

CHOICE OF CRITERION MEASURE MATTERS FOR
VALIDATING ESTIMATES OF YOUTH SEDENTARY
BEHAVIOR: DIRECT OBSERVATION VERSUS INDIRECT
CALORIMETRY VERSUS BOTH

A Thesis Presented for the
Master of Science
Degree
The University of Tennessee, Knoxville

Minttu Karoliina Hukka

May 2020

© 2020 Minttu Karoliina Hukka.

All rights reserved.

DEDICATION

This work is dedicated to my parents Mari and Tero Hukka. Thank you, *äiti* and *iskä*, for always believing in me and letting me to chase my dreams – even if it means that I must move across the world to go after them. Thank you for empowering me and making me strong enough to believe that no obstacle is greater than my will. Due to your constant love and support, I have not been afraid of taking these big steps in my life that lead to the completion of this big milestone: my master’s thesis. *Kiitos*.

With Love,

Mimmi

ACKNOWLEDGEMENTS

Firstly, I want to thank my committee members. My advisor Dr. Scott E. Crouter, thank you for your incredible patience and support throughout this process, and the opportunity to work in your laboratory. Being Dr. Crouter's student was challenging at times but he pushed me to excel and take responsibility for my work and actions and helped me to grow so much during the process.

Next, I want to thank Dr. Eugene Fitzhugh who was one of the first professors I got to work with at The University of Tennessee. Dr. Fitzhugh was in an important role in the process of completion of this degree. Dr. Fitzhugh taught me so much about how to be a better researcher and scholar and challenged me to think more critically and search for different solutions. The knowledge I attained from Dr. Fitzhugh's classes were the cornerstones of my thesis.

Dr. Dawn Podulka Coe, I want to thank you for providing both academic and emotional support for always being available for questions and listening when I needed to blow off steam. The material I learned in Dr. Coe's classes served an important purpose for my thesis.

Next, I want to thank my "APL 307 crew", Samuel LaMunion and Paul Hibbing for their incredible mentorship. Thank you for your guidance and help during this process. Thank you for answering all my silly questions and everything you taught me and helped me with to make this thesis possible. You were the real "MVPs" and made the process easier. You two will do incredible things in whatever you end up doing. Best of luck!

Lastly, I would like to thank Ms. Cary Springer for statistical consultation.

ABSTRACT

BACKGROUND: Sedentary behavior (SB) is defined by an energy expenditure (EE) ≤ 1.5 METs while in a seated, lying, or reclining posture. Criterion measures for SB, direct observation [(DO); posture] and indirect calorimetry [(IC); EE], are used to validate objective estimates of SB. It is unclear whether the criterion measures provide different measurement for SB time and if this impacts validation studies. **PURPOSE:** To compare the measured SB time by DO, IC, and combined approach (DO+IC) and to examine how changing the criterion measure impacts the validity of estimated SB time in youth from accelerometer-based devices. **METHODS:** One-hundred youth (6-18 years) had their resting metabolic rate measured and completed 16 activities ranging from sedentary to vigorous intensities while wearing an ActiGraph GT9X (AG) on the right hip, and an activPAL 3C (AP) on each thigh. Criterion measures for SB time were the Cosmed K4b² IC (EE), DO (posture), and DO+IC (EE+posture). Estimated SB by the AP and AG inclinometer and AG SB thresholds [vertical axis (VA) < 25 counts/10 sec, vector magnitude (VM) < 75 counts/10 sec; VA < 100 counts/60 sec and VA < 150 counts/60 sec] were compared against each criterion measure using repeated measures ANOVAs. **RESULTS:** Final sample was 65 youth. DO SB time was significantly greater than IC and DO+IC SB time by 7.0 min and 4.5 min, respectively ($p < 0.001$). DO+IC SB time was significantly greater than IC by 2.5 min ($p = 0.014$). All estimates, except AP_{left}, VA < 150 counts/60 sec and VM < 75 counts/10 sec, were significantly different from DO SB time by 3.1 (VA < 100 counts/60 sec) to 32.8 min (AG inclinometer) ($p < 0.05$). All estimates significantly overestimated IC SB time by 3.9 min (VA < 100 counts/60 sec) to 39.9 min (AG inclinometer; $p < 0.001$). VA < 100 counts/60 sec was the only estimate not significantly different from DO+IC SB time (mean difference = 1.4 min; $p = 0.212$); all other estimates significantly overestimated SB time by 2.7 (VA < 150 counts/60 sec)

to 37.3 minutes (AG inclinometer; $p<0.05$). **CONCLUSION:** Validation studies using different criterion measures for SB should not be compared as DO, IC, and DO+IC provide different measured SB time. Standardization of criterion measures for SB time is needed to ensure consistent results going forward.

TABLE OF CONTENTS

LIST OF TABLES	ix
LIST OF FIGURES	x
ABBREVIATIONS	xii
CHAPTER I: INTRODUCTION	1
Introduction	2
Statement of problem.....	7
Statement of purpose	7
Research questions	8
Delimitations	9
Limitations	10
CHAPTER II: REVIEW OF LITERATURE	11
Introduction	12
Sedentary behavior health implications.....	13
Definition and measurement of sedentary behavior	15
1. Criterion measures.....	16
2. Subjective and objective measures.....	28
Conclusion.....	57
CHAPTER III: MANUSCRIPT	60
Introduction	61

Methodology	64
1. Participants	64
2. Procedures	65
3. Equipment	66
4. Data processing	69
5. Statistical analysis	73
6. Results	74
Discussion	77
Conclusion.....	82
Acknowledgements.....	83
REFERENCES.....	84
APPENDICES	109
APPENDIX A: TABLES	110
APPENDIX B: FIGURES	118
APPENDIX C: RECRUITMENT FLYER.....	124
APPENDIX D: INFORMED CONSENT	126
APPENDIX E: INFORMED ASSENT	130
APPENDIX F: HEALTH HISTORY QUESTIONNAIRE	133
VITA.....	136

LIST OF TABLES

Table 1. Description of the activities.	111
Table 1. Continued	112
Table 2. Descriptive characteristics of the participants.....	113
Table 3. Comparison of the measured total sedentary time by the criterion measures.....	114
Table 4. Comparison of the measured total sedentary time by DO and the estimated by the objective measures.	115
Table 5. Comparison of the measured total sedentary time by IC and the estimated by the objective measures.	116
Table 6. Comparison of the measured total sedentary time by the combined approach and the estimated by the objective measures.....	117

LIST OF FIGURES

Figure 1. Flow chart to describe the process of reaching the final sample.	119
Figure 2. Bland-Altman plots depicting error scores (measured minus estimation) against the combined approach for the time spent in SB during the two-day activity protocol using A) direct observation (DO), B) indirect calorimetry (IC), i.e. the Cosmed K4b ² , and C) DO and IC vs. combined approach (DO+IC). The solid line represents the mean and the dashed line the 95% prediction interval.	120
Figure 3. Bland-Altman plots depicting error scores (measured minus estimation) against DO for the time spent in SB during the two-day activity protocol using A) the activPAL on the left thigh (AP _{left}) , B) the activPAL on the right thigh (AP _{right}) ^l , C) vertical axis (VA) <25 counts/10 sec, D) vector magnitude (VM) <75 counts/10 sec, E) VA <100 counts/60 sec, F) VA <150 counts/60 sec, and G) the ActiGraph inclinometer function (AG _{inclinometer}). The solid line represents the mean and the dashed line the 95% prediction interval.	121
Figure 4. Bland-Altman plots depicting error scores (measured minus estimation) for the time spent in SB during the two-day activity protocol using indirect calorimetry (IC), i.e. the Cosmed K4b ² and A) the activPAL on the left thigh (AP _{left}) , B) the activPAL on the right thigh (AP _{right}) ^l , C) vertical axis (VA) <25 counts/10 sec, D) vector magnitude (VM) <75 counts/10 sec, E) VA <100 counts/60 sec, F) VA <150 counts/60 sec, and G) the ActiGraph inclinometer function (AG _{inclinometer}). The solid line represents the mean and the dashed line the 95% prediction interval.	122
Figure 5. Bland-Altman plots depicting error scores (measured minus estimation) for the time spent in SB during the two-day activity protocol using the combined approach (DO+IC) and A) the activPAL on the left thigh (AP _{left}) , B) the activPAL on the right thigh (AP _{right}) ^l , C) vertical	

axis (VA) <25 counts/10 sec, D) vector magnitude (VM) <75 counts/10 sec, E) VA <100 counts/60 sec, F) VA <150 counts/60 sec, and G) the ActiGraph inclinometer function ($AG_{inclinometer}$). The solid line represents the mean and the dashed line the 95% prediction interval. 123

ABBREVIATIONS

ADL Activities of daily living

AP activPAL accelerometer-based device

AG ActiGraph accelerometer-based device

AUC Area Under Curve

BMI Body mass index

BMR Basal metabolic rate

BxB Breath-by-breath

CARS Children Activity Rating Scale

CV Coefficient of variation

DB Douglas bag

DMC Dynamic mixing chamber

DO Direct observation

EE Energy expenditure

ENMO Euclidean norm minus one

F_EO₂ Fraction of oxygen in expired air

F_ECO₂ Fraction of carbon dioxide in expired air

IC Indirect calorimetry/calorimeter

ICC Intra-class correlation coefficient

IPAQ The International Physical Activity Questionnaire

K4b² Cosmed K4b² portable indirect calorimeter

K5 Cosmed K5 portable indirect calorimeter

LL Lower limit

LOA Limits of agreement

MAPE Mean absolute percent error

MDC Minimum detectable change

MET Metabolic equivalent

MET_y Youth-MET (measured activity VO_2 /measured RMR)

MOCA Movement Observation in Children and Adolescents, direct observation coding system

MOST Measurement of Older Adults Sedentary Time, a survey method

PA Physical activity

MPA Moderate intensity physical activity

MVPA Moderate-to-vigorous intensity physical activity

VPA Vigorous intensity physical activity

PPV Positive predictive value

Pre-PAQ Preschool-aged Physical Activity Questionnaire, a survey method

PVO₂max Power output at the measured *VO₂max*

RER Respiratory exchange ratio

RMR Resting metabolic rate

RMSE Root mean squared error

ROC Receiver Operator Characteristics

SB Sedentary behavior

SBRN Sedentary Behavior Research Network

SOFIT System for Observing Fitness Instruction Time

SOPARC System for Observing Play and Recreation in Communities, direct observation coding system

SOPLAY System for Observing Play and Leisure Activity in Youth, direct observation coding system

STR Parental screen time recreation

TV Television

UL Upper limit

VA Vertical axis

V_E Minute ventilation

VCO₂ Carbon dioxide production

VO₂ Oxygen consumption

VO_2max Maximal oxygen consumption

VM Vector magnitude

CHAPTER I: INTRODUCTION

Introduction

Increased time spent sitting and in other low energy expenditure (EE) pursuits may contribute to the higher prevalence of obesity and several other negative health outcomes in today's society (1, 2). Throughout the day, individuals are exposed to a variety of sedentary behaviors (SB) with differing motives, including more "purposeful" or "required" SB (e.g. sitting at school, work, and for transportation), and leisure time SB (e.g. watching television, playing video games, and using a computer and other devices), resulting in a high volume of sedentary time each day (3-6). Generally, all SB has similar negative health implications despite the motive behind the behavior (7). However, some SB may be associated with other health diminishing behaviors, such as an increased consumption of fattening foods and physical inactivity, which may contribute to positive energy balance and weight gain overtime. Researchers have identified SB as one of the most important independent, yet preventable risk factors for all-cause mortality (5, 8-10) and the development of cardiovascular disease (9-11) and Type 2 diabetes (5, 7, 12). These associations have found to persist even after controlling for the level of physical activity (PA) (5, 7, 13), body mass index (BMI) or weight (5, 9), and diet (10).

In 2017, The Sedentary Behavior Research Network (SBRN) proposed a consensus definition for SB with an intention to standardize the terminology and include the consideration of both the activity EE and posture to separate SB from any other behaviors that might appear sedentary but do not necessarily meet the criteria of "any waking behavior characterized by EE less than or equal to 1.5 metabolic equivalents (METs), while in a sitting, reclining or lying posture" (14). SB can be quantified by utilizing either criterion measures which are the most reliable and valid measures, or secondary measures, which only provide estimates and thus come with some degree of bias. Because no single criterion measure exists that could simultaneously

measure both the activity posture and EE, a separate criterion measure has been determined for each component of the consensus definition: direct observation (DO) is used to assess the posture, and indirect calorimetry (IC) the EE (15).

DO is the criterion measure for posture but it can also be used to provide more detailed information regarding the activity type, duration, and other variables, such as the physical and social environment (i.e. the contextual variables). Several different DO coding systems have been created and validated for these purposes. A downside of using DO is that it is time-consuming, and also heavily relies on the observer's ability to correctly classify behaviors (16, 17). Therefore, it might be less feasible in larger studies. IC is used as the criterion measure for the measurement of oxygen consumption (VO_2), which can then be used to calculate EE.

Measurement of EE by VO_2 with IC is often impractical in field studies, but portable metabolic systems such as the Cosmed K4b² and K5 (Cosmed Inc, Rome, Italy) and the Oxycon Mobile (Vyair Medical, Mettawa, IL) have become more commonly available and found to be valid methods for measuring EE when compared against the traditional Douglas Bag (DB) method (18-21).

Measuring SB with either DO or IC may result in a different amount of time spent in SB. Because DO is mainly concerned with the postural allocation, some behaviors could be misclassified as sedentary even when they might be classified as non-sedentary by EE. IC, on the other hand, does not distinguish between postures, and some behaviors may have a MET value greater than 1.5 METs even when performed in a seated or lying; or a MET value less than 1.5 in an upright position. For example, the Compendium of Physical Activities (22) gives an EE value of 1.8 METs for "sitting, fidgeting feet" and for "sitting, writing, desk work, typing"; and 1.3 METs for "standing quietly, standing in a line". Based on DO, the first two behaviors would be

classified as sedentary and the last one as non-sedentary, but if the classification was based on the EE values, the first two would be classified as non-sedentary and the last one as sedentary. Because utilizing the criterion measures in research studies is often less feasible and comes with an increased participant burden, they are used to validate objective measures, such as accelerometer-based devices for estimating SB. The potential differences between the measured SB by the criterion measures might become problematic as some objective measures are validated against DO (23-26) and others against IC (27-30). The consensus definition of SB includes the consideration of both activity EE and posture, yet no studies have utilized the combined approach (DO+IC) with both measured posture and EE to validate objective estimates of SB.

Accelerometer-based devices have become the preferred method in SB research if criterion measures are not available. Typically, the devices contain one or multiple accelerometers that collect raw acceleration data at a specific sampling frequency and range, and an inclinometer function (calculated from the raw acceleration data) that estimate the change in angle of the body part where the device is placed. The acceleration data can be used to estimate activity intensities, and the inclinometer data the time spent in different postures. The count-based estimates of SB tend to be closer to the measured by IC compared to DO as the prediction algorithms are often related to activity intensity, whereas the inclinometer-based estimates tend to perform better against DO (31-35).

ActiGraph [(AG); ActiGraph LLC, Pensacola, FL] devices are one of the most commonly used accelerometer-based devices in SB research. The AG is typically worn on the hip or the wrist, and the raw acceleration data can be uploaded from the device to the manufacturer's ActiLife software. From the raw acceleration data, "counts" and inclinometer data can be derived

and then used to estimate time spent in SB. From the counts, time spent in SB can be estimated by quantifying the total time where the counts are below a specified SB threshold. The AG inclinometer has four different outputs: seated, lying, standing, and non-wear, and SB equals the time where the output is either sitting or lying.

The AG SB thresholds have been validated in both youth and adults and against DO (36-39) and IC (28, 30, 36, 37, 39, 40), with varying outcomes. The vertical axis (VA) <100 counts/60 sec is one of the most commonly used AG SB thresholds in youth (23, 41-44) and adults (25, 38). Recently, more evidence has been found to support the use of higher SB thresholds, e.g. VA <150 counts/60 sec to obtain better estimates of SB (25, 26, 38). Studies have also suggested that using SB thresholds based on 10 sec counts, such as, VA <25 counts/10 sec and vector magnitude (VM) <75 counts/10 sec may provide closer estimates of SB in youth compared to the more traditional 60 sec thresholds (45). Crouter et al. (30) also found that the SB threshold VA <50 counts/60 sec provided better estimates of SB in adults compared to the measured values by IC than the traditionally used VA <100 counts/60 sec threshold.

A few studies have investigated the validity of the AG inclinometer function against the criterion measures (25, 26, 29, 33, 46-48), and the findings have been contradicting. Hänggi et al. (29) reported only 15% classification accuracy for lying but 94% for sitting when the AG inclinometer output was compared against DO in a study with child participants. Mitchell et al. (46) studied the validity of the AG inclinometer against DO in a sample of overweight and obese youth. The AG inclinometer overestimated the hourly sit-to-stand transitions by 282.6% with wide 95% limits of agreement. Carr et al. (26), and Peterson et al. (25) found about 70% classification accuracy for the AG inclinometer function for the estimated SB compared to the measured by DO among adult participants .

