaca, A., Demirci, N., Yılmaz, V., Hazır Aytar, S., Can, S., & Ünver, E. (2021). Validation of the ActiGraph wGT3X-BT Accelerometer for Step Counts at Five Different Body Locations in Laboratory Settings. *Measurement in Physical Education and Exercise Science*. https://doi.org/10.1080/1091367X.2021.1948414
- Katzmarzyk, P. T., Denstel, K. D., Beals, K., Bolling, C., Wright, C., Crouter, S. E., McKenzie, T. L., Pate, R. R., Saelens, B. E., Staiano, A. E., Stanish, H. I., & Sisson, S. B. (2016). Results from the United States of America's 2016 report card on physical activity for children and youth. *Journal of Physical Activity and Health*, *13*(11), S307–S313. https://doi.org/10.1123/jpah.2016-0321
- Kim, Y., Beets, M. W., & Welk, G. J. (2012). Everything you wanted to know about selecting the "right" Actigraph accelerometer cut-points for youth, but...: A systematic review. In *Journal of Science and Medicine in Sport* (Vol. 15, Issue 4, pp. 311–321). https://doi.org/10.1016/j.jsams.2011.12.001
- Kuhn, M. (2008). Building Predictive Models in R Using the caret Package. Journal of Statistical Software, 28(5), 1 26. doi:http://dx.doi.org/10.18637/jss.v028.i05
- Lakens, D. (2017). "Equivalence tests: A practical primer for t-tests, correlations, and metaanalyses." *Social Psychological and Personality Science*, 1, 1–8. doi: <u>10.1177/1948550617697177</u>.
- Larson, T. A., Normand, M. P., Morley, A. J., & Hustyi, K. M. (2014). The Role of the Physical Environment in Promoting Physical Activity in Children Across Different Group Compositions. *Behavior Modification*, 38(6), 837–851. https://doi.org/10.1177/0145445514543466
- Latorre-Román, P. A., Martínez-Redondo, M., Salas-Sánchez, J., García-Pinillos, F., & Pérez-Jiménez, I. (2017). Suid-Afrikaanse Tydskrif vir Navorsing in Sport. South African Journal for Research in Sport, Physical Education and Recreation, 39(3), 57–66.
- Lee, I. M., Shiroma, E. J., Lobelo, F., Puska, P., Blair, S. N., Katzmarzyk, P. T., & Lancet Physical Activity Series Working Group (2012). Effect of physical inactivity on major non-communicable diseases worldwide: an analysis of burden of disease and life

expectancy. *Lancet (London, England)*, *380*(9838), 219–229. https://doi.org/10.1016/S0140-6736(12)61031-9

- Leitzmann, M. F., Park, Y., Blair, A., Ballard-Barbash, R., Mouw, T., Hollenbeck, A. R., & Schatzkin, A. (2007). *Physical Activity Recommendations and Decreased Risk of Mortality*. https://jamanetwork.com/
- Loprinzi, P. D., Cardinal, B. J., Loprinzi, K. L., & Lee, H. (2012). Benefits and environmental determinants of physical activity in children and adolescents. In *Obesity Facts* (Vol. 5, Issue 4, pp. 597–610). https://doi.org/10.1159/000342684
- Luke, A., Dugas, L. R., Durazo-Arvizu, R. A., Cao, G., & Cooper, R. S. (2011). Assessing physical activity and its relationship to cardiovascular risk factors: NHANES 2003-2006. BMC public health, 11, 387. https://doi.org/10.1186/1471-2458-11-387
- McKenzie, T. L., Sallis, J. F., Elder, J. P., Berry, C. C., Hoy, P. L., Nader, P. R., Zive, M. M., & Broyles, S. L. (1997). Physical activity levels and prompts in young children at recess: a two-year study of a bi-ethnic sample. *Research quarterly for exercise and sport*, 68(3), 195–202. https://doi.org/10.1080/02701367.1997.10607998
- McGarty, A. M., Penpraze, V., & Melville, C. A. (2016). Calibration and Cross-Validation of the ActiGraph wGT3X+ Accelerometer for the Estimation of Physical Activity Intensity in Children with Intellectual Disabilities. *PloS one*, *11*(10), e0164928. https://doi.org/10.1371/journal.pone.0164928
- McHugh M. L. (2012). Interrater reliability: the kappa statistic. *Biochemia medica*, *22*(3), 276–282.
- Meyer, U., Roth, R., Zahner, L., Gerber, M., Puder, J. J., Hebestreit, H., & Kriemler, S. (2013). Contribution of physical education to overall physical activity. *Scandinavian journal of medicine & science in sports*, 23(5), 600–606. https://doi.org/10.1111/j.1600-0838.2011.01425.x
- Montoye, A. H. K., Clevenger, K. A., Pfeiffer, K. A., Nelson, M. B., Bock, J. M., Imboden, M. T., & Kaminsky, L. A. (2020). Development of cut-points for determining activity intensity from a wrist-worn ActiGraph accelerometer in free-living adults. *Journal of Sports Sciences*, 2569–2578. https://doi.org/10.1080/02640414.2020.1794244
- Montoye, A. H. K., Nelson, M. B., Bock, J. M., Imboden, M. T., Kaminsky, L. A., MacKintosh, K. A., McNarry, M. A., & Pfeiffer, K. A. (2018). Raw and Count Data Comparability of Hip-Worn ActiGraph GT3X+ and Link Accelerometers. *Medicine and Science in Sports* and Exercise, 50(5), 1103–1112. https://doi.org/10.1249/MSS.000000000001534
- Mota, J., Silva, P., Santos, M. P., Ribeiro, J. C., Oliveira, J., & Duarte, J. A. (2005). Physical activity and school recess time: Differences between the sexes and the relationship

between children's playground physical activity and habitual physical activity. *Journal of Sports Sciences*, *23*(3), 269–275. https://doi.org/10.1080/02640410410001730124

- Neishabouri A, Nguyen J, Samuelsson J, et al. (2022) Quantification of Acceleration as Activity Counts in ActiGraph Wearables. Research Square. DOI: 10.21203/rs.3.rs-1370418/v1
- Nettlefold, L., McKay, H. A., Warburton, D. E. R., McGuire, K. A., Bredin, S. S. D., & Naylor, P. J. (2011). The challenge of low physical activity during the school day: At recess, lunch and in physical education. *British Journal of Sports Medicine*, 45(10), 813–819. https://doi.org/10.1136/bjsm.2009.068072
- Nyström, C., Pomeroy, J., Henriksson, P., Forsum, E., Ortega, F. B., Maddison, R., Migueles, J. H., & Löf, M. (2017). Evaluation of the wrist-worn ActiGraph wGT3x-BT for estimating activity energy expenditure in preschool children. *European journal of clinical nutrition*, 71(10), 1212–1217. https://doi.org/10.1038/ejcn.2017.114
- Osborn, W., Simm, P., Olds, T., Lycett, K., Mensah, F. K., Muller, J., Fraysse, F., Ismail, N., Vlok, J., Burgner, D., Carlin, J. B., Edwards, B., Dwyer, T., Azzopardi, P., Ranganathan, S., & Wake, M. (2018). Bone health, activity and sedentariness at age 11-12 years: Crosssectional Australian population-derived study. *Bone*, *112*, 153–160. https://doi.org/10.1016/j.bone.2018.04.011
- Palmberg, L., Rantalainen, T., Rantakokko, M., Karavirta, L., Siltanen, S., Skantz, H.,
  Saajanaho, M., Portegijs, E., & Rantanen, T. (2020). The Associations of Activity
  Fragmentation With Physical and Mental Fatigability Among Community-Dwelling 75-,
  80-, and 85-Year-Old People. *The journals of gerontology. Series A, Biological sciences and medical sciences*, 75(9), e103–e110. https://doi.org/10.1093/gerona/glaa166
- Parrish, A. M., Chong, K. H., Moriarty, A. L., Batterham, M., & Ridgers, N. D. (2020). Interventions to Change School Recess Activity Levels in Children and Adolescents: A Systematic Review and Meta-Analysis. *Sports medicine (Auckland, N.Z.)*, 50(12), 2145– 2173. https://doi.org/10.1007/s40279-020-01347-z
- Pate, R. R., Dowda, M., Brown, W. H., Mitchell, J., & Addy, C. (2013). Physical activity in preschool children with the transition to outdoors. *Journal of physical activity & health*, 10(2), 170–175. https://doi.org/10.1123/jpah.10.2.170
- Pellegrini, A. D., & Smith, P. K. (1998). Physical activity play: the nature and function of a neglected aspect of playing. *Child development*, 69(3), 577–598.
- Peterson, N. E., Sirard, J. R., Kulbok, P. A., DeBoer, M. D., & Erickson, J. M. (2015). Validation of Accelerometer Thresholds and Inclinometry for Measurement of Sedentary Behavior in Young Adult University Students. *Research in nursing & health*, 38(6), 492– 499. https://doi.org/10.1002/nur.21694