The activPAL [(AP); PAL Technologies Limited, Glasgow, Scotland, UK] is another commonly used research-grade accelerometer-based device used for quantifying SB. Unlike the AG, the AP devices are worn on the thigh, and usually attached to the skin with non-allergenic adhesive “PALStickies.” The AP measures acceleration of the limb that can be used to calculate thigh incline. This data can then be used to estimate time spent in three different postural categories: sitting/lying, standing, and walking using the AP Professional Research Edition software and the specific proprietary algorithms (49). The time spent in SB can be estimated by quantifying the total time in the seated/lying postural category.

Aminian et al. (31) found a perfect correlation ($r = 1.00$) between the AP and DO for the time spent in sitting/lying in a sample of school-aged children. A study by Dowd et al. (36) reported a perfect agreement (100%) between the AP and DO for time spent sitting, and demonstrated that the AP is a valid objective measure for the purpose of measuring SB among adolescent females. Adult studies have also demonstrated that the AP has high correlations and agreement with the DO sedentary minutes. Grant et al. (50) reported an excellent intraclass correlation coefficient [(ICC); >0.99], a mean difference ranging from 0.2 to 0.3%, and a percentage agreement ranging from 93.5 to 98.5% between the AP and DO for sitting in adults during both controlled laboratory and free-living activities. Lyden et al. (35) reported also an ICC of 0.99 and a mean difference of 0.8 min between the AP and DO for total sedentary time during a three-day free-living study in adults.

The criterion measure used during the validation process may influence the outcome of the study: the count-based SB estimates tend to be closer to the measured SB by IC because these models are often related to activity intensity, i.e. the EE (23, 27, 33). The inclinometer-based estimates tend to perform better against DO (31-35). Crouter et al. (33) compared the

estimated SB by the AP and AG inclinometer against the measured SB by DO in laboratory, and the AG and AP inclinometer and AG counts against IC under free-living conditions. The AG inclinometer had a better classification accuracy for sitting/lying against DO compared to the AP (74.6 vs. 54.7%). Under free-living conditions, the mean difference between the estimates and the measured SB by IC were not significantly different for any estimate, except AG VA <200 counts/60 sec. The AG inclinometer had the smallest mean bias (0.2 min), but the AG 10 sec thresholds (VA <25 counts/10 sec and VM <75 counts/10 sec) had the smallest mean absolute percent error [(MAPE); 75.3 and 69.2%, respectively]. The MAPEs for the 60 sec thresholds (VA <50, <100, <150, and <200 counts/60 sec) were similar to the AP but the mean differences were larger.

Statement of problem

There is discrepancy in how studies validate accelerometer-based estimates of SB. The consensus definition of SB consists of both posture (“seated, reclining, or lying”) and EE (“ ≤ 1.5 METs”), yet the estimates are typically validated using only one half of the consensus definition, i.e. either against DO for posture, or IC for EE. No studies have observed the differences between the measured SB by the different criterion measures, and how the validity of the SB estimates would be influenced by comparing them against these different criteria measured values.

Statement of purpose

The purpose of this study is to see how the measurements of SB determined by the different criterion measures (DO, IC, and DO+IC) compare, and how the validity of the

estimates of SB will be influenced if different criterion measures are used during validation. The estimated SB will be derived from the AG GT9X and AP 3C accelerometer-based devices. The AG data will be processed using four SB thresholds (VA <100 counts/60 sec, VA <150 counts/60 sec, VA <25 counts/10 sec, and VM <75 counts/10 sec). The inclinometer function from the AG and AP will also be used to estimate posture.

Research questions

Question 1. Is there a difference in the measured youth sedentary time by DO, IC, and the combined approach (DO+IC)?

Hypothesis: It is expected that the youth sedentary time measured by DO will be greater than the time measured by IC and DO+IC, but the measured time by IC will be closer to the DO+IC than DO measured youth sedentary time.

Question 2. Is there a difference in the measured youth sedentary time by DO and estimated by the AG SB thresholds (VA <100 counts/60 sec, VA <150 counts/60 sec; VA <25 counts/10 sec, and VM <75 counts/10 sec) and the inclinometer function in both the AG and AP devices?

Hypothesis: The estimated youth sedentary time by the AP and AG inclinometer are close to the measured by DO. The VA <150 counts/60 sec will result in a closer estimate of the measured youth sedentary time compared to the VA <100 counts/60 sec. The estimates from the VA <25 counts/10 sec and VM <75 counts/10 sec will be closer to the measured youth sedentary time by DO compared to the 60 sec thresholds.

Question 3. Is there a difference between the measured youth sedentary time by IC and estimated by the AG SB thresholds (VA <100 counts/60 sec, VA <150 counts/60 sec; VA <25 counts/10 sec, and VM <75 counts/10 sec) and the inclinometer function in both the AG and AP devices?

Hypothesis: The estimated youth sedentary time by the AG thresholds will be close to the measured SB time by IC. The AG and AP inclinometer function estimated youth sedentary time will be greater than the measured by IC.

Question 4. Is there a difference between the measured youth sedentary time by the combined approach (DO+IC) and estimated by the AG SB thresholds (VA <100 counts/60 sec, VA <150 counts/60 sec; VA <25 counts/10 sec, and VM <75 counts/10 sec) and the inclinometer function in both the AG and AP devices?

Hypothesis: The estimated youth sedentary time by the AP and AG inclinometer and SB thresholds will be greater than the measured SB time due to the higher stringency of the definition of SB when DO+IC is being utilized.

Delimitations

1. All the participants are youth, 6-18 years old.
2. Participants must be healthy without current or previous musculoskeletal injuries, asthma, cardiovascular conditions, medications or pregnancy that could influence participation to PAs.
3. The activity protocol is structured and performed in a controlled setting at The University of Tennessee, Knoxville Applied Physiology Laboratory.

4. Only the K4b², DO, the AG GT9X and the AP 3C accelerometer-based devices are utilized in the study, excluding other methods and devices that are available to be used for the same purposes.

5. Only the selected AG SB thresholds and the inclinometer function in both the AG and AP devices are used in the data analysis, excluding other possible ones.

Limitations

1. All the participants are youth, 6-18 years old, which may influence the generalizability of the findings.

2. The results may only be applicable to youth and this specific sample, as well as the devices used in the study.

3. This is a secondary analysis of data that was collected for a larger study that observed PA in children. Thus, changes in the methodology are not possible and the activity protocol does not include that many activities that are purely sedentary.

CHAPTER II: REVIEW OF LITERATURE

Introduction

Growing evidence supports the link between sedentary lifestyle and premature morbidity and mortality, and increased time spent sitting may partly explain the current obesity trends across different age groups (2, 51). Any waking behavior that is performed in a seated, reclining or lying posture paired with a low energy expenditure [(EE); ≤ 1.5 metabolic equivalents (METs)] is classified as sedentary behavior (SB) (14, 52). It is recognized that one can be both highly active and meet the public physical activity (PA) guidelines, but also at the same time spend a large amount of waking hours in SB (53). Therefore, SB should be studied independently, as a separate construct from PA due to its own health effects, determinants, and related biological processes (54).

Quantifying SB is important for better understanding the dose-response relationship between SB, obesity, and other negative health outcomes in both youth and adults. Valid data of SB is essential for the purposes of creating effective preventive strategies to address the previously mentioned issues (23). Measures to quantify SB can be categorized into subjective or objective, and criterion level measures. The criterion measures are regarded as the most accurate assessment tools, and other measures are compared against these “gold-standards” to assess their validity.

What makes the application of the SB consensus definition problematic is that it includes the consideration of both activity EE and posture, but no single measure is available that can simultaneously capture them both. Because of that, separate criterion measures have been used for the postural [direct observation (DO)] and EE [indirect calorimetry (IC)] components. The validity of the secondary measures may be highly influenced by the criterion measure utilized in the validation studies. Typically, studies use either the criterion measure for posture or EE to

validate the estimates of SB. Little is known how the measured SB by the criterion measures compare against one another, and how using different criterion measures influences the validity of estimates of SB. This literature review going forward will focus on reviewing validation studies of objective measures against the SB criterion measures IC and DO. Because this thesis project's primary focus is on SB and how the validity of objective measures is determined and influenced by the criterion measures, the following literature review will focus on studies done on SB only.

Sedentary behavior health implications

Every day, individuals are exposed to a variety of sedentary pursuits with differing motives, including more “purposeful” or “required” SB (e.g. sitting at school, work, and for transportation), and leisure time SB [e.g. watching television (TV), playing video games, or using a computer and other devices], resulting in a high volume of sedentary time (3-6). Generally, all SB is treated the same regardless of the motive when the health implications of prolonged sedentary time is of concern. However, some SB may also be accompanied by other health-related behaviors that further influence the metabolism and lead to a positive energy balance, such as eating fattening foods while watching TV. A positive dose-response relationship between the total time spent watching TV, and body mass index (BMI) and adiposity in youth representing different nationalities has been established (55-57).

SB has been found to greatly contribute to the increased prevalence of obesity across different age groups (2, 51, 58-62). Because it often takes several years until lifestyle-related diseases become prevalent, most studies have examined the relationship between SB and other negative health outcomes in adults. Increased time spent in SB has been associated with a higher

risk of all-cause mortality (5, 8-11, 63, 64). A metanalysis (8) examined the link between increased time spent sitting and mortality of all causes in 54 countries worldwide. In these countries, the weighted mean sitting time was 4.2 h/day, ranging from 4.2 h/day (American countries) to 6.2 h/day (Western Pacific countries). It was suggested that sitting >3 h/day might account for 3.8% of the annual all-cause mortality (approximately 433,000 deaths/year) in these 54 countries, and a 10% reduction in the daily sitting time (approximately 30 min less sitting time per day) could reduce all-cause mortality by 0.6%. Even more drastic effects could be seen if the mean sitting time was reduced by 50% (approximately 2 h less sitting per day), which would result in at least threefold fewer mortality of all causes compared to the 10% decrease in sitting time.

Prolonged bouts of SB have been associated with increased risk of Type 2 diabetes (5, 7, 12, 65) and cardiovascular disease risk factors (9-11, 63-65), even after controlling for several cofounders, such as moderate-to-vigorous intensity PA (MVPA) (5, 7, 13, 61, 62, 65), BMI or weight (5, 9), diet (10), cigarette smoking (5), sex, and age (5), or the motive behind SB (e.g. “purposeful”/required vs. leisure) (7). Researchers have become interested in studying the possible health effects of breaking up prolonged sedentary time. It is speculated that breaks from SB, also known as sit-to-stand transitions, may balance out the negative health effects related to SB by increasing muscle fiber recruitment, which produces specific cellular signals that regulate cardiovascular disease risk factors (66). It is unclear whether similar effects exist among children and adolescents (67-70). A study (71) found evidence to suggest that decreasing the total weekly sedentary time among children with low MVPA participation may have more important, independent of the MVPA status, effects on weight than increasing the total amount of weekly MVPA. A study in older adults with Type 2 diabetes (72) found evidence to suggest that breaks

from SB could have a positive effect on daily glucose control. Another study (65) done in adults with Type 2 diabetes found that total number of breaks from sedentary time, independent from the total sedentary time and MVPA, were beneficially associated with lower waist circumference, BMI, blood triglycerides, and 2-hr plasma glucose.

There seems to be enough evidence to support the notion that SB is an independent risk factor for chronic diseases and other negative health outcomes in different age groups. These effects may become more prevalent in younger populations, which highlights the need to study and accurately quantify SB across different age groups to draw stronger dose-response relationships and formulate effective public health guidelines that are population specific. Breaks from prolonged sitting may be an important factor to attenuate the negative health-effects related to SB, but this needs further investigation specifically in younger populations.

Definition and measurement of sedentary behavior

In 2017, the Sedentary Behavior Research Network (SBRN) proposed a consensus definition to describe SB in hopes to standardize the terminology so that outcomes from different studies would be more comparable. Behaviors that fall under the definition of “any waking behavior characterized by EE less than or equal to 1.5 METs, while in sitting, reclining or lying posture” are sedentary (14). The combination of the low EE and sedentary posture makes the consensus definition unique yet more stringent than simply measuring either posture or EE alone. What should be noted is that the EE in children during quiet rest and when performing different activities differs from that of adults. Specifically, children have a higher resting metabolic rate (RMR) than adults per unit of body mass that declines gradually as a result of growth and maturation and reaches the adult value by the age of 18 (73). For adults, the value of 1 MET is

$3.5 \text{ ml} \cdot \text{kg}^{-1} \cdot \text{min}^{-1}$ on average, but the respective value for youth can be two to three times higher (74, 75). Therefore, a youth-MET (MET_y) has been established as a more appropriate metric for the purposes of describing EE in youth (74, 75). To calculate the MET_y values for different activities, the individual's measured activity oxygen consumption (VO_2) value is divided by the estimated basal metabolic rate (BMR) (73) or measured RMR.

Criterion measures are the most valid measures against which other measures are compared to. Because the consensus definition of SB is two-fold, a separate criterion measure is used for EE (IC), and posture (DO). The criterion measures are used to validate other measures for estimating SB. Thus, the estimates of time spent in SB are based on posture or EE that is being estimated from raw data. For example, accelerometer-based devices collect raw acceleration data that can be processed with algorithms to obtain estimates of time spent in different postures or activity intensities. Survey methods that collect subjective information of different behaviors can also be compared against the measured EE and posture.

1. Criterion measures

The section that follows will first review studies where IC has been compared against the gold-standard Douglas bag (DB) method. The section ends with a brief discussion of DO and studies done utilizing different DO coding systems.

1.1. Indirect calorimetry

IC is the criterion measure for the EE component of SB. The method is based on the measurement of an individual's VO_2 and carbon dioxide production (VCO_2), which can then be

used to calculate EE. When IC is used to quantify SB, the time periods where the EE is less than or equal to 1.5 METs are regarded as sedentary.

The DB method has traditionally been considered as the criterion measure, i.e. the gold-standard, for respiratory gas exchange measurements (76). However, the method has some limitations, including increased burden for both the participant and investigator, and its inability to detect rapid changes in ventilation (V_E), VO_2 and VCO_2 during higher metabolic loads (77). To overcome these major limitations, portable metabolic systems have been developed and validated against the DB method. Today, several systems are available to be used in laboratory and field studies (e.g. free-living environments) to obtain gas exchange measurements. Some of the most commonly used portable metabolic systems include the Oxycon Mobile (Carefusion Netherlands, Houten, The Netherlands) and Cosmed K4b² (K4b²) and Cosmed K5 [(K5); Cosmed Inc, Rome, Italy]. Studies have investigated the validity of the portable metabolic systems against the DB method (78-81), and other commonly used gas exchange measures, such as metabolic carts (19, 21, 82-87), metabolic simulators (87), and a mass spectrometer (85), with varying outcomes.

Because the DB method is considered as the criterion of all the EE measures, and the K4b² and K5 are the devices relevant for this thesis project, only studies where the portable systems are validated against the DB method will be discussed further below. The search of existing literature revealed a gap in knowledge of the validity of the K4b² and K5 against the DB method in youth. No studies were identified where the devices were validated against the DB method in youth, thus this part of the review will only include studies using adult participants.

1.1.1.1. Cosmed K4b²

The K4b² is a portable IC that can be used in both clinical and field studies for breath-by-breath (BxB) analysis of gas exchange. The BxB procedure calculates $\dot{V}O_2$ and $\dot{V}CO_2$ instantly by pairing continuous flow and gas signals over every breath cycle, requiring a correction of the time delay between flow and gas measurements (80). The BxB mode allows for studying fast changes in respiratory gas exchange and ventilation, but the time delay correction is a critical issue and may result in large measurement errors in $\dot{V}O_2$ and $\dot{V}CO_2$ (80, 88, 89), which often become inflated at higher ventilatory frequencies during heavier workloads (88, 90-92). The K4b² has a galvanic fuel cell for the O_2 analysis, and a non-dispersive infrared CO_2 analyzer, which allows for the calculation of the $\dot{V}O_2$ and $\dot{V}CO_2$ which can then be used to calculate the respiratory exchange ratio (RER) (77). A turbine and sampling line are connected to a sealed rubber mask on one end and transport the gas samples to the gas analyzer on the other end. A mesh headpiece is used to hold the mask in place, and the system is attached to the participant with a specific harness provided by the manufacturer where the device is mounted on the participant's chest. The system weighs approximately 1.5 kg. Prior to each test, O_2 and CO_2 analyzers must be calibrated according to the manufacturer's instructions. The calibration process consists of performing a room air calibration, a flow turbine calibration with a 3.00L Hans-Rudolph syringe (Hans Rudolph Inc, Shawnee, KS), and a reference gas calibration with a known gas mixture (15.93% O_2 , 4.92% CO_2 , N_2 balance). Lastly, a delay calibration is performed to adjust for the time lag between the expiratory flow measurement and the gas analyzers, i.e. the time it takes for the gas sample to be transported from the facemask to the analyzers, and the analyzer response time (77).