- Pfeiffer, K. A., McIver, K. L., Dowda, M., Almeida, M. J. C. A., & Pate, R. R. (2006). Validation and calibration of the actical accelerometer in preschool children. *Medicine* and Science in Sports and Exercise, 38(1), 152–157. https://doi.org/10.1249/01.mss.0000183219.44127.e7
- Pope, Z. C., Huang, C., Stodden, D., McDonough, D. J., & Gao, Z. (2020). Effect of children's weight status on physical activity and sedentary behavior during physical education, recess, and after school. *Journal of Clinical Medicine*, 9(8), 1–10. https://doi.org/10.3390/jcm9082651
- Puyau, M. R., Adolph, A. L., Vohra, F. A., & Butte, N. F. (2002). Validation and calibration of physical activity monitors in children. *Obesity research*, 10(3), 150–157. https://doi.org/10.1038/oby.2002.24
- Rachele, J. N., McPhail, S. M., Washington, T. L., & Cuddihy, T. F. (2012). Practical physical activity measurement in youth: a review of contemporary approaches. *World journal of pediatrics : WJP*, 8(3), 207–216. https://doi.org/10.1007/s12519-012-0359-z
- Rastogi, T., Backes, A., Schmitz, S. *et al.* (2020). Advanced analytical methods to assess physical activity behaviour using accelerometer raw time series data: a protocol for a scoping review. *Syst Rev* **9**, 259 https://doi.org/10.1186/s13643-020-01515-2
- Ridgers, N. D., Stratton, G., & Fairclough, S. J. (2005). Assessing physical activity during recess using accelerometry. *Preventive Medicine*, 41(1), 102–107. https://doi.org/10.1016/j.ypmed.2004.10.023
- Ridgers, N. D., Stratton, G., Fairclough, S. J., & Twisk, J. W. (2007). Long-term effects of a playground markings and physical structures on children's recess physical activity levels. *Preventive medicine*, 44(5), 393–397.https://doi.org/10.1016/j.ypmed.2007.01.009
- Ridgers, N. D., Fairclough, S. J., & Stratton, G. (2010). Variables associated with children's physical activity levels during recess: the A-CLASS project. *The international journal of behavioral nutrition and physical activity*, 7, 74. https://doi.org/10.1186/1479-5868-7-74
- Ridgers, N. D., Timperio, A., Cerin, E., & Salmon, J. (2014). Compensation of physical activity and sedentary time in primary school children. *Medicine and Science in Sports and Exercise*, 46(8), 1564–1569. https://doi.org/10.1249/MSS.00000000000275
- Ridgers, N. D., Timperio, A., Crawford, D., & Salmon, J. (2012). Five-year changes in school recess and lunchtime and the contribution to children's daily physical activity. *British Journal of Sports Medicine*, 46(10), 741–746. https://doi.org/10.1136/bjsm.2011.084921
- Romanzini, M., Petroski, E. L., Ohara, D., Dourado, A. C., & Reichert, F. F. (2014). Calibration of ActiGraph GT3X, Actical and RT3 accelerometers in adolescents. *European journal of sport science*, *14*(1), 91–99. https://doi.org/10.1080/17461391.2012.732614