1.1.1.1. Validation against DB in adults

A review article by Overstreet et al. (77) summarized validation studies done with 11 portable metabolic systems against the DB method during rest and submaximal and maximal exercise intensities. The intra-class correlation coefficient (ICC) values for the test-retest reliability in the studies done with the K4b² ranged from 0.53 to 0.87 during exercise, 0.85 to 0.96 for gross VO_2 and 0.66 to 0.93 for net VO_2 (gross minus resting) (86, 93, 94). Out of all the 11 portable metabolic systems included in the review, the K4b² was one of the two portable metabolic systems that seemed to provide the most valid estimates of VO_2 in all activity intensities when compared to the criterion DB method. All the other systems tended to overestimate the VO_2 especially in the higher exercise intensities.

Darter et al. (93) determined the test-retest agreement and minimum detectable change (MDC) of VO_2 measures done with the K4b² at different submaximal walking speeds in 22 healthy young adults. The participants completed a 10-min resting VO_2 measurement, followed by two identical treadmill walking tests on nonconsecutive days at five leg-length standardized speeds (0% grade) for 5 min each. The last 2 min of each stage was used to determine the gross VO_2 and net VO_2 . No systematic differences between the tests for gross or net VO_2 were identified ($p > 0.05$). During walking, the ICC values for gross VO_2 (ICC range, 0.85-0.96) were higher than those for net VO_2 (ICC range, 0.74-0.93). The coefficient of variation (CV) for VO_2 was 7.3% at rest and changed little for the gross VO_2 with increasing speed but decreased progressively for the net VO_2 (CV range, 2.0-2.6% for gross VO_2 and 3.0-7.3% for net VO_2). The MDC values for gross VO_2 were $0.88 \text{ mL} \cdot \text{kg}^{-1} \cdot \text{min}^{-1}$ at rest (25.5% of the test means), and $0.75 \text{ mL} \cdot \text{kg}^{-1} \cdot \text{min}^{-1}$ to $1.27 \text{ mL} \cdot \text{kg}^{-1} \cdot \text{min}^{-1}$ (6.9-9.6% of the rest means) at the different walking speeds. The respectable values for the net VO_2 were higher than for gross VO_2 (1.11

$mL \cdot kg^{-1} \cdot min^{-1}$ to $1.49 mL \cdot kg^{-1} \cdot min^{-1}$; 20.3% to 9.8% of the rest means). The results suggested excellent reliability and acceptable measurement errors for the measurement of VO_2 using the K4b², but it might be recommended to report the gross measurements rather than net measurements due to the higher contribution of the RMR to the steady state VO_2 .

McLaughlin et al. (78), compared the accuracy of the K4b² against the DB method. After an initial 10 min rest period, 10 healthy adult male participants completed five 5 min stages on a cycle ergometer at 50-250 W (in 50 W increments) while their VO_2 , VCO_2 , V_E and RER were measured by the K4b² or DB method. The trials were performed on two consecutive days, and the treatments were counter-balanced so that one half of the participants performed the K4b² trial first, and the other half the DB trial first. The subjects remained connected to the K4b² throughout the protocol, and immediately after each protocol a test for analyzer drift with a calibration gas was performed. Values from the last two minutes of each stage were averaged and compared across the two measures. No significant differences in VO_2 ($L \cdot min^{-1}$) were found between the K4b² and the DB method at rest and 250 W, but the K4b² values were significantly higher than DB values at all the other intensities between 50 and 200 W ($p < 0.05$). However, these differences were not practically significant due to their small magnitude [mean difference range, $0.088 L \cdot min^{-1}$ (50 W) to $0.096 L \cdot min^{-1}$ (150 W)]. The K4b² values for the VCO_2 and V_E were not significantly different from the DB values at rest and 50-150 W but were significantly lower at 200 and 250 W ($p < 0.05$). Despite the significant, yet small differences between the measures, the findings suggested that the K4b² is acceptable for measuring VO_2 , VCO_2 , and V_E across different activity intensities.

1.1.2. Cosmed K5

The K5 is an updated version of the K4b² released by the company in 2015. The K5 includes a BxB mode as the K4b², but the updated version also provides a dynamic mixing chamber (DMC) mode. The DMC mode uses a constant flow pump that takes gas samples from multiple breaths and uses them to obtain proportional, fraction of expired O_2 (F_{EO_2}) and fraction of expired CO_2 (F_{ECO_2}) (80). The DMC calibration process includes CO_2 scrubber calibration to obtain real zero CO_2 . In the DMC mode, delay calibration is not necessary. The BxB mode includes the same calibration steps as in the K4b². Unlike the K4b², the K5 is worn on the upper-back, and the system weights approximately 900 g.

1.1.2.1. Validation against DB in adults

Crouter et al. (79) studied the accuracy of the K5 DMC and BxB modes against the criterion DB method in trained cyclists. Fifteen adult participants (14 male) performed an exercise protocol consisting of five 5 min stages on a cycle ergometer at 50-250 W power outputs. The order of the respiratory gas collection methods was randomized, and the gas sampling was performed so that the first method was used during the first four minutes to reach steady state, followed by two minutes of gas collection, except during the rest period, when the sampling period was three minutes for the K5 and five minutes for DB. The second and third collection periods were preceded by a washout period (1 to 3 min). When the participants were switched to a different respiratory gas collection method, additional 2 min samples were collected (except for rest), and this step was repeated for the third respiratory gas collection method. The K5 DMC V_E values were not significantly different from the respective values, but the VO_2 values at rest were significantly higher from the DB (mean difference = $0.05 L \cdot min^{-1}$;

$p = 0.006$). The K5 BxB values for V_E and VO_2 were not significantly different from the respective DB values (both $p \geq 0.05$). The K5 DMC values were significantly lower than the respective BxB values for V_E at 50 W (mean difference = $-1.1 \text{ L} \cdot \text{min}^{-1}$) and for VO_2 at 100 W (mean difference = $0.07 \text{ L} \cdot \text{min}^{-1}$) ($p < 0.05$). The findings suggest that both the DMC and BxB methods are acceptable methods to be used for the measurement of VO_2 and V_E across a wide range of cycling intensities.

Winkert et al. (80) evaluated the reliability and validity of the K5 BxB and DMC modes against the criterion DB method at different metabolic rates. Two K5 devices operated simultaneously in BxB or DMC modes. Twelve endurance trained adult male participants performed the same cycling protocol during three consecutive days [8 min at 30, 40, and 50% of the power output at their measured VO_{2max} (PVO_{2max}), and 4 min at 85% PVO_{2max}]. The biological validity decreased with increasing metabolic rate. The BxB mode significantly underestimated the DB measured VO_2 at 85% PVO_{2max} (mean difference, -0.27%, $p < 0.01$), and it was also significantly less than the respective DMC value (mean difference, -5.74%, $p < 0.01$). Both modes significantly underestimated V_E at 50% and 85% PVO_{2max} , and the DMC also at 30% PVO_{2max} [mean difference range -6.32 (DMC at 30% PVO_{2max}) to -13.34% (DMC at 85% PVO_{2max}); $p < 0.05$], and the DMC values were also significantly less than the respective BxB V_E values at each stage ($p < 0.05$). The conclusion was that the DMC mode might be better suited to be used when performing testing at high metabolic workloads compared to the BxB, which was significantly different from the criterion more often than the DMC mode.

1.1.3. Conclusion

The existing literature is missing validation studies of the K4b² and K5 against DB in youth. The adult validation studies suggest small differences when the K4b² and K5 are compared against the DB during exercise; and for gross and net VO_2 . The limited number of validation studies done with the K5 makes it difficult to draw conclusions, but despite the small yet statistically significant differences between the BxB and DMC modes in the K5, both modes seem to provide similar physiological measures and can be used to measure VO_2 and V_E across a wide range of activity intensities.

1.2. Direct observation

DO is the criterion measure for the postural component of SB and can provide specific information regarding the type, duration and contextual variables (e.g. physical and social environment) of the activity. When DO is used for quantifying SB, time stamps with a seated, reclining, or lying posture are regarded as sedentary. DO comes with a low participant burden as no specific equipment or actions from the participants are required. The method is also not limited by recall bias that is typically related to survey methods, and is therefore particularly useful with younger participants who may not be able to provide accurate and detailed descriptions of their behaviors (95). However, some downsides of DO are that it is time-costly for the investigator and requires practice and an adequate level of skill to provide valid records of the behaviors (17). The method may also not be suitable for larger populations, and another common issue related to its use is the probability of participant reactivity to the observer's presence.

Several coding systems for DO have been used and validated for the purposes of quantifying different behaviors. DO can be accomplished by using either second-by-second coding (continuous focal sampling) where all the occurring behaviors are coded, such as in the Movement Observation in Children and Adolescents (MOCA) (96); or with momentary time sampling, where a set observation interval is followed by a set recording interval, such as in the Children Activity Rating Scale (CARS), the System for Observing Play and Leisure Activity in Youth (SOPLAY) (97), and the System for Observing Play and Recreation in Communities (SOPARC) (98). Different coding systems utilize different activity and intensity categories, which may influence the validity of the records. For example, The CARS has separate codes for stationary activities with and without limb movement; the SOPLAY and SOPARC include standing posture in the sedentary category, and the MOCA uses the Compendium of Physical Activities (22) EE categories to distinguish between SB and light-intensity PA (LPA).

A variety of computer programs have been developed for the purposes of DO and pairing the observation records with data from other measures. Some programs allow the creation of customized coding schemes that can be accessed on handheld devices and used for live coding. One of the most commonly used DO software is the Observer XT by the Noldus Informational Technologies (Wageningen, the Netherlands), which includes these features. The Observer XT software has an extension application called the Pocket Observer that can be downloaded on a tablet for live coding.

1.2.1. DO in youth

CARS (95) is one commonly used DO coding system in youth studies that includes five intensity levels [1. stationary/motionless, 2. stationary/movement of limb(s) or trunk, 3.

translocation (slow speed/easy), 4. translocation (medium speed/moderate), and 5. translocation (fast or very fast/hard)]. The observations are performed in 1 min intervals, and each intensity level can only be coded once within an interval; and the observed behavior must occur for at least 3 sec before it can be coded. If two different activity levels occur for less than 3 sec each but are repeated three or more times in a 15-20 sec interval both activity levels should be coded.

Puhl (95) calibrated the CARS in a sample of 25 5-6-year-old children against a Beckman LB-2 CO_2 analyzer, an Applied Electrochemistry SA-3 O_2 analyzer, and a Vacumed air flow meter. The gas samples were collected in 30 sec intervals and averaged over 60 sec. VO_2 was measured for eight activities from the Levels 1-5 [lying, sitting, standing and coloring, standing and throwing/catching a NERF ball; walking on a treadmill at 2.5 mph for 1 min followed by walking on a treadmill at 2.5 mph at four different grades (0%, 5%, 10%, and 15%); running on the treadmill at 4.0 mph at 10% grade, followed by running on the treadmill at 4.0 mph with increasing grade in 2% increments until exhaustion]. The first four activities were performed for 5 min, and the mean values for minutes 2-5 were used. The calculated inter-observer agreement was $84.1 \pm 10.1\%$. No significant differences between the sexes for the mean absolute VO_2 during all the eight activities, and for the mean relative VO_2 during Level 1 and 2 activities were observed. The measured VO_2 values were significantly different across the different activity levels for both sexes ($p < 0.05$), suggesting that the CARS levels distinguish between different intensity levels.

The MOCA is a study that uses free-living accelerometer data and DO for studying youth behavior. Cox and Petrucci (96) developed the MOCA DO system that includes codes for four main outcomes [1. Whole-body movement, 2. Locomotion, 3. Activity type, and 4. MET value (based on the Youth Compendium of Physical Activities (75))]. After the DO development they

performed an initial agreement analysis and refinement of the system with a sample of 28 youth (6-10 years). Each participant performed a 30-min period of free-play while being video recorded. The video recordings were later coded by six observers. Intra-rater agreement was acceptable (>80%) for all variables besides the activity MET value ($69 \pm 27\%$). Inter-rater agreement (against an “expert coder”) ranged from 50 to 95% for the four main outcomes.

1.2.2. DO in adults

A study by Lyden et al. (99) compared the agreement between DO and IC for time spent in different intensity categories. Fifteen adult participants were directly observed on three separate days (two hours per session) using a hand-held personal digital assistant with custom software (Noldus Inc., Wageningen, The Netherlands). Twelve general PA categories with the corresponding MET values from the Compendium of Physical Activities (22) were used to categorize the behaviors during DO. The participants also wore an Oxycon Mobile IC for the measured EE. The measured SB by DO was not significantly different compared to the measured SB by IC (mean difference, -7.6%; ICC, 0.80). The findings provided evidence that a carefully designed DO method may be equally valid compared to IC as a criterion measure of SB based on the measurement of EE in free-living validation studies.

Welch et al. (100) evaluated the accuracy of DO for estimated EE (METs), intensity category (sedentary-vigorous), and intensity category by domain (e.g. occupational, leisure, etc.) against the measured values by the K4b². Nine older adults (71.0 ± 6.9 years, 67% male) performed 1-hr of unstructured activity in laboratory and 3-hr under free-living conditions. A custom DO platform was built in the Noldus Observer XT and included codes for activity type and modifiers for the estimated MET value. Of the in-laboratory activities, 45% of the estimated

MET values by DO were within 0.5 METs of the measured EE by the K4b². DO also correctly classified 45% of the activity intensity categories, but SB was the most often misclassified (overestimated) intensity category with 68.9% of the observations being misclassified. The correct intensity classification by domain was the lowest for sedentary/occupational activities (41.6%). Of the free-living activities, 61% of the estimated activity MET values by DO were within 0.5 METs of the measured values by the K4b². The overall activity intensity classification accuracy was 50.9%. Based on the results, the DO accuracy for estimating EE and activity intensity categories were higher under free-living conditions compared to the laboratory. Due to the lower accuracy of the estimates, using a continuous DO sampling method may not be the best option for obtaining accurate estimates of activity METs. Despite the low variability between the mean MET values between the estimated by DO and measured by the K4b² (0.5 METs), further investigation is needed to refine this approach.

In the third phase of the previously mentioned study by Cox and Petrucci (96) the intra- and inter-rater variability of the MOCA coding system was investigated in a subsample of ten young adults (18-24 years). The participants completed four 1-hr sessions in a distinct free-living setting (home, school, community, and PA/exercise). Each session was video recorded, and each occurring activity labeled for the main outcomes [1. Whole-body movement, 2. Locomotion, 3. Activity type, 4. MET value (based on the Compendium of Physical Activities (22)), and 5. Intensity category]. The intra-rater reliability for all variables and environments ranged from 91.9% to 100%. Of the observation environments, the home environment had the highest intra-rater variability (93-100%), and the community environment the lowest (91.9-96.4%). The ICC for the activity intensity category ranged from 0.86 (community) to 1.00 (school and home). The inter-rater reliability for all the variables and environments ranged from 88.2% (community) to

100% (school). The ICC values suggested high intra- and inter-rater reliability for the MOCA DO coding system in adults across different environments.

1.2.3. Conclusion

Several DO coding systems are available for the purposes of quantifying behaviors. However, most of the systems have been developed and validated for the purposes of measuring PA behaviors and may have limited utility in SB research due to a wide range of heterogeneity across the systems, and the differences in categorizing SB and LPA.

2. Subjective and objective measures

Validity means the extent to which an assessment accurately measures what it intends to (101). There are different types of validity, including content-, construct-, and criterion-referenced validity. The criterion-referenced validity is the one of interest when validating objective measures, such as accelerometer-based devices for quantifying SB; the former two are more relevant when validating subjective measures, such as questionnaires (101, 102). There are two types of criterion-referenced validity: concurrent validity and predictive validity (102). When determining the concurrent validity of a measure, simultaneous data collection is done with a criterion measure and the measure that is being validated, and the data are then compared or correlated to observe the agreement (102). Predictive validity, on the other hand, means the extent to which an objective measure is able to predict the outcomes obtained with a criterion measure (102). Studies have validated secondary measures against the criterion measures for estimating EE, time spent in different postures and in SB. The techniques used in research and to guide program development and establishment should be as valid, reliable, practical, and non-

reactive as possible (16). What follows is a brief discussion of self-reports and surveys (subjective methods) for quantifying SB. Lastly, validation studies where estimates of SB from accelerometer-based devices (objective methods) have been compared against the SB criterion measures will be presented.

2.1. Self-report

Self-report questionnaires and diaries have traditionally been used to collect health-related information among different populations. Compared to several other measures, the method of using self-reports is often considered more suitable for larger-scale studies as it typically allows for larger sample sizes. The method is relatively simple and does not require the participants to change their current behaviors. The problem is, however, that several survey methods are accompanied with a large degree of error. Studies have validated surveys that include questions related to SB against criterion measures, such as DO (103, 104); and against objective measures such as accelerometer-based devices (105-114), heartrate monitors (115), and pedometers (114-116), with a variety of outcomes. Designing a survey that captures as much detail of SB as possible yet is easy to administer and interpret by the participant is difficult, as the following summary of some SB surveys demonstrate.

Many questionnaires focus on TV viewing habits as an indicator of sedentary time (112). A review article by Prince et al. (117) identified 16 pediatric and 18 adult surveys and epidemiological studies that assessed at least one modality of SB. Variability was present across the questionnaires regarding the types of SB, but TV viewing time seemed to be the most frequently assessed SB (pediatric surveys = 100%, adult surveys = 72%). Time spent playing computer and/or video games were also frequently assessed, specifically among the pediatric

populations (pediatric surveys = 88%, adult surveys = 39%). Reading was included in about half of the surveys in both populations (pediatric surveys = 44%, adult surveys = 50%), and a smaller number of surveys were interested in sedentary transport (pediatric surveys = 13%, adult surveys = 33%). A cross sectional study by Carson and Janssen (67) reported that TV viewing and computer use are the most common forms of screen-based SB in youth. A review article by Marshall et al. (118) reported that the youth TV viewing habits have barely changed over the past 50 years, but the amount of time spent engaging with electronic entertainment media at home has experienced a rapid increase. It is estimated that youth spend up to five hours each day with these media accounts (118). Although relevant, focusing on TV viewing or other screen time as a sole indicator of SB leaves out several other behaviors that largely contribute to the total sedentary time each day. The validity of most survey measures related to TV viewing is of unknown quality (119), and no self-report measure is comprehensive to assess all SB (39).

Among preschool-aged children, Prince et al. (117) found only two questionnaires that demonstrated excellent levels of reliability for TV viewing: The Preschool-Aged Physical Activity Questionnaire (Pre-PAQ) (120) and the proxy-report used in the Health, Eating and Play Study (121) (ICC, range 0.70-0.88 and 0.78, respectively). The Pre-PAQ also demonstrated high levels of reliability for playing computer and video games (ICC range, 0.82-0.85). For youth 9-12 years old, the COMPASS questionnaire (122) had the highest ICC (0.56) for TV viewing, and for playing computer and video games (ICC, 0.65) and surfing the internet (ICC, 0.71) among high school students. Among adults, the Sedentary Behavior Questionnaire (105) had excellent reliability for weekday and weekend TV viewing (ICC, 0.86 and 0.83, respectively). Reliable questions for the assessment of TV viewing were also included in The Past Week Modifiable Activity Questionnaire (ICC, 0.67), Salmon SB questionnaire (ICC, 0.82)/Measure of Older

Adults Sedentary Time (MOST; ICC, 0.76); Sedentary, Transportation and Activity Questionnaire (ICC, 0.79), and the SIT-Q (ICC, 0.84) (108, 123-125). The MOST also had very high levels of reliability for the question that assessed computer and internet use (ICC, 0.79).

When using self-reports and questionnaires to quantify SB, relying on a participant's memory and ability to correctly answer the questions and classify behaviors introduces a potential source of error. The manner of recording activities also influences the degree of bias, i.e. whether the activities are reported on a minute-by minute basis over a short period of time, or recalled over a longer period of time (126). Surveys often include questions that ask the participant to quantify the time spent in different SB during a typical weekday and a weekend day, typically in blocks of 15 to 30 min (127), which may result in a great degree of error in the estimates especially if extrapolated out for a longer period of time. Also, the participant's interpretation of the questions may influence the validity of a questionnaire as discussed by Prince et al. (117). They found that there were differences in the wording of questions across surveys and many included a specific time period for recall (pediatric surveys = 38%, adult surveys = 44%), and questions about a "typical" or "usual" day or week in the specific time period. Some studies focused on hours per day of each mode of SB (e.g. hours of TV viewing in a typical day), whereas others focused on hours per week, average of the days of the week, or school/work/weekdays and weekends separately. Several surveys separated the questions for each modality of SB (pediatric surveys = 63%, adult surveys = 28%); some surveys used a grid or list of SB modalities where the participant entered the daily time spent in each modality.

2.1.1. Self-report in youth

Self-reports of SB obtained from children may be biased due to their unpredictable behavioral patterns and lower cognitive function compared to adults (16). Using proxy-reports with parents can be unreliable especially if the children are school-aged and not constantly around the parents (128). Salmon et al. (121) used a proxy-report to observe whether children's TV viewing is a good indicator of obesity-promoting vs. healthy eating behaviors. A sample of 1560 Australian primary school pupils [n = 613, age 5-6 years (50% boys), and n = 947, age 10-12 years (46% boys)] and their parents participated in the study. The parents reported the time the child spent watching TV during a typical school day and a weekend day (on a scale of 0 to 6 or more hours per day in half an hour segments); as well as the child's participation in organized PA, and their intake of different foods. The child's PA levels were objectively assessed for one week using an ActiGraph (AG) model AM7164 (ActiGraph LLC, Fort Walton, FL, USA). For both TV viewing and MVPA, the data was dichotomized into ≤ 2 -hrs/day and > 2 -hrs/day. An inverse relationship between TV viewing and frequency of participating in organized MVPA, and a non-significant inverse association with objectively assessed MVPA were observed. The researchers suggested that TV viewing might be an important indicator of a range of lifestyle behaviors that are related to SB and weight gain and asking about children's TV viewing habits could reveal associations with unhealthy and inactive behaviors and eating patterns.

The Pre-PAQ was developed to measure the behaviors of young children specifically in their home environment (120). The Pre-PAQ includes questions related to PA and SB, and it was designed as a 3-day recall. One hundred and five preschool-aged children (age 3.0-5.9 years) and their parents participated in the validation study where the Pre-PAQ was compared against accelerometry outputs. The test-retest reliability was also assessed by asking the parents to fill

out the survey in two different occasions one to two weeks apart. Participants wore an AG 7164 device on their right hip for 4-5 days, and the first day was excluded to minimize the participant reactivity. Sedentary thresholds developed by Sirard et al. (129) [vertical axis (VA) <300 counts/15 sec] and Reilly et al. (130) (VA <1100 counts/60 sec) were used to distinguish SB from other behaviors. The lowest reliability of the sedentary items included in the Pre-PAQ was for the total time the child spent in car for transportation during weekend (Saturday ICC, 0.37; and Sunday ICC, 0.31). “Parental screen time recreation” (STR) measurement error ranged from 5.5 to 13.8 min for time spent on the computer on a weekend and TV viewing during the week, respectively. The ICC for the agreement of the time child spent in stationary activities was also low (0.44). The agreement between the Pre-PAQ sedentary level was poor compared to the sedentary time derived using both Sirard et al.’s and Reilly et al.’s thresholds (mean difference = -235.4 min/day and -208.6 min/day, respectively). The conclusion was that the use of the Pre-PAQ for obtaining estimates of SB in children may not be a recommended practice.

2.1.2. Self-report in adults

Mensah et al. (127) examined the reliability and validity of the Sedentary, Transportation, and Activity Questionnaire for assessing context specific SB and transport-related activities in a sample of 96 adult participants (51 women). The questionnaire collects past-month information regarding behaviors in different domains (domestic, work, commuting and transportation, and leisure). The participants wore an AG GT3X+ on the right hip for seven consecutive days and filled out an activity logbook. For a participant subgroup (n = 45), the study period was extended to 14 days. The participants self-administered the questionnaire at the end of the study (on days 7 or 14). After a month, 32 participants completed the questionnaire second time for the

assessment of test-retest reliability. SB threshold VA <150 counts/60 sec was used with the AG. The test-retest reliability was fair for total sitting time (ICC, 0.52). A significant correlation and a fair-to-good agreement between the questionnaire-derived total sitting time and the AG-derived SB was observed ($r = 0.54$; ICC = 0.44). However, the questionnaire significantly underestimated the total weekly sitting time (mean difference = -17-hrs/wk) with wide 95% limits of agreement (LOA). The time spent viewing TV/DVD had the highest ICC (0.79) of the domestic activities. Sitting time at work had the strongest association ($r = 0.88$; ICC = 0.82) of all the different activities and domains. Generally, the questionnaire overestimated the work and leisure-related SB and sitting time and underestimated the total sitting time and sitting time related to transportation. Based on the findings, the Sedentary, Transportation, and Activity Questionnaire can be regarded as an easy-to-use tool that could be used for the assessment of SB and transportation in different contexts in free-living settings.

Craig et al. (105) examined the reliability and validity of the IPAQ sitting items against estimated SB time using an AG 7164 in 289 adults (44.6% male) from 12 different countries. Participants were asked to wear an AG 7164 for seven days, and both the short and long version of the IPAQ were administered. For the AG 7164, VA <100 counts/60 sec was used as the sedentary threshold. No differences in the reliability coefficients for all the sitting items between men and women using the long form were observed. Most correlations indicated a very good-to-excellent test-retest reliability for the sitting items in the long version ($r > 0.75$; range, 0.40-1.0 for men, and 0.64-0.96 for women) for the sitting items. For the short form, the most test-retest reliabilities were also in acceptable ranges for both men and women ($r > 0.70$; range, 0.77-0.98 for men, and 0.62-0.96 for women). For weekday and weekend sitting, the reliabilities were slightly lower compared to those achieved with the long form, but similar for the overall sitting.

The results indicated that the IPAQ sitting items have acceptable reliability and validity for men and women representing different nationalities.

Visser and Koster (112) investigated the agreement between self-reported sedentary time and the sedentary time derived from an AG GT3X using the SB threshold VA <100 counts/60 sec in a sample of 83 older men and women from the Longitudinal Aging Study Amsterdam. The participants performed a protocol of ten sedentary activities (napping, reading, listening to music while sitting or lying; watching TV, video or DVD; sitting at the computer for work or leisure, seated working, seated hobbies, talking on the phone, transportation, church/theater) completed on a regular weekday and a weekend day. The self-reports were compared against the data from eight days of AG wear. The researchers determined the activities that should be used in a new SB questionnaire: napping, reading, listening to music, watching TV, seated hobbies, and talking were the activities that were found to best correlate with the objectively estimated SB ($r = 0.46$, $p < 0.05$), yet the total self-reported SB was still underestimated (mean difference = -2.1 h) when using these activities only. Watching TV was found to be the most common SB among the older persons, but the researchers concluded that it should not be used independently as an indicator of total SB in this population. The researchers concluded that because developed questionnaire had moderate correlations with the objectively measured SB, it might compliment other SB questionnaires used to assess the total daily activities across a range of intensities.

Studies that have used DO as the criterion to validate SB-related surveys have reported both overestimation and underestimation of SB compared to the observations. In a study by Douwes et al. (104), self-reports obtained from 99 bank employees overestimated daily computer use by 99 min (39%; $p = 0.001$). Self-reports of SB from 44 adult participants in a study by Klesges et al. (103) underestimated the measured time spent in SB by over 300%. The studies

discussed above demonstrate that there is a lot of inconsistency when the self-reported estimates of EE and total sedentary time are compared against the measured EE and sedentary time by the criterion measures, or against the estimates from accelerometer-based devices. The findings suggest that using self-reports to assess SB may not be a recommended practice. The type of the questionnaire, recall period, and activities included in the analysis all influence the accuracy of the records and thus, the validity of the measure against the criterion. Capturing all SB with subjective measures may be difficult, and another limitation related to their use is that they are often unable to assess possible breaks from SB, which has recently become a new topic of interest. Despite the several drawbacks related to the self-report and survey methods, they still come with some advantages compared to objective measures, including an opportunity for collecting data from a larger sample (131), and possibly providing more information about activities that are difficult to capture by objective measures.

2.1.3. Conclusion

A variety of questionnaires that include the assessment of SB have been utilized in research, but the validity of the methods may be questionable due to a wide range of outcomes from validation studies. The validity of the SB self-reports seems to be influenced by the specificity of the questions, the recall period, and the SB components included. TV viewing and other screen time are some of the most commonly assessed SB in both youth and adults and found to correlate well with total sedentary time in both groups. However, no questionnaire is comprehensive to assess all SB.

2.2. Accelerometer-based devices

Accelerometer-based devices are small devices that can be used to obtain estimates of sedentary time. The devices collect raw acceleration data from the movement of the body-part where the device is worn (usually hip, ankle, or wrist) (132, 133). Researchers have developed and validated several algorithms against criterion measures for the purposes of processing and quantifying different activity components from the raw acceleration data, including intensity, duration, and frequency of behaviors (134). Typically, accelerometer-based devices contain at least one, sometimes multiple accelerometers that are used to obtain samples in different ranges and rates; and an inclinometer function that is used to measure the changes in the angle of the body-part where the device is attached to.

A drawback of using accelerometer-based devices and estimating activity intensity from the accelerometer “counts” is that they may be less accurate for detecting SB and low-intensity activities. The estimates are based on lack of movement and not exceeding the thresholds, thus, they may be unable to discriminate between sitting and non-sitting stationary behaviors, and sometimes activities that require muscular activity but are accompanied with little movement of the body part where the device is placed may be misclassified as SB (39, 135). The epoch length may also influence the study results and significantly influence the estimated time spent in SB (136). Shorter epoch lengths (e.g. 15 sec) may result in lower estimated sedentary time and higher active time compared to longer epoch lengths (e.g. 60 sec) (136). Most studies have used thresholds based on 60 sec epochs, probably due to their ease of use and converting the accelerometer counts to activity minutes (44).

2.2.1. ActiGraph

The most widely used accelerometer-based devices in both SB and PA research are the different generations of the AG devices (Pensacola, FL) (131). The AG devices are commonly worn on the wrist or hip, but other possible wear-locations are the thigh and the ankle. The raw acceleration data can be uploaded from the device to the manufacturer's ActiLife software on a PC. Count and inclination estimates can be obtained from the raw acceleration data and used to estimate time spent in SB and other activity intensities, and in different postures. DO has been used in several studies to validate the estimates of SB (23, 24, 26, 33, 39, 137) and time spent in different postures (48, 138). Studies have also used IC to validate the estimates of activity intensity (27, 30, 42, 45, 139, 140), and time spent in SB (33). What follows is a brief summary of validation studies of the AG SB estimates, broken into youth and adult studies, and studies where either DO or IC served as the criterion measure of SB.

2.2.1.1. Validation against IC in youth

Crouter et al. (30) previously developed a two-regression model for estimating PA and SB in youth from data from hip-worn accelerometer-based devices. The validity of the two-regression model to estimate SB was examined against the K4b² under free-living conditions. Youth participants (11-15 years; 57% boys) performed approximately two hours of free-living activity while wearing an AG GT3X+ device on the right hip, and a K4b² for measured EE. The accelerometer data were processed using three different SB thresholds [vector magnitude (VM) <75 counts/10 sec, and VA <25 counts/10 sec (28), and VA <100 counts/60 sec). The measured and predicted EE was expressed as MET_y, and MET_y <1.5 was equated as SB. None of the estimates were significantly different from the measured SB [mean difference range, 0.6 min

(2.1%; VA <25 counts/10 sec) to 6.9 min (23.0%; VM <75 counts/10 sec); all $p \geq 0.05$]. The findings suggest that objective measures may be appropriate to obtain estimates of SB in youth.

Crouter et al. (27) validated methods for analyzing wrist worn AG GT3X+ and GT3X data against the K4b² as the criterion measure for EE. The used wrist SB thresholds used VA ≤ 35 counts/5 sec and VM ≤ 100 counts/5 sec. Forty-two youth (8-15 years) completed approximately 2-hr of unstructured activity. During the unstructured activities, the estimated time spent in SB by VA ≤ 35 VA counts/5 sec and VM ≤ 100 counts/5 sec were not significantly different from the measured time spent in SB (mean bias = 2.1 and 2.5%, respectively; both $p > 0.05$), but both thresholds had large individual errors.

2.2.1.2. Validation against IC in adults

Crouter et al. (30) compared three SB thresholds [VA ≤ 25 counts/10 sec (141), VA ≤ 8 counts/10 sec (40), and VA <100 counts/60 sec (58)] against the K4b² during free-living activities. Twenty-nine adult participants (58.6% female) wore a K4b² for measured EE, and an AG GT1M on the right hip during a 5-6-hr sampling period. Twenty-three participants were measured at work, and three during leisure time. VA ≤ 8 counts/10 sec significantly underestimated the measured SB by the K4b² (mean difference = -34.9 min, -20.8%, $p < 0.05$). VA ≤ 50 counts/60 sec was not significantly different from the measured SB by the K4b² (mean difference = -3.0 min, -1.8%, $p > 0.05$). VA <100 counts/60 sec SB threshold overestimated the measured SB by the K4b² (mean difference = 16.5 min, 9.9%, $p > 0.05$). VA <50 counts/60 sec had the smallest mean bias compared to the K4b² for the time in SB.

2.2.1.3. Validation against DO in youth

Fischer et al. (23) used an AG Actitrainer to compare the SB thresholds VA <1100, <800, <300, and <100 counts/60 sec against measured SB by DO in 29 children (5-11 years old, 28% boys). The participants performed a protocol of supervised sedentary activities in free-living settings. Each activity was performed for 20 min while wearing the AG device on the right hip. The protocol was designed to reflect the common SB among children. After the protocol, a period of free play outside was used to determine whether the thresholds could distinguish between sedentary and higher intensity activities. The data was analyzed as median, 25th, and 75th percentiles, and the results suggested that the estimated time spent in SB is a factor of the different thresholds used. VA <1100 and 800 counts/60 sec were considered as too high and resulted in an overestimation of the actual time spent sedentary, whereas the lower thresholds, namely VA <100 counts/60 sec underestimated the total sedentary time according to the 75th percentile, but provided the best estimate when based on the median.