- Rooney, L. (2018). Contribution of Physical Education and Recess towards the overall Physical Activity of 8-11 year old children. *Journal of Sport and Health Research*, *10*(2), 303-316.
  [8]. http://www.journalshr.com/index.php/issues/70-vol-10-n2-may-august-2018/307-rooney-l-mckee-d-2018-contribution-of-physical-education-and-recess-towards-the-overall-physical-activity-of-8-11-year-old-children-journal-of-sport-and-health-research-102303-316
- Routen, A. C., Upton, D., Edwards, M. G., & Peters, D. M. (2012). Discrepancies in accelerometer-measured physical activity in children due to cut-point non-equivalence and placement site. *Journal of sports sciences*, 30(12), 1303–1310. https://doi.org/10.1080/02640414.2012.709266
- Rowlands, A. v. (2018). Moving forward with accelerometer-assessed physical activity: Two strategies to ensure meaningful, interpretable, and comparable measures. In *Pediatric Exercise Science* (Vol. 30, Issue 4, pp. 450–456). Human Kinetics Publishers Inc. https://doi.org/10.1123/pes.2018-0201
- Rowlands, A. v., Dawkins, N. P., Maylor, B., Edwardson, C. L., Fairclough, S. J., Davies, M. J., Harrington, D. M., Khunti, K., & Yates, T. (2019). Enhancing the value of accelerometer-assessed physical activity: meaningful visual comparisons of data-driven translational accelerometer metrics. *Sports Medicine - Open*, 5(1). https://doi.org/10.1186/s40798-019-0225-9
- Rowlands, A. v., Edwardson, C. L., Davies, M. J., Khunti, K., Harrington, D. M., & Yates, T. (2018). Beyond Cut Points: Accelerometer Metrics that Capture the Physical Activity Profile. *Medicine and Science in Sports and Exercise*, 50(6), 1323–1332. https://doi.org/10.1249/MSS.00000000001561
- Rowlands, A. V., & Eston, R. G. (2007). The Measurement and Interpretation of Children's Physical Activity. *Journal of sports science & medicine*, 6(3), 270–276.
- Rowlands, A. v., Rennie, K., Kozarski, R., Stanley, R. M., Eston, R. G., Parfitt, G. C., & Olds, T. S. (2014). Children's physical activity assessed with wrist- and hip-worn accelerometers. *Medicine and Science in Sports and Exercise*, 46(12), 2308–2316. https://doi.org/10.1249/MSS.00000000000365
- Rowland T. W. (1998). The biological basis of physical activity. *Medicine and science in sports and exercise*, *30*(3), 392–399. https://doi.org/10.1097/00005768-199803000-00009
- Sallis, J. F., Prochaska, J. J., & Taylor, W. C. (2000). A review of correlates of physical activity of children and adolescents. *Medicine and science in sports and exercise*, *32*(5), 963–975. https://doi.org/10.1097/00005768-200005000-00014
- Schrack, J. A., Kuo, P. L., Wanigatunga, A. A., Di, J., Simonsick, E. M., Spira, A. P., Ferrucci, L., & Zipunnikov, V. (2019). Active-to-Sedentary Behavior Transitions, Fatigability, and

Physical Functioning in Older Adults. *Journals of Gerontology - Series A Biological Sciences and Medical Sciences*, 74(4), 560–567. https://doi.org/10.1093/gerona/gly243