2.2.1.4. Validation against DO in adults

Carr and Mahar (26) examined the validity of SB thresholds and the inclinometer output from an AG GT1M and GT3X+ against DO. Thirty-six college-aged adults (55% female) wore the devices on their right hip during a protocol that consisted of six activities, four of which were sedentary. Each activity was performed for eight minutes but the first and last minute was excluded, and the middle six minutes was used for analysis. The measured DO activity intensities were derived from the Compendium of Physical Activities (22) (*Note: The researchers classified standing still as a sedentary activity based on the Compendium of Physical Activities activity intensity categories*). SB thresholds VA <100 counts/60 and VA <150

counts/60 sec were applied to the GT1M and GT3X+ data, and the total minutes that were under the SB thresholds were compared against the SB measured by DO. The total number of minutes categorized as lying, seated, or standing by the GT3X+ inclinometer function was compared against the measured SB by DO. Percent agreement with DO and root mean squared error (RMSE) were calculated for each estimate. For the GT1M, VA <100 counts/60 sec and VA <150 counts/60 sec correctly identified greater than 96.8 to 100% of the measured sedentary time. The GT3X+ VM <100 counts/60 sec and VM <150 counts/60 sec correctly identified 82.9 to 97.7% of the measured sedentary time. The SB-threshold VA <100 counts/60 sec had a higher RMSE compared to the <150 counts/60 sec for both devices (0.13 vs. 0.07 for the GT1M, and 0.86 vs. 0.68 for the GT3X+, respectively). The GT3X+ inclinometer function correctly identified 60.6 to 67.7% of the sedentary time and had the highest RMSE of all estimates (2.15). The conclusion was that using a higher SB threshold (VA <150 counts/60 sec) might decrease the error of the SB estimates for correctly identifying SB time, but caution should be taken when basing the SB estimates on the inclinometer outputs.

Peterson et al. (25) explored the degree of construct validity of the AG GT3X+ inclinometer function and the ability to distinguish between sedentary and non-sedentary behaviors using both VA <100 counts/60 sec, VA <150 counts/60 sec, VM <100 counts/60 sec, and VM <150 counts/60 sec against DO as the criterion measure. Twenty-eight young adults (18-20 years) wore an AG on the right hip while performing a protocol of nine activities (six sedentary). From the GT3X+ inclinometer data, SB was defined as time categorized as sitting or lying. All the SB thresholds were less biased compared to the inclinometer function for estimating SB: the percent agreement with DO for all the sedentary activities were 98.8-100% for VA <100 counts/60 sec, 100% for VA <150 counts/60 sec, and 91.7-100% for VM <100

counts/60 sec and 95.2-100% for VM <150 counts/60 sec. The inclinometer had more variability, ranging from 44.9 to 80.8% agreement with DO for SB. The conclusion was that the inclinometer function was more biased for estimating SB compared to the SB thresholds when DO was used as the reference. Therefore, it may be recommended to use the AG counts over the inclinometer function to estimate SB.

Marcotte et al. (24) used DO and EE values derived from the Compendium of Physical Activities (22) to determine the validity of AG SB thresholds for estimating SB during free-living conditions. Forty-eight young adults (18-25 years) were observed using the MOCA system in four settings (home, school, community, and exercise) while wearing two AG devices, one on the non-dominant wrist, and the other on the right hip. Thirteen different SB thresholds were used to obtain the estimates of SB time [hip: VA <100 counts/60 sec (58), VA <150 counts/60 sec (39), and VM <200 counts/60 sec (142) with the low frequency extension (LFE) filter enabled and disabled; Sojourn-1x and Sojourn-3x (143), and Euclidean norm minus one [(ENMO); ENMO 47.4 (144)]; wrist: VM <1853 counts/60 sec and VM <376 counts/15 sec (145), ENMO 27.9 (146) and ENMO 44.8 (144); the sedentary sphere (147), and two machine learning algorithms (random forest and decision tree) (148)]. For the hip, VM <200 counts/60 sec underestimated the measured time in SB (mean difference = -5.5 min, $p < 0.001$) and the ENMO 47.4 overestimated measured SB time (mean difference = 12.2 min, $p < 0.001$). The other methods were not significantly different from measured SB time. For the wrist, VM <376 counts/15 sec, VM <1853 counts/60 sec, ENMO 27.9, and the decision tree underestimated measured SB time by -9.5 to -2.5 min ($p < 0.001$); and the sedentary sphere and ENMO 44.8 overestimated measured SB time by 3.9 and 6.1 min, respectively ($p < 0.001$). In general, the hip methods had smaller mean biases [near zero for all but ENMO 47.4 (about 12.5min) and LFE

VA <200 counts/60 sec (about -6 min)] compared to the wrist methods [mean bias range, about -2.5 mi to -10 min for all besides the sedentary sphere and ENMO 44.8 (about 5-6 min)]. The hip methods had a greater classification accuracy and lower individual-level errors [mean absolute percent error (MAPE) hip range, ~18% (LFE VA <150 counts/60 sec) to ~38% (ENMO 47.4) vs. wrist range, about 30% (ENMO 44.8) to about 40% (random forest)]. The wrist methods underestimated measured SB time at home, school, and community. The sitting behaviors and activities of daily living (ADLs) had the largest number of false negatives across methods (>91.5%). The findings suggest using hip-based methods may be a more recommended practice for estimating SB time.

2.2.2. activPAL

The activPAL [(AP), PAL Technologies Limited, Glasgow, Scotland, UK] is a small accelerometer-based device that is worn on the anterior midline of the thigh. The device measures static and dynamic acceleration and has an inclinometer function that measures the changes of the thigh incline. The raw data can be uploaded to a PC and the activPAL Professional Research Edition software used to distinguish between different postures (“sitting/lying”, “standing”, and “stepping”). The software uses “Intelligent Activity Classification Routines” which are proprietary algorithms used to create an “activity classification profile” that includes both the posture, intensity, and number of steps taken during the activity (49, 138).

The PAL family includes both uniaxial and triaxial accelerometer-based devices. Numerous studies have validated the AP posture algorithms against DO (31, 34, 39, 48, 149) and IC (33, 36). The device is not a criterion measure, but due to its potential in accurately estimating

time spent in, and breaks from, seated and lying postures (35), it has been widely utilized in SB research and often considered as the best alternative if DO is not an option. What follows is a brief summary of validation studies of the AP algorithm, broken into youth and adult studies, and studies where either DO or IC served as the criterion measure of SB.

2.2.2.1. Validation against IC in youth

Dowd et al. (36) validated the AP against IC in a sample of 30 female adolescents (15-18 years). The participants wore an AP device on the right thigh, and a K4b² for the measured EE. The participants performed sitting, standing, and three locomotor activities (slow and brisk walking and light jogging at self-selected pace). For sitting, the AP had a perfect (100%) agreement with the DO measured time spent sitting.

2.2.2.2. Validation against IC in adults

Harrington et al. (150) validated the AP MET equation against IC in a sample of 62 young adult females (15-25 years). The participants performed a laboratory protocol consisting of RMR and locomotor activities while wearing an AP on the right thigh and an AMIS 2000 metabolic measurement system (Innovision, Odense, Denmark) for measured EE. The AP step-based MET estimates were compared to the measured METs. The estimates significantly underestimated the measured values (mean difference = -0.5 METs, ICC = 0.57, $p < 0.001$).

2.2.2.3. Validation against DO in youth

Aminian et al. (31) validated the AP for distinguishing between different postures, and the number of sit-to-stand transitions against DO in a sample of 25 children (9-10 years). The

participants wore an AP device on the right thigh while performing an activity protocol consisting of activities ranging from sedentary to vigorous intensities, performed for 2 min each, as well as activity patterns that simulated typical classroom behaviors of children in this age. The AP data from each 2 min activity was summed in 15 sec epochs. A perfect correlation ($r = 1.00$) was observed between the AP and DO for time spent in “seated/lying” posture, and the activity patterns, one of which included transitions between seated and upright posture. A high correlation ($r = 0.99$) between the AP and DO was observed for the total number of sit-to-stand and stand-to-sit transitions for each 2 min activity. The results suggest that the AP is a valid tool for the measurement of time spent “seated/lying” and the total number of transitions between the seated and lying posture in school-aged children.

Davies et al. (34) validated the AP postural output against DO. Thirty-two children (3.1-4.9 years, 66% female) wore an AP device on the right thigh while engaging in their typical activities. The AP and DO data were time synchronized and analyzed on a second-by-second basis for posture (seated/lying, standing, walking). Sensitivity, specificity, and positive predictive value (PPV) were calculated for each AP posture category. When analyzed at the individual level, 46% (755 min) of the time was in the “seated/lying” category according to DO compared to 42% (688 min) estimated by the AP. The AP underestimated the total sitting time (mean difference = -4.4%, $p < 0.01$). The median AP “seated/lying” sensitivity was 92.8%, specificity 97.3%, and PPV 97.0%. Based on the results, the AP could be utilized for estimating time spent in sitting and/or lying in young children during free-living conditions.

Janssen et al. (32) validated the AP time in SB and breaks from SB against DO. Forty children (4-6 years) performed a 150 min structured activity protocol in a laboratory consisting of SB, LPA and MPA. The data collections were video recorded and later coded on a second-by-

second basis for posture, which included coding for sitting/lying, standing and stepping to match the AP. An “other” category was also coded in the DO that included activities such as crawling, kneeling, leg hanging off a chair, etc., that didn’t have a clear posture identified. The number of breaks from “seated/lying” were also calculated for each participant. Sensitivity, specificity, and receiver operating area under the curve (AUC) were used to evaluate the AP classification accuracy for posture allocation. When postures classified as “other” by DO were included in the analysis, the AP misclassified 51.7% of the other posture category as “seated/lying”.

Additionally, the total amount of time classified as “seated/lying” by the AP was 46.6% compared to the 40.8% measured by DO (mean difference = 5.9%, $p < 0.05$; AUC = 0.84, sensitivity = 87.6%, specificity = 81.0%). When the postures coded as “other” by DO were excluded from the analysis, the AP classified 45.6% of the time as “seated/lying” compared to the 45.2% measured by DO (mean difference = 0.3%, $p = 0.58$; AUC = 0.88, sensitivity = 87.6%, specificity = 88.1%). The AP significantly overestimated the total number of breaks from SB compared to DO when the other posture category was excluded (55 vs. 20, respectively, $p < 0.01$). Reclassifying the other posture category as “seated/lying” and including in the analysis decreased the degree of overestimation but did not change the statistical significance (79 (AP) vs. 55 (DO), $p < 0.01$). The observed degree of overestimation by AP of the total time spent “seated/lying” (5.9%) would result in 3.5 min more time classified as sedentary by AP per every hour compared to DO, which might have practical or clinical significance if extrapolated out for a longer period of time. The primary conclusion was that the AP seems to provide accurate estimates of total time spent in SB compared to DO in young children, but its utility in quantifying breaks from SB is yet unclear.

2.2.2.4. Validation against DO in adults

Grant et al. (50) examined the reliability and validity of the AP against DO for estimating posture and postural transitions in a simulated free-living setting in adults. Ten adults (60% female) performed a controlled section (video recorded standing, sitting, and walking at self-selected speed), and an activity protocol consisting of six ADLs randomly selected from a list of 19 activities. The participants wore three AP monitors on the same leg, one placed mid-thigh, second positioned immediately distally, and third “piggybacked” on top of the second monitor. The participants performed each posture and activity once with different order and duration (2-9 min) assigned for each participant. The total time, time spent in the different postures, and the number of postural transitions were calculated by each observer and compared to those obtained from each AP device. The validity assessment was done by randomly selecting one of the AP devices and comparing the data against one observer. The ICC was excellent (>0.99) for sitting. The mean difference for sitting between DO and the AP device was 0.2% during the controlled section, and 0.3% during the ADL section. The percent agreement, sensitivity, and PPV between the AP and DO for sitting for the controlled time was 98.5%, 99.7%, and 99.8%, respectively, and 93.5%, 99.4%, and 99.5%, respectively, for the ADL time. The number of transitions between sitting and standing were identical between DO and the AP for all tests. The findings suggested that the AP is not only a convenient measure by its small size and ease of use, but it is also a valid and reliable device for estimating sitting time during ADLs in a healthy adult population.

Lyden et al. (35) validated the AP MET output for estimating time in SB in a free-living setting against DO in adults. Thirteen participants (61.5% women) were directly observed for a 10-hr period during three separate days while wearing an AP device on the right thigh. The AP

Professional Research Edition software was used to derive the estimated time spent in SB based on the posture-based MET prediction equation in the software. The MET value assigned for standing was adjusted from the default of 1.4 METs to 1.5 METs. The AP default values for sitting/lying posture (1.2 METs) and stepping (internal algorithm estimated the METs from cadence) were used. The epochs with a sitting/lying posture and an EE less than 1.5 METs (sedentary) were summed together to obtain the total time in SB. The AP estimated time in SB was not significantly different from DO measured time in SB (mean bias = 0.8 min, $p > 0.05$, RMSE = 11.8 min, ICC = 0.99). There were no significant differences between the AP and DO for the total number of sedentary bouts >30 min or >60 min in duration (mean bias = 0.3% and 0.0%, respectively; $p > 0.05$, RMSE = 1.0 min and 0.6 min, respectively, ICC = 0.86 and 0.80, respectively), or time spent in sedentary bouts that were >30 min or >60 min in duration (mean difference = 19.2 min (18.0%) and 6.7 min (5.6%), respectively; $p > 0.05$, RMSE = 65.1 min and 60.2 min, respectively; ICC = 0.84 and 0.78, respectively). The conclusion was that the MET outputs from the AP can be used to estimate SB in healthy adults in free-living settings.

2.2.3. ActiGraph and activPAL

This section includes a brief review of validation studies where SB estimates from both AG and AP devices have been compared against IC and DO in the same study. The section is organized as that first studies with IC as the criterion measure in both youth and adults will be discussed, followed by studies with DO as the criterion measure following the same format.

2.2.3.1. Validation against IC in youth

Crouter et al. (33) validated the AG inclinometer function and SB thresholds and the AP for estimating SB against IC. Forty-four youth (11-15 years) completed 2 h of free-living activity while wearing an AG GT3X+ device on the right hip and an AP device on the right thigh. A K4b² was used as the criterion for EE, and the VO₂ data averaged over one-minute periods and converted to MET_y units (SB was defined as <1.5 MET_y). For the device estimates, SB was estimated as time below the SB thresholds (VM <75 counts/10 sec, VA <25 counts/10 sec, VA <50, <100, <150, and <200 counts/60 sec) and for the inclinometer output when it indicated “seated” or “lying” (AG), or “seated/lying” (AP). The AG inclinometer function had the smallest mean difference compared to IC (0.2 min) but the second highest RMSE (26.3 min) and the highest MAPE (107.1%) of all the estimates. The AP underestimated the K4b² measured sedentary time (mean difference = -1.6 min, RMSE = 28.5 min, MAPE = 85.0%). The mean differences for the SB thresholds ranged from 0.6 min (VA <25 counts/10 sec) to 12.0 min (VA <200 counts/60 sec). Both the MAPE and RMSE increased with increasing 60 sec SB threshold, and only VA <200 counts/60 sec SB estimates were significantly different from the measured SB by IC ($p < 0.05$). VA <25 counts/10 sec had the smallest RMSE (20.6 min), and the VM <75 counts/10 sec had the smallest MAPE (69.2%) of all the estimates. Most methods tended to overestimate lower measured SB time and underestimate higher measured SB time. The researchers suggest further refinement and validation of the AP and AG inclinometer function before using them as criteria for measuring posture in free-living studies as they seemed to perform worse than the commonly used AG SB thresholds. The 10 sec AG thresholds might also provide better estimates of youth SB compared to the 60 sec SB thresholds.

2.2.3.2. Validation against DO in youth

Ridley et al. (47) validated the estimated SB time using AG SB thresholds, AG inclinometer function, and the AP against DO. Sixty-five youth participants (9-12 years) performed their typical classroom activities without researcher intervention other than fitting an AG GT3X device on the right hip and the AP on the right thigh. SB time was estimated using the AG inclinometer output (time spent “sitting” or “lying”), the AP output (time spent “sitting/lying”), and time spent below the AG SB thresholds (VA <13, <25, <38, and <75 counts per 15 sec). The System for Observing Fitness Instruction Time (SOFIT) (151) coding system was used for DO. Of the AG thresholds, VA <75 counts/15 sec had the smallest mean bias (-0.2 min, or 2.4%) and the highest correlation with DO for sitting ($r = 0.80$, $p < 0.01$), but wide 95% LOA [lower limit (LL) -10.5, upper limit (UL) 9.9 min]. The AP mean bias for sitting was larger than that for the best performing AG SB threshold (VA <75 counts/15 sec), but had a stronger correlation and narrower 95% LOA (mean bias = -1.9 min, or 5.6 %, $r = 0.86$, $p < 0.01$; 95%LOA -9.0 to 5.2 min). The results suggest that both devices could be utilized for the purpose of assessing youth SB under free-living conditions. The AP device might provide closer estimates of youth sitting time compared to the AG SB thresholds, but the results also provided evidence for the use of the SB threshold VA <75 counts/15 sec.