- Stenholm, S., Pulakka, A., Leskinen, T., Pentti, J., Heinonen, O. J., Koster, A., & Vahtera, J. (2021). Daily Physical Activity Patterns and Their Association With Health-Related Physical Fitness Among Aging Workers—The Finnish Retirement and Aging Study. *The Journals of Gerontology: Series A*, 76(7), 1242–1250. https://doi.org/10.1093/gerona/glaa193
- Strath, S. J., Pfeiffer, K. A., & Whitt-Glover, M. C. (2012). Accelerometer use with children, older adults, and adults with functional limitations. *Medicine and Science in Sports and Exercise*, 44(SUPPL. 1). https://doi.org/10.1249/MSS.0b013e3182399eb1
- Stratton, G., & Mullan, E. (2005). The effect of multicolor playground markings on children's physical activity level during recess. *Preventive medicine*, 41(5-6), 828–833. https://doi.org/10.1016/j.ypmed.2005.07.009
- Sugiyama, T., Leslie, E., Giles-Corti, B., & Owen, N. (2009). Physical activity for recreation or exercise on neighbourhood streets: associations with perceived environmental attributes. *Health & place*, 15(4), 1058–1063. https://doi.org/10.1016/j.healthplace.2009.05.001
- Sylvia, L. G., Bernstein, E. E., Hubbard, J. L., Keating, L., & Anderson, E. J. (2014). Practical guide to measuring physical activity. *Journal of the Academy of Nutrition and Dietetics*, 114(2), 199–208. https://doi.org/10.1016/j.jand.2013.09.018
- Telama, R. (2009). Tracking of physical activity from childhood to adulthood: A review. In *Obesity Facts* (Vol. 2, Issue 3, pp. 187–195). https://doi.org/10.1159/000222244
- Telford, R. M., Telford, R. D., Olive, L. S., Cochrane, T., & Davey, R. (2016). Why Are Girls Less Physically Active than Boys? Findings from the LOOK Longitudinal Study. *PloS* one, 11(3), e0150041. https://doi.org/10.1371/journal.pone.0150041
- Troiano, R. P. (2005). A timely meeting: Objective measurement of physical activity. Medicine and Science in Sports and Exercise, 37(11 SUPPL.). https://doi.org/10.1249/01.mss.0000185473.32846.c3
- Troiano, R. P., Berrigan, D., Dodd, K. W., Mâsse, L. C., Tilert, T., & Mcdowell, M. (2008). Physical activity in the United States measured by accelerometer. *Medicine and Science in Sports and Exercise*, 40(1), 181–188. https://doi.org/10.1249/mss.0b013e31815a51b3
- Trost, S. G., Fees, B. S., Haar, S. J., Murray, A. D., & Crowe, L. K. (2012). Identification and validity of accelerometer cut-points for toddlers. *Obesity*, 20(11), 2317–2319. https://doi.org/10.1038/oby.2011.364