Mitchell et al. (46) examined the validity of the SB estimates and sit-to-stand transitions from hip- and wrist-worn AG GT3X and AP devices, compared to DO. Nine overweight and obese (BMI >95th percentile) youth (10-17 years) wore an AG device on the right hip and on the non-dominant wrist, and an AP device on the right anterior thigh during a weight-management session. Data from the hip-worn AG device were analyzed with the inclinometer function and with the LFE filter enabled and disabled using SB threshold VA <25 counts/15 sec. The SB

threshold for the wrist was $VA < 105$ counts/15 sec. The AP (1-sec epochs) and AG (15-sec epochs) were time synchronized with the DO data for analysis at the epoch level for each device. Estimates were considered in agreement with DO if the difference was not more than ± 1 epoch different and data was also collapsed to report values per hour. On epoch-level, the AP had a 96.1% agreement with DO for sit-to-stand transitions and was slightly prone for false positives. On hour-level, the AP slightly overestimated the sit-to-stand transitions (6.8 vs. 5.8 transitions/hours; 17.1% difference. $ICC = 0.26$, $p = 0.27$). On epoch-level, the AG hip; the AG hip with LFE enabled and the AG wrist had similar agreement with DO for the sit-to-stand transitions [agreement range, 85.7 (AG wrist) to 88.0% (AG hip with LFE enabled)], but the AG inclinometer agreement with DO was significantly lower (54.1%). The sensitivity and PPV were low for all the AG estimates and indicated tendency towards false positives. On hour-level, the AG hip, AG hip with LFE enabled, and AG wrist overestimated the sit-to-stand transitions by 189.6 to 202.8%, and the AG inclinometer by 282.6% [ICC range, 0.0 (AG wrist) to 0.12 (inclinometer)]. The results indicate that further investigation is needed to determine the validity of the AP for quantifying breaks from prolonged SB, and that the commonly used hip- and wrist-worn accelerometer SB thresholds may not be valid for measuring sit-to-stand transitions and detecting breaks from SB.

In the previously mentioned study by Crouter et al. (33), the posture estimate by the AP and the AG GT3X+ inclinometer function were compared against measured posture by DO during structured activities in a laboratory setting. Forty-four youth participants (11-15 years) completed a 30 min period of supine rest for the measurement of RMR and an activity protocol consisting of seven structured activities, ranging from sedentary to vigorous intensities (8 min each), while wearing an AG GT3X+ on the right hip and an AP on the right thigh. DO was used

as the criterion for posture, and a postural label was assigned for each activity in the structured protocol based on minutes 5-7. The AG inclinometer output that was “seated” or “lying” and the AP output “seated/lying” were equated as SB. Compared to DO, the AG inclinometer correctly classified supine rest 45% of the time, 26.8% for all lying activities, and 47.8% for all sitting activities. Combining the “lying” and “sitting” categories so that the AG and AP data could be compared resulted in correct classification of 74.6% of all the activities designated as SB by DO. The AP correctly classified 54.7% of all the lying/sitting activities compared to DO. The findings were similar to Carr, et al. (26), Hänggi, et al. (29), and Peterson, et al. (25); and suggest that more studies are needed to examine the validity of the AG inclinometer function, and the AP before solely relying on them as valid measures of SB.

2.2.3.3. Validation against DO in adults

Kozey-Keadle et al. (39) validated AG SB thresholds (VA <50, <100, <150, <200, and <250 counts/60 sec) and AP against DO for estimating SB and reductions in sitting time. Twenty overweight adult participants performed two seven-day conditions (baseline, and an intervention, where the participants were encouraged to reduce sitting time) while wearing an AG GT3X on the right hip and an AP on the right thigh. The total SB from both conditions were determined from the DO records by summing the time spent in lying and sitting postures. VA <100 counts/60 sec underestimated the total sedentary time compared to the measured by DO (mean difference = -16.9 min, or 4.9%). VA <150 counts/60 sec had the lowest bias [mean difference = -0.9 min, or 1.8%), but the other tested thresholds had larger mean biases (mean difference range, -60.0 (VA <50 counts/60 sec) to 32.0 min (VA <250 counts/60 sec); or -22.0 to 17.8%]. The AP underestimated the total sedentary time compared to the measured by DO (mean

difference = -7.7 min, or 2.8%). None of the AG thresholds were sensitive for the reductions in sitting time. Average The AP and DO had a strong correlation for the percent time spent in SB ($R^2 = 0.94$), and the AP was found to be sensitive for detecting the reduction in sitting time between the two conditions ($p < 0.01$). Based on these findings, VA <150 counts/60 sec SB threshold may provide a slightly better estimates of SB compared to the commonly used VA <100 counts/60 sec as the mean difference against DO when using this threshold was minimal. The relatively small mean difference between the AP and DO, and the device's ability to detect reductions in sitting time suggested that it could be used for these purposes in adult free-living studies.

Lyden et al. (38) validated AG GT3X SB thresholds (VA <100 counts/60 sec and VA <150 counts/60 sec) and the AP for quantifying and detecting changes in breaks from SB against DO using a “break rate” ($brk \cdot sed - h^{-1}$). Thirteen physically active adults performed free-living activities for about 10-hrs on two separate days while wearing two AG GT3X on the right hip, one had the LFE filter enabled, and the other operated with the default filter, and an AP on their right thigh. First, the participants performed their normal ADLs and PA (baseline) and were then asked to decrease their time in SB and break up the SB as much as possible with provided guidance and examples (intervention). The measured time spent in SB by DO was 6.3-hrs at baseline, and the absolute number of breaks was 30.9 and the break rate was $5.1 brk \cdot sed - h^{-1}$. The LFE filter VA <100 counts/60 sec was not significantly different from the measured total sedentary time by DO at baseline (mean percent error = 5.6%), but all the other AG estimates were significantly different from DO [mean percent error range, 12.1% (default filter VA <100 counts/60 sec) to 17.1% (default filter VA <150 counts/60 sec); $p < 0.05$]. The measured time spent in SB by DO was 4.5-hrs during the intervention condition, and the absolute

number of breaks was 27.8 and the break rate was $7.3 \text{ rk} \cdot \text{sed} - h^{-1}$. All the AG estimates were significantly different from the measured time in SB by DO during the intervention condition [mean percent error range, 35.9% (LFE filter VA <100 counts/60 sec) to 50.7% (default filter VA <100 counts/60 sec); $p < 0.05$]. There were significant differences between DO and AG for number of sedentary breaks at baseline [mean percent error range, 77.8% (default filter VA <150 counts/60 sec) to 110.7% (LFE filter AG VA <100 counts/60 sec); $p < 0.05$] and after the intervention [mean percent error range, 98.1% (default filter VA <150 counts/60 sec) to 133.3% (LFE filter VA <100 counts/60 sec); $p < 0.05$], and for the break-rate at the baseline [mean percent error range, 53.7% (default filter VA <150 counts/60 sec) to 102.3% (LFE filter VA <100 counts/60 sec); $p < 0.05$] and after the intervention [mean percent error range, 30.4% (default filter VA <150 counts/60 sec after intervention) to 75.8% (LFE filter VA <100 counts/60 sec); $p < 0.05$]. The AP was not significantly different compared to DO for the total time in SB at baseline or after the intervention (mean percent error = 1.6% and 0.1%, respectively); the absolute number of breaks (mean percent error = 0.3% and 10.9%, respectively), or break-rate (mean percent error = 1.0% and 10.9%, respectively). The findings indicated that the AP may provide better estimates of the total sedentary time, absolute number of breaks and the break rate than the selected SB thresholds for a hip-worn AG, which were not sensitive for the changes in the number of breaks from SB when compared to the measured values by DO.

Steeves et al. (48) compared the posture classification accuracy between the AG inclinometer function and AP against DO in a laboratory setting. Twenty-one adults (27% male, 42% Hispanic) wore both devices on the anterior midline of the thigh, midway between the hip and knee joint, placed in a sequence with a slight gap between. One half of the participants wore

the AG device proximally, another half distally. The participants performed an activity protocol that consisted of different sitting and standing postures, walking scenarios, standing while writing on a whiteboard with intermittent steps, and cycling on an upright cycle ergometer. Each activity was performed for 2 min, except for self-paced fixed distance activities. The data collections were directly observed for the start and stop time of each activity and the postural categories. The middle 90 sec of each 2 min laboratory activity were used for the analysis. The AG inclinometer function correctly classified five out of six sitting activities 100% of the time compared to the AP that classified four out of six sitting activities correctly >95% of the time. The classification accuracy for sitting on a laboratory stool was 86% for the AG and only 5% for the AP. The results from this study suggest that, compared to DO, an AG device when worn on the thigh, may have a higher classification accuracy for SB than the AP.

Van Loo et al. (137) compared the estimated SB by the AG GT3X+ and the AP against DO. Fifty-seven children (split into two age groups: 5-8 years, and 9-12 years) wore an AG device on each wrist and an AP on the right thigh while performing a protocol of 15 semi-structured activities (sedentary to vigorous intensities). The SB thresholds used with the wrist AG were VA <1756 counts/60 sec and VM <3958 counts/60 sec (152), VA <161 counts/5 sec, VA <202 counts/5 sec, and VM <305 counts/5 sec (153); and VA <35 counts/5 sec, VA <105 counts/5 sec, VM <100 counts/5 sec, and VM <275 counts/5 sec (27). The SB threshold used with the hip AG was VA <25 counts/15 sec (139). VM <3958 counts/60 sec estimated SB was equivalent to the measured time by DO in both age-groups [mean bias = 4.1% (5-8 years), and -1.7% (9-12 years); ICC range, 0.82-0.85; $p = 0.02$ and $p < 0.001$, respectively]. VA <1756 counts/60 sec was equivalent to DO in the older age-group and the closest estimate in the younger group [mean bias = -1.7% and -6.5% respectively, ICC = 0.86 for both, $p < 0.01$ (9-12

years) and $p = 0.08$ (5-8 years)]. VA <35 counts/5 sec estimated SB was also equivalent with DO in 9-12 years (mean bias = 2.5%; $p < 0.01$); and VM <100 counts/5 sec approached equivalence (mean bias = 5.3%, $p = 0.05$). The hip SB threshold VA <25 counts/15 sec had a greater mean bias compared to the best wrist SB thresholds VA <1756 counts/60 sec and VM <3958 counts/60 sec (mean bias = -15.8% (5-8 years) and -17.8% (9-12 years); ICC = 0.85 for both). The AP was not significantly different from the measured SB by DO for 9-12-year-olds (mean bias = 1.4%). All the SB thresholds in both age groups, wrists, and hip overestimated the number of breaks from SB (mean difference range, 1.8-160.8 for the wrist estimates, $p < 0.05$, and 29.3-33.2 for the hip estimates, $p < 0.01$); except for VM <3758 counts/60 sec in the non-dominant wrist (mean difference = 1.4 (5-8 years), $p = 0.24$; and 1.8 (9-12 years), $p = 0.05$), and the 5 sec thresholds were more biased than the 60 sec thresholds (154.4 vs. 2.9 for 5-8 years, and 129.9 vs. 2.5 for 9-12 years). The AP was better at detecting breaks from SB (8.5 (5-8 years), and 3.2 for (9-12 years); both $p < 0.01$). The results indicated that VA <1756 counts/60 sec and VM <3958 counts/60 sec wrist SB thresholds could be utilized with a similar level of confidence than the hip threshold VA <25 counts/15 sec for estimating time in SB, but in general, the AP may provide better estimation of the sedentary time in children.

2.2.3.4. Conclusion

The summarized studies demonstrated that the AG SB thresholds have been validated against both DO and IC in youth and adults. Fewer studies exist where the AP devices have been validated against IC. Based on the existing literature, the AG SB thresholds seem to generally provide better estimates of SB with smaller error scores against IC. The validity of the AG inclinometer is still unclear due to the contradicting findings, and its utility should be

investigated further. But this observation may simply be a result of the criterion measure utilized for the validation. The AP generally demonstrated smaller error compared to DO for estimated SB time, but more studies where the AP postural output is compared against IC are needed. The findings may also suggest that laboratory studies tend to result in smaller degree of error between the estimated and measured SB compared to free-living studies. Therefore, the validity of objective estimates should be investigated and compared against the criterion measures in both settings.

Conclusion

SB consists of any waking behavior performed in a seated, reclining or lying posture with an EE less than or equal to 1.5 METs. (14, 52). Reviewing the existing literature showed that there is substantial evidence to support the negative, health-deteriorating effects of increased time spent in SB on the risk of morbidity and mortality across the population. Therefore, studying the dose-response relationship between the volume of SB, as well as the effects of breaking up SB on these outcomes becomes relevant.

Quantifying SB can be accomplished with criterion measures, which include DO for posture and IC for EE. Secondary measures, which include both subjective and objective measures, are widely utilized in SB research due to their higher feasibility. Criterion measures are used to validate the secondary measures. Due to the negative health implications associated with SB, it is important that the measures used for quantifying the behaviors are as valid and reliable as possible.

Many self-reports include the assessment of TV viewing and other screen time as the primary indicator of SB (108, 120-125). The validity of the self-report questionnaires is of

unknown quality, and the use of this method may be limited in younger populations due to a greater extent of recall and misclassification bias. Nevertheless, the method may be more practical in large-scale population studies where other measures might be less feasible and increase both the participant and observer burden.

Accelerometer-based devices have gained more ground as potential tools in SB research due to their ease of use and potentially higher accuracy compared to subjective measures. The SB prediction algorithms for accelerometry data have been validated against both postural and EE criterion measures. The reviewed studies suggested high agreement with DO SB time and posture for the AP devices, but the findings were contradicting for the AG devices. These differences may be due to the criterion measure used for the validation. The AP device has only one algorithm that is more often compared against DO than IC as the proprietary algorithm is designed to estimate posture. There is a wide range of SB thresholds and epoch lengths that can be utilized with the AG devices. The cited studies demonstrated that the choice of the criterion measure matters for the validity of the AG estimates of SB. Most validation studies were done in structured laboratory settings, and the AG devices were more often validated against IC, and the AP devices against DO. Changing the criterion measure and moving from laboratory to free-living settings influenced the validity, as was seen in the study by Crouter et al. (33). The effect of the criterion measure used to validate subjective and objective measures is an important factor and may greatly influence the outcome of the validation studies. Because the consensus definition of SB includes the consideration of both the activity EE and posture, validating the secondary measures against the DO and IC together (“combined approach”) would be an ideal practice and provide insights of the secondary measure’s validity to capture true SB.

The search of the existing literature did not result in any studies that had utilized the combined approach to measure SB. A few studies were identified that looked at both the EE and posture together in the same study, but none of them used both postural and EE criterion together (24, 33, 99). A few studies done in adults utilized the Compendium of Physical Activities to derive the intensity level for activities (24, 35), but none used IC as the criterion measure for EE together with the criterion measure for posture. One youth study by Crouter et al. (33) utilized both the K4b² and DO to validate objective measures for estimating SB. The strength of the study was that it was the first one to utilize both criterion measures to validate objective measures in a single study, however, the criterion measures were used separately, not together as the combined approach.

This literature review demonstrated that validation studies do not generally follow the consensus definition of SB and validate objective measures against both the postural and EE criterion measures. That means only one half of the consensus definition of SB is being utilized, and it is unknown how valid the estimates are against both postural and EE criterion measures of SB. This means that there is a gap in the existing literature of how the validity of SB estimates derived from commonly used accelerometer-based devices is affected by the criterion measure used for the validation – whether the estimates are compared against the postural and EE criterion measures independently, or together as the combined approach.

CHAPTER III: MANUSCRIPT

Introduction

Growing evidence suggests that sedentary behavior (SB) is an important independent risk factor for premature morbidity and mortality of all-causes (5, 8-10), and contributes to the increased prevalence of obesity in youth and adults (53, 154-157). Individuals spend a large amount of the waking hours in different SB (58), with motives ranging from leisure [e.g. watching television (TV)] to more “purposeful” or “required” SB (e.g. transportation and occupational). Most evidence about the impact of SB on different health outcomes in children and youth is based on studies of TV watching and other screen-time (158-161). The association between TV watching and increased prevalence of obesity may be due to the biological effects associated with increased time spent sitting, decreased active time, or eating unhealthy foods in front of TV, or some combination of them all. These behaviors can lead to overeating and positive energy balance (161).

According to the Sedentary Behavior Research Network (SBRN), SB is “any waking behavior characterized by energy expenditure (EE) less than or equal to 1.5 metabolic equivalents (METs)”, and is performed in “a sitting, reclining or lying posture” (14). Quantifying SB is important to obtain accurate data of the current trends so that dose-response relationships between SB and different health outcomes can be drawn, and effective intervention programs developed to reduce the time spent in sedentary pursuits. Criterion measures are the most accurate at capturing different components of a behavior, but no single measure is available that can simultaneously capture both posture and EE. Because the SB consensus definition consists of two components, separate criterion measures have been determined for posture [direct observation (DO)] and EE [indirect calorimetry (IC)]. In addition to the activity type and posture, DO may be used to obtain information regarding the contextual variables, such as the physical

and social environment where the activity is being performed. IC is a technique that measures an individual's oxygen consumption (VO_2) and carbon dioxide (CO_2) production, which is used to calculate EE. Using the criterion measures is often less feasible in research studies, and that is why they are used to validate objective measures, such as prediction algorithms for accelerometer-based devices for estimating SB.

ActiGraph [(AG), Pensacola, FL] and activPAL [(AP), Glasgow, Scotland, UK] are commonly used research-grade accelerometer-based devices capable of quantifying SB. Time spent in SB can be estimated from the AG using SB thresholds or the inclinometer function to get an estimate of posture. Typically, the SB thresholds are applied to AG proprietary counts (although they can also be used with the raw acceleration data). Counts that are below a specified SB threshold or an inclinometer output of “lying” and “sitting”. Using counts to estimate sedentary time is based on EE and does not take posture into account; rather assumes that the individual is in a seated or lying position while engaging in the low-EE behaviors (33). Inclinometer measures the distribution of the gravitational acceleration among the three axis of the accelerometer and can be used to estimate time spent in SB based on posture (33). The AP proprietary algorithm uses accelerometer-derived information about thigh position and acceleration to determine body posture and transition between postures (162). The time spent in “seated/lying” posture is equated as SB.

Some studies have used DO as the criterion measure (23-26), others IC (27-30), and the outcomes have varied based on the criterion measure used. Of the AG SB thresholds, the vertical axis (VA) <100 and VA <150 counts/60 sec have been shown to provide the closest estimates of time spent in SB in both youth (23) and adults (25, 26) when DO has been used as the criterion measure of SB. VA <25 counts/10 sec and vector magnitude (VM) <75 counts/10 sec have been

shown to provide closer estimates of SB against IC in youth under free-living conditions (33, 45, 47), and the VA <50 counts/60 sec in adults under free-living conditions (30).

Posture classification by the AP proprietary algorithm has been shown to have high agreement with DO for estimating SB in youth (31, 163) and adults (35, 39, 50). However, in youth, when IC has been used as the criterion measure of SB, the AP has been shown to have greater error for the estimated time in SB (33). The use of the AG inclinometer function to estimate SB is a less studied concept, and studies have reported contradicting results (25, 26, 29, 33). Crouter et al. (33) found the AG inclinometer and AP provided similar estimates of posture in youth during structured laboratory activities when, using DO as the criterion measure. In the same study, IC was used as the criterion measure during free-living conditions, and the AG inclinometer SB estimates had the lowest error, compared to IC, for time spent in SB, but had greater individual error than the AP.

It is possible that differences exist for the measured SB between DO and IC, and that these differences influence the degree and direction of the error in SB estimates derived from the accelerometer-based devices. When either DO or IC is used independently to quantify SB, only one half of the SB consensus definition is being utilized and the other half is being left out of consideration. Using DO as the criterion measure of SB may result in a larger amount of measured SB compared to IC because the method quantifies SB based on the activity posture; and some activities performed in SB posture may in fact have a higher than 1.5 MET EE (22), leading to misclassification of the activity as sedentary. In a recent free-living study in young adults, Marcotte et al. (24) compared the objective estimates of SB against posture and EE together, but the EE values were not measured by IC but were derived from the Compendium of Physical Activities (22). No studies have validated objective measures against the “combined

approach” (DO+IC), i.e. using both measured EE and posture together that align with the SBRN consensus definition of SB (14). It is unclear how the postural and EE criterion measures compare against one another and against the DO+IC method, and how the validity of objectively estimated SB would be influenced by using different criterion measures for the validation. Therefore, the primary purpose of this study was to compare the measured sedentary time by the SB criterion measures (DO, IC, and DO+IC). The secondary purpose was to determine how changing the criterion measure influences the direction and degree of error between the estimated SB by the AG SB thresholds and inclinometer function and the AP inclinometer function, and the measured SB by the criterion measures.

Methodology

1. Participants

This study was a secondary analysis of data collected for a larger study (164, 165). One hundred youth participants aged 6 to 18 years old were recruited from the Knoxville, TN community via flyers (*Appendix C*), and word of mouth. Written informed consent (*Appendix D*) and a prescreening health history questionnaire (*Appendix F*) were obtained from each participant’s parent or legal guardian before enrollment in the study. A written informed assent (*Appendix E*) was obtained from the participant before beginning the data collection. The recruitment process aimed to achieve a sample with an evenly distributed range of ages (grouped as <12 years and ≥ 12 years old) and sexes. Participants with any contraindications to exercise reported by the parent or legal guardian, including musculoskeletal injuries during the past six months that would not permit safe participation in PA; cardiorespiratory conditions, including asthma; metabolic conditions or prescription medications that could influence metabolism, were

excluded from the study. The study methods and procedures were approved by The University of Tennessee, Knoxville Institutional Review Board.

2. Procedures

The study consisted of two laboratory visits at The University of Tennessee, Knoxville Applied Physiology Laboratory. During the first visit, following the informed consent and assent and the pre-participation screening process, anthropometric measurements (height and seated height (cm), body mass (kg), and body fat percentage (%)) were obtained in light clothing with shoes and socks removed. Height measurements were obtained using a mounted portable stadiometer (SECA 213, Hamburg, Germany), and body mass and body fat percentage with a bioelectrical impedance analysis (Tanita BIA BC-418, Tokyo, Japan). Body mass index (BMI) was calculated by dividing the participant's measured body mass in kilograms by their measured height in meters squared. A combination of research-grade and consumer-grade PA monitors were fitted on each participant in different body locations. For the purpose of the subsequent analyses, only the AG GT9X (ActiGraph LLC, Pensacola, FL) worn on the right hip, and the two AP 3C (PALTechnologies Ltd, Glasgow, Scotland, the UK) worn on each thigh (AP_{left} on the left side and AP_{right} on the right side) are reported here. For the measurement of EE, each participant also wore a Cosmed K4b² [(K4b²); Cosmed Inc, Rome, Italy] portable IC, that was mounted on a chest harness provided by the manufacturer. Classification of posture was done with DO using a custom coding scheme created on the Noldus Observer XT software (Noldus Information Technology, Wageningen, The Netherlands). As a secondary record, the laboratory visits were also video recorded using a Canon Vixia HFR700 camcorder (Canon Inc, Ōta, Tokyo, Japan).

During the first laboratory visit, the participants were asked to complete a 30 min period of supine rest in a quiet room to determine their resting metabolic rate (RMR), followed by a semi-structured activity protocol that consisted of eight activities that were randomly selected from a list of 16 activities, five of which were sedentary (*Appendix A – Table 1*). Each activity was performed in duplicate, one bout being 60-90 sec and the other 4-5 min in duration. Participants were given the freedom to select the order in which they wanted to perform the activities with as little limitations as possible, but some constraints were held in place; related to space availability and scheduling issues and optimizing the transition time between activities. The order of the bout duration (i.e. the long or short bout first) was randomized by the research team. At 60 sec or 4 min of an activity, the participant was provided with a verbal prompt to transition to the next activity when they were ready. If the participant did not transition before reaching the maximal bout length (90 sec or 5 min), they were asked to transition to the next activity. The remaining eight activities were performed during the second laboratory visit following the same procedures.

3. Equipment

3.1. PA monitors

ActiGraph GT9X (AG) is a small (3.5 cm x 3.5 cm x 1.0 cm), lightweight (14 g), and water-resistant (1 m, 30 min) triaxial research-grade accelerometer-based device with a programmable display window that can be worn on the hip, wrist, or ankle. The primary triaxial accelerometer measures acceleration with a dynamic range of ± 8 gravitational units (g's) with the sampling frequency ranging from 30 to 100 Hz (in 10 Hz increments). The device also contains an inertial measurement unit (IMU) that samples at 100 Hz, and houses the secondary

triaxial accelerometer with a dynamic range of ± 16 g, a triaxial gyroscope with a dynamic range of $\pm 2,000$ °/s, a magnetometer with a dynamic range of 4,800 μ -Tesla, and a thermometer that measures the sensor temperature.

In this study, the primary accelerometer was set to collect data at 90 Hz, the IMU enabled (100 Hz), and the Idle Sleep Mode disabled. At the beginning of each data collection, the device was initialized and synchronized with an external digital clock so that the start time and activity time stamps could be synchronized with the data from the K4b², DO, and APs. During all testing, the AG was attached to an adjustable nylon belt and worn at waist level on the right anterior axillary line. The raw acceleration data from the AG was exported to a PC using the manufacturer's ActiLife software.

activPAL™ 3C (AP) is a small (35 mm $\times$ 53 mm $\times$ 7 mm), lightweight (15 g), triaxial accelerometer-based device that is typically worn on the anterior midline of the thigh, midway between the inguinal crease and the proximal border of patella. The AP contains a microprocessor, a sensing element, a recording element, associated electronics and a power supply (49). The device has a sampling frequency of 20 Hz and it is intended to be used to detect posture allocation based on the accelerations related to the changes in thigh inclination. The device has the capacity to record and store raw acceleration data up to 14 days.

In the current study, an AP device was attached on each thigh with a non-allergenic adhesive tape. The AP was synchronized with an external digital clock so that the start time and activity time stamps could be synchronized with the data from the K4b², DO, and AG. After each data collection, data from the devices were exported to a PC using the activPAL Professional Research Edition software.

3.2 Indirect calorimetry

Cosmed K4b² is a portable IC that provides breath-by-breath (BxB) measures of $\dot{V}O_2$ and $\dot{V}CO_2$. The K4b² has been reported to be valid for the measurement of EE compared to the Douglas bag (DB) method (78). Prior to each test, the oxygen (O_2) and carbon dioxide (CO_2) analyzers were calibrated according to the manufacturer's instructions. The calibration process consists of a room air calibration, a flow turbine calibration with a 3.00-L Hans-Rudolph syringe (Hans Rudolph Inc., Shawnee, KS), and a reference gas calibration with a known gas mixture (15.93% O_2 , 4.92% CO_2 , N_2 balance); and a delay calibration to adjust for the lag of time that occurs between the measurement of expiratory flow and the gas analyzers. The K4b² was mounted to the participant's chest with a specific harness provided by the manufacturer and connected to a sealed rubber mask with a sampling line that transports gas samples to the analyzers. The system weighs approximately 1.5 kg. Data from the K4b² was exported to a PC using the manufacturer's software and synchronized with the data from DO and the activity monitors.

3.3 Direct observation

Observer XT is a computer program designed for DO and coding of activity behaviors. This can be done both in real-time (i.e. live coding) or in retrospective using video recordings that are uploaded to the program. In the current study, all the coding was done in real-time using the Noldus Pocket Observer application (version 3.2), which is an extension of the Observer XT for Android Tablets. A coding scheme consisting of the 16 activities in the protocol (*Appendix A – Table 1*) plus labels for the start and stop time; activity posture, possible transitions between activities, unknown activity, and social and environmental context were created in the Observer

XT. The coding scheme was then downloaded to the Pocket Observer for real time live coding during the data collections. Following each data collection, the observation records were imported to the Observer XT. As a secondary record, the laboratory visits were also video recorded using a Canon Vixia HFR700 camcorder (Canon Inc, Ōta, Tokyo, Japan). The video records were imported to the Observer XT and synchronized with the manually coded DO records from the Pocket Observer. The DO data were then synchronized with the IC and the activity monitor data from the AG, APs.

4. Data processing

Data were processed using the R computer programming language and software environment (<http://cran.r-project.org/>). Only the long activity bouts (4-5 min) were used in the analysis because the length of the short activity bouts (60-90 sec) were not adequate to reach metabolic steady state. The data was obtained from all the measures and converted into a comparable form so that a second-by-second analysis could be performed.

4.1. Cosmed K4b²

The K4b² was the IC criterion measure for EE. The BxB absolute VO_2 data ($\text{ml} \cdot \text{min}^{-1}$) were imported from the device to the COSMED software. The VO_2 data from the long activity bouts were collapsed into 15 sec epochs by averaging the VO_2 values over the time length using the manufacturer's instructions. The VO_2 data were then exported as a .csv file for further processing in R.

The RMR was calculated from the VO_2 data obtained during the 30 min supine rest by using rolling averages (i.e. the “sliding window” approach), where the first 10 min and last 1 min

were excluded, and the breaths were averaged with any breaths in the preceding five minutes. This process was repeated until the 30 min period was up. As a result, the 5 min period with the lowest mean VO_2 value was used as the measured RMR for the participant. The RMR data were screened for outliers, and participants whose RMR was more than two standard deviations from the mean of their age group (i.e. <12 years or ≥ 12 years old) were removed.

The 15 sec absolute VO_2 data ($L \cdot \min^{-1}$) were converted to relative VO_2 values [$(mL \cdot kg^{-1} \cdot \min^{-1})$; the volume of O_2 consumed by unit of body mass)] by dividing the values by the participant's body mass. To account for the additional mass of the devices worn by the participant, 2 kg were added to each participant's measured body mass for all weight-bearing activities. Youth-METs (MET_y) were then calculated by dividing the measured relative activity VO_2 by the measured RMR. This was done because youth have a higher RMR than adults (73), and using the adult value would result in an overestimation of the EE and activity intensity. The MET_y values were then expanded back out to 1 sec epochs, where the same MET_y values were replicated across all the individual seconds within each 15 sec epoch. After this process, each second of the long activity bouts from both laboratory visits were categorized as sedentary ($MET_y \leq 1.5$) or non-sedentary ($MET_y > 1.5$).

4.2. Direct observation

DO was used as the criterion measure for posture. The DO data were first imported from the Pocket Observer application to the Observer XT software, after which the 1 sec event logs were imported to R for processing.

A custom R script was used to process the DO event logs and separate the long activity bouts from the entire trial for further processing. Each timestamp was analyzed for the activity

and posture labels. In general, activities performed in seated, reclining or lying posture were labeled as sedentary, and activities performed in other postures were labeled as non-sedentary. However, stationary cycling was categorized as a non-sedentary activity due to higher EE ($\text{MET}_y > 1.5$) associated with the activity, regardless of the seated posture.

4.3. Combined approach

The 1 sec IC and DO data were merged together for the “combined approach” (DO+IC) using a custom R script. All the seconds where the MET_y were ≤ 1.5 and the postural label sedentary were categorized as sedentary; the seconds that did not meet both criteria were categorized as non-sedentary.

4.4. ActiGraph GT9X

The 1 sec raw acceleration data from the AG primary accelerometer (90 Hz) and gyroscope (100 Hz) were exported to a PC using the ActiLife software. The data were then converted to 10 sec and 60 sec epochs, and imported to R in .csv files for analysis using four SB thresholds [VM <75 counts/10 sec (28), VA <25 counts/10 sec (28), VA <100 counts/60 sec (166), and VA <150 counts/60 sec] and the inclinometer function. The epochs where the accelerometer counts were below these thresholds were categorized as sedentary, and those where the count thresholds were exceeded were categorized as non-sedentary. The AG inclinometer has four possible outputs: lying, sitting, standing, and non-wear. The seconds where the inclinometer output was either sitting or lying were categorized as sedentary, and those where the output was standing, or non-wear were categorized as non-sedentary.

4.5. activPAL 3C

Data from both the left and right APs were exported to a PC using the ActivPAL Professional Research Edition software. The event files were imported to R and a custom R script was used to transform the event files into 1 sec files. The 1 sec files were then analyzed for the postural categories (sitting/lying, standing, and stepping). The seconds where the AP postural output was sitting/lying were categorized as sedentary, and the seconds where the output was something else (standing or stepping) were categorized as non-sedentary.

4.6. Final dataset

After obtaining and processing data from each measure with each second labeled as sedentary or non-sedentary, all the timestamps were aligned with a Windows system clock. The AG timestamps were aligned automatically during the initialization process. The DO records from the Pocket Observer included the start and stop times that were obtained from a known Windows time. These timestamps were used to calculate the offset between the Android timestamps and Windows system clock, which was then subtracted from all Android timestamps to obtain Windows timestamps. The K4b² timestamps were in a relative form (minutes and seconds since the beginning of the measurement). The relative timestamps were added to the original start time of the measurement recorded from a Windows system clock to calculate a Windows timestamp for each data point.

Participants who completed both laboratory visits and had valid data from all the measures were included in the analyses. Data from all the measures (DO, IC, all the AG comparison measures, and both AP devices) from each laboratory visits were first merged together into two separate Microsoft Excel spreadsheets for each participant. The two laboratory

visits were first examined independently. Timestamps where data from any measure was momentarily missing were removed: missing data from one measure lead to the deletion of the entire data row for that specific timestamp so that data from all the devices could be compared in that timestamp and no values were missing. The number of timestamps categorized as sedentary were summed together in seconds and divided by 60 to obtain the total number of sedentary minutes for both data collection days. The sedentary time from the two separate visits, using only the long activity bouts, were then added together for the total number of sedentary minutes from each measure and participant. These data were then included in the master participant file on Microsoft Excel.

5. Statistical analysis

Statistical analyses were performed with IBM SPSS statistics software version 25.0 (IBM, Armonk, NY). The statistical significance level was set to $p \leq 0.05$.

Initially, between-subjects factors effect of participant's sex or age on the measured SB was tested by performing two one-way repeated measures ANOVAs. In these analyses, the participant's sex (male or female) and age (<12 and ≥ 12 years) were the independent variables, and the measured SB by DO, IC, and DO+IC were the dependent variables.