- 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., & O'Neil, M. (2014). Clinical use of objective measures of physical activity. In British Journal of Sports Medicine (Vol. 48, Issue 3, pp. 178–181). https://doi.org/10.1136/bjsports-2013-093173
- Trost, S. G., Pate, R. R., Sallis, J. F., Freedson, P. S., Taylor, W. C., Dowda, M., & Sirard, J. (2002). Age and gender differences in objectively measured physical activity in youth. In *Med. Sci. Sports Exerc* (Vol. 34, Issue 2). http://www.acsm-msse.org
- Vähä-Ypyä, H., Vasankari, T., Husu, P., Suni, J., & Sievänen, H. (2015). A universal, accurate intensity-based classification of different physical activities using raw data of accelerometer. *Clinical physiology and functional imaging*, 35(1), 64–70. https://doi.org/10.1111/cpf.12127
- Vähä-Ypyä, H., Vasankari, T., Husu, P., Mänttäri, A., Vuorimaa, T., Suni, J., & Sievänen, H. (2015). Validation of cut-points for evaluating the intensity of physical activity with accelerometry-based Mean Amplitude Deviation (MAD). *PLoS ONE*, 10(8). https://doi.org/10.1371/journal.pone.0134813
- van Hees V, Fang Z, Zhao J, Heywood J, Mirkes E, Sabia S, Migueles J (2022). *GGIR: Raw Accelerometer Data Analysis*. doi:10.5281/zenodo.1051064, R package version 2.7-1, https://CRAN.R-project.org/package=GGIR.
- van Hees, V. T., Fang, Z., Langford, J., Assah, F., Mohammad, A., M da Silva, I. C., Trenell, M. I., White, T., Wareham, N. J., Brage, S., Hees, van V., & Silva, da I. (2014). Autocalibration of accelerometer data for free-living physical activity assessment using local gravity and temperature: an evaluation on four continents. *J Appl Physiol*, *117*, 738–744. https://doi.org/10.1152/japplphysiol.00421.2014.-Wearable
- van Hees, V. T., Gorzelniak, L., Dean León, E. C., Eder, M., Pias, M., Taherian, S., Ekelund, U., Renström, F., Franks, P. W., Horsch, A., & Brage, S. (2013). Separating Movement and Gravity Components in an Acceleration Signal and Implications for the Assessment of Human Daily Physical Activity. *PLoS ONE*, 8(4). https://doi.org/10.1371/journal.pone.0061691
- Vanderloo, L. M., Di Cristofaro, N. A., Proudfoot, N. A., Tucker, P., & Timmons, B. W. (2016). Comparing the Actical and ActiGraph Approach to Measuring Young Children's Physical Activity Levels and Sedentary Time. *Pediatric exercise science*, 28(1), 133–142. https://doi.org/10.1123/pes.2014-0218
- Vanderloo, L. M., Tucker, P., Johnson, A. M., & Holmes, J. D. (2013). Physical activity among preschoolers during indoor and outdoor childcare play periods. *Applied physiology*,

*nutrition, and metabolism = Physiologie appliquee, nutrition et metabolisme, 38*(11), 1173–1175. https://doi.org/10.1139/apnm-2013-0137

- Wanigatunga, A. A., Di, J., Zipunnikov, V., Urbanek, J. K., Kuo, P. L., Simonsick, E. M., Ferrucci, L., & Schrack, J. A. (2019). Association of Total Daily Physical Activity and Fragmented Physical Activity with Mortality in Older Adults. *JAMA Network Open*, 2(10). https://doi.org/10.1001/jamanetworkopen.2019.12352
- Wanigatunga, A. A., Ferrucci, L., & Schrack, J. A. (2019). Physical activity fragmentation as a potential phenotype of accelerated aging. In *Oncotarget* (Vol. 10, Issue 8). www.oncotarget.com
- Wanigatunga, A. A., Simonsick, E. M., Zipunnikov, V., Spira, A. P., Studenski, S., Ferrucci, L., & Schrack, J. A. (2018). Perceived Fatigability and Objective Physical Activity in Midto Late-Life. *The journals of gerontology. Series A, Biological Sciences and medical sciences*, 73(5), 630–635. https://doi.org/10.1093/gerona/glx181
- Welk, G. J., Corbin, C. B., & Dale, D. (2000). Measurement issues in the assessment of physical activity in children. *Research quarterly for exercise and sport*, 71(2 Suppl), S59–S73.
- Welk, G. J., McClain, J., & Ainsworth, B. E. (2012). Protocols for evaluating equivalency of accelerometry-based activity monitors. *Medicine and Science in Sports and Exercise*, 44(SUPPL. 1). https://doi.org/10.1249/MSS.0b013e3182399d8f
- Yu, H., Kulinna, P. H., & Mulhearn, S. C. (2021). The Effectiveness of Equipment Provisions on Rural Middle School Students' Physical Activity During Lunch Recess. *Journal of physical activity & health*, 18(3), 287–295. https://doi.org/10.1123/jpah.2019-0661