Subsequently, a one-way repeated measures ANOVA was used to compare the means of the measured SB by the criterion measures (DO, IC, and DO+IC). Subsequently, three one-way repeated measures ANOVAs were performed to compare the criterion measures (DO, IC, and DO+IC) against the objective estimates of SB derived from the AG and AP accelerometer-based devices: first, the measured SB by DO was compared against the objectively estimated SB; then, the measured SB by IC was compared against the objectively estimated SB, and lastly, the

measured SB by DO+IC was compared against the objectively estimated SB. In these three comparisons, the independent variable was the criterion measure (DO, IC, or DO+IC), and the dependent variable was the validity of the objective estimates of SB (i.e. whether or not the differences were statistically significant).

Bland-Altman plots, mean absolute percent error (MAPE), and percent difference between the measured and estimated sedentary time were utilized to assess further agreement. The MAPE was calculated by subtracting the estimated value from the measured value and dividing the outcome by the measured value, times 100. The percent difference was calculated by taking an average of all the individual differences to obtain a group level mean difference.

6. Results

Of the 100 youth who volunteered to participate, 7 withdrew before completing the study, 13 were removed due to issues related to the K4b² on one or both days, and 15 were removed due to missing AG data ($n = 1$) and missing AP ($n = 14$) data on one or both days (*Appendix B – Figure 1*). The final sample included 65 youth (mean $\pm$ SD age: 12.6 ± 3.2 years, 47.7% male; *Appendix A – Table 4*).

6.1. Interaction effect of sex or age

Separate repeated measures ANOVAs for between subject factors controlling for the participant's sex (male or female) and age (grouped as <12 years old and ≥ 12 years old) were performed prior to the primary analysis. No significant interaction effect on the total sedentary time were found for the participant's sex ($F(1, 63) = 1.58, p = 0.214$) or age ($F(1, 63) = 0.472, p = 4.95$), so these factors were not included as potential cofounders in the primary analysis.

6.2. Comparison of the criterion measures

The measured total sedentary time by the criterion measures were significantly different from one another ($F(2, 63) = 64.9, p < 0.001$). The IC was significantly less than that measured by DO+IC (mean difference = -2.5 min, MAPE = 10.1%, $p = 0.017$; *Appendix A – Table 3*). The measured total sedentary time by DO was significantly greater compared to the combined approach (mean difference = 4.5 min, MAPE = 14.7%, $p < 0.001$). The IC measurement of total sedentary time was significantly less compared to the DO measurement (mean difference = -7.0 min, MAPE = 12.9%, $p < 0.001$).

Bland-Altman plots (*Appendix B - Figure 2, A-C*) showed mean biases ranging from -4.5 to 7.0 min. Both the IC and DO had similar, wide limits of agreement with the combined approach, but the IC had the smallest mean bias compared to the combined approach. When the DO and IC measurements were compared, the 95% limits of agreement (95% LOA) were the narrowest [lower limit (LL) -2.7 min, upper limit (UL) 16.7 min], but the mean bias the highest (7.0 min).

6.3. DO as the criterion measure

The estimated total sedentary time by the AP_{left}, VM <75 counts/10 sec, and the VA150 were not significantly different from the measured by DO ($F(1,64) = 0.081, 3.22, \text{ and } 3.56$, respectively; $p = 0.777, 0.077, \text{ and } 0.064$, respectively; mean difference = -0.3, -1.5, and -1.8 min, respectively; MAPE = 12.0, 9.2, and 11.4%, respectively; *Appendix A – Table 4*). The AP_{right}, VA <25 counts/10 sec, and the AG inclinometer function significantly overestimated the total sedentary time by DO ($F(1,64) = 17.46, 52.78, \text{ and } 417.73$; all $p < 0.001$; mean difference = 3.9 min, 6.2 min, and 32.8 min, respectively; MAPE = 12.4%, 13.6%, and 59.3%, respectively).

The VA <100 counts/60 sec significantly underestimated the total sedentary time measured by DO ($F(1,64) = 12.06$, $p = 0.001$; mean difference = -3.1 min, MAPE = 10.8%).

Bland-Altman plots (*Appendix B – Figure 3, A-G*) showed a large negative mean bias for the AG inclinometer, but the mean biases for the other estimates were near 0 min. The inclinometer also had the widest 95% LOA (LL -58.1 min, UL -7.4 min). All the estimates had relatively wide 95% LOA, but the VM <75 counts/10 sec had the narrowest ones (LL -11.4 min, UL 14.3 min).

6.4. IC as the criterion measure

All the objective measures significantly overestimated the measured sedentary time by the IC (*Appendix A – Table 5*), ranging from 3.9 min (VA <100 counts/60 sec; $F(1,64) = 23.83$, $p < 0.001$) to 39.9 min (AG inclinometer; $F(1,64) = 649.29$, $p < 0.001$). The smallest MAPE was for the VA <100 counts/60 sec (12.7%) and the greatest for the AG inclinometer (82.6%).

The Bland-Altman plots (*Figure 3. A-G, Appendix B*) showed mean biases below zero (range, -3.9 min (VA <100 counts/60 sec) to -39.8 min (AG inclinometer). The AG had the widest 95% LOA (LL -64.5 min, UL -15.1 min), and the VM <75 counts/10 sec had the narrowest 95% LOA (LL -16.4 min, UL 5.3 min).

6.5. Combined approach as the criterion measure

The measured SB by DO+IC was compared to the estimated SB by the objective measures (*Appendix A – Table 6*). The estimated sedentary time by the VA <100 counts/60 sec was not significantly different from the measured by DO+IC ($F(1,64) = 1.59$, $p = 0.212$; mean difference = 1.4 min, MAPE = 12.7%). The AP_{left} and AP_{right}; VA <25 counts/10 sec, and the AG

inclinometer function significantly overestimated the measured total sedentary time by DO+IC ($F(1, 64) = 52.87, 59.12, 108.61, \text{ and } 427.45$, respectively; all $p < 0.001$; mean difference = 4.2 min, 8.4 min, 10.7 min, and 37.3 min, respectively; MAPE = 10.2%, 19.3%, 23.5%, and 75.3%, respectively). The VM <75 counts/10 sec and VA150 also significantly overestimated the measured total sedentary time by DO+IC, but the overestimation was smaller compared to the other estimates ($F(1,64) = 8.58 \text{ and } 5.29$, respectively, $p = 0.05 \text{ and } 0.025$, respectively; mean difference = 3.0 min and 1.4 min, respectively; MAPE = 12.6% and 13.6%, respectively).

The Bland-Altman plots (*Appendix B, Figure 4, A-G*) showed negative mean biases (range, -1.4 min (VA <100 counts/60 sec) to -37.3 min (AG inclinometer). The AG inclinometer function had the widest 95% LOA (LL -65.8 min, UL -8.8 min), and the AP_{left} had the narrowest 95% LOA (LL -13.3 min, UL 4.9 min) but several outliers.

Discussion

Studies use DO and IC to validate accelerometer-based devices for estimating SB. Not much is known about the potential differences between these criterion measures for the measured SB and how using only one half of the SBRN SB consensus definition (14) compares against the true measured SB time by DO+IC with both measured posture and EE. This was the first study that utilized the DO+IC method and examined the differences between the criterion measured SB time using all the SB criterion measures, and how utilizing a different criterion measure influences the interpretation of the accelerometer-based estimates of SB. The primary finding from this study was that there were significant differences between all three criterion measures for sedentary time; IC measured the least amount of sedentary time followed by DO+IC and DO only. Typically, DO is only concerned about the postural allocation of a behavior. Due to this, it was expected that the method might result in misclassification of some behaviors as sedentary,

i.e. a greater amount of measured SB than by IC or DO+IC. Some behaviors performed in a sedentary posture may include additional limb movement that results in an EE that is higher than 1.5 METs. On the other hand, some behaviors performed in a non-sedentary posture may have a lower EE than what the posture might indicate, and thus result in misclassified as sedentary. For example, standing quietly in a line has an EE of 1.3 METs according to the Compendium of Physical Activities (22). This behavior would be classified as non-sedentary by DO and sedentary by IC. The DO+IC method balances out these potential misclassifications by considering both posture and EE together. The measured SB by this method in the current study was indeed between the IC and DO measured SB.

The degree and direction of error of the objective estimates of SB were influenced by the criterion measure used. The AP devices are widely utilized in SB research and have been found to provide close estimates of time spent in different postures in youth (31, 34, 163) and adults (35, 50) compared to DO as the criterion measure. In the current study, when DO was the criterion measure the AP worn on the left thigh (AP_{left}) provided the closest estimate of SB with a mean difference of 0.3 min (-0.2%). This finding is similar to a study by Grant et al. (50) that used adult participants and found a mean difference that was similar in magnitude (0.2%) between AP and DO for total time spent sitting. The AP device is typically worn on the anterior aspect of the right thigh (138), but in the current study, one AP device was placed on each thigh. What was interesting and unexpected was that the AP worn on the right thigh (AP_{right}) provided a significantly different estimate of total sedentary time compared to both DO and the AP_{left} (mean difference against DO = 3.1 min, and between the limbs = 4.2 min). The reason for the difference between the limbs is unknown but can be speculated to be related to limb-dominance or bending of one of the legs during certain activities, such as while the RMR was being measured. This

would position the device vertically and lead to misclassification of the activity as non-sedentary (standing). Unlike the AP proprietary algorithms, posture classification by the AG inclinometer is a less studied concept. In a study by Crouter et al. (33), the AG inclinometer provided closer estimates of SB time than the AP when DO was used as the criterion in a laboratory setting. In the current study, the AG inclinometer had the greatest degree of error compared to DO, and it overestimated the measured SB by 33.2 min (-59.3%). Of the AG SB thresholds, previous studies have shown that $VA < 100$ counts/60 sec and $VA < 150$ counts/60 sec may provide the closest estimates of SB compared to DO in youth (23) and adults (25, 26, 43). In the current study, $VA < 150$ counts/60 sec and the $VM < 75$ counts/10 sec estimated SB were the only SB thresholds not significantly different from DO (mean difference = 1.9 and 2.7%, respectively). $VA < 25$ counts/10 sec overestimated by -11.8% and the $VA < 100$ counts/60 sec underestimated by 4.8% the measured SB time by DO. Although the $VA < 100$ counts/60 sec was significantly different from DO, it had a smaller MAPE than the $VA < 150$ counts/60 sec. Peterson et al. (25), using adult participants, had similar overall findings to the current study with $VA < 150$ counts/60 sec providing the closest estimate of SB time compared to DO time, and the inclinometer function had the highest error (66.5% agreement with DO).

When IC was used as the criterion measure, all the estimates were significantly different from the measured value. Previous studies have suggested that $VA < 25$ counts/10 sec and $VM < 75$ counts/10 sec may provide the closer estimates of SB compared to IC in youth (33, 45). In the current study, $VA < 100$ counts/60 sec had the smallest bias of all the estimates against IC (mean difference = 4.1 min, -8.8%; MAPE = 12.7%); but $VM < 75$ counts/10 sec had the narrowest 95% LOA (LL -16.38 min, UL 5.27 min) and the second smallest MAPE and percent difference (13.6% and -12.0%, respectively). In the previously mentioned study by Crouter et al.

(33), the AG SB thresholds and inclinometer and AP SB estimates were also compared against IC under free-living conditions. The mean difference in sedentary time between the AG inclinometer and IC (0.2 min) was lower than the AP and IC (1.6 min), however, the AP had a lower root mean squared error (RMSE) and MAPE. The SB thresholds provided lower individual error for estimating SB time than the inclinometer-based estimates based on the RMSE and MAPE: VA <25 counts/10 sec and VM <75 counts/10 sec had the smallest MAPE, and the VA <25 counts/10 sec had the smallest RMSE of all the estimates. In contrast, the current study found that the AG inclinometer SB estimates had the greatest mean error compared to IC, overestimating the time spent in SB by 39.8 min (82.6%). The AP_{left} and AP_{right} and the VA <25 counts/10 sec also had a high degree of error [mean difference range, 6.7 min (14.5%; AP_{left}) to 13.2 min (28.0%; VA <25 counts/10 sec)].

When DO+IC was the criterion measure, VA <100 counts/60 sec was the only estimate not significantly different from the measured SB time (mean difference = 1.4 min, -4.3%); the other estimates had errors ranging from 2.7 min (VA <150 counts/60 sec) to 37.4 min (AG inclinometer). The AP_{left} had the narrowest 95% LOA (LL -13.3 min, UL 4.9 min) and the smallest MAPE (10.2%), but the 95% LOA were large for all estimates. Because no studies have previously used this method to validate SB estimates from the accelerometer-based devices, the results from the current study cannot be compared to any previous results. However, the results from the current study indicate that there are differences between the criterion measures used to measure SB and it influences the degree and direction of bias of the objective estimates of SB. The error scores against the DO+IC method might give a better idea of how well the estimates represent true SB because the SBRN SB consensus definition (14) includes both sedentary posture and EE. For example, one of the most commonly used SB thresholds VA <100 counts/60

sec was not significantly different from the DO+IC measured SB (mean difference = 1.4 min) but was significantly different from both DO and IC when utilized independently (mean difference = -3.1 and 3.9 min, respectively). However, other estimates (e.g., VA <100 counts/60 sec) appear to perform better at the individual level based on the lower MAPE, compared to the other estimates.

The observed differences between the criterion measures used to measure SB time and the influence these differences have on interpreting the accelerometer-based estimates of SB may have further practical implications. On average, the biases for the SB estimates from the accelerometer-based devices ranged from 2 to 7 min over approximately one hour of measured SB time. Studies have shown that individuals spend about 7-9 hours a day in SB (58, 167). The observed 2-7 min differences between the measured and estimated SB may become practically significant if extrapolated out over 24-hrs and result in almost 1-hr/day difference in the total sedentary time depending on the measure used. This 1-hr difference would translate into even greater differences across weeks, months, and years. Thus, accurate and precise quantification of SB may have important implications for both individual and public health interventions and outcomes, and the seemingly marginal differences could become more significant over time.

Like any study, this study has both strengths and limitations. The study examined the validity of objective measures in estimating sedentary time during a semi-structured protocol that included everyday activities performed both inside and outside of the laboratory setting. The final sample size of 65 youth was relatively large and evenly distributed by sex (52.3% female) and age-groups (<12 years: n = 30; ≥12 years: n = 35). Including both the IC, DO, and IC+DO as the criterion measures in a single study observing SB is another unique feature of this study, and the study provided more information about the differences between criterion measured SB when

the measurements are based on posture, EE, or both. The variety of activities and different environments utilized in the protocol should also improve the generalizability of the findings into free-living settings.

Because this study was a secondary analysis of data collected for a larger study observing PA in youth, the limitations include using an activity protocol that was initially designed for a different purpose than observing SB. Therefore, the protocol may not have included enough sedentary time or a variety of sedentary pursuits and transitions between the different postural allocations for the primary purpose of this study. Other limitations include a homogeneous sample by race and the fact that the participants came from the same geographical area. It is also possible that the participants' RMR was elevated due to performing PA or eating or drinking before reporting to the lab. If the RMR was elevated, the result would be lower calculated MET_y value for the activities, which could explain some of the observed differences.

Conclusion

Using DO and IC independently and together as the combined approach provide different measurements of youth sedentary time. These differences influence the degree of bias and interpretation of the SB estimates from accelerometer-based devices. The combined approach to measure SB includes consideration of both posture and EE together as it has been suggested by the SBRN consensus definition. The current study found evidence to support that validating objective estimates against only one half of the SB consensus definition (either posture or EE, i.e. either against DO or IC) may not be a recommended practice. The observed differences in the measured sedentary time between the criterion measures may influence the validity of the estimated SB time by the accelerometer-based devices and result in data that may not accurately

represent true SB or be comparable across studies that have validated SB time estimates against different criterion measures. Future studies should validate accelerometer-based estimates of SB against the combined approach both in youth and adults. Studies using this method should also be done under free-living conditions as the validity of the estimates under these conditions with this approach is not yet studied.

Acknowledgements

This research was supported by NIH grant R01HD083431.

REFERENCES

1. Martinez-Ramos E, Beltran AM, Martin-Borras C, Lasasa-Medina L, Real J, Trujillo JM, Sola-Gonfaus M, Puigdomenech E, Castillo-Ramos E, Puig-Ribera A, Gine-Garriga M, Serra-Paya N, Rodriguez-Roca B, Gascon-Catalan A, Martin-Cantera C, group S. Patterns of sedentary behavior in overweight and moderately obese users of the Catalan primary-health care system. *PLoS One*. 2018;13(1):e0190750. doi: 10.1371/journal.pone.0190750.
2. Rosenberger M. Sedentary behavior: target for change, challenge to assess. *Int J Obes Suppl*. 2012;2(Suppl 1):S26-S9. doi: 10.1038/ijosup.2012.7.
3. Owen N, Sugiyama T, Healy GN, Dunstan DW. ‘Too much sitting’ and metabolic risk – has modern technology caught up with us? *Eur Endocrinol*. 2010;06:19. doi: 10.17925/EE.2010.06.00.19.
4. Pate RR, Mitchell JA, Byun W, Dowda M. Sedentary behaviour in youth. *Br J Sports Med*. 2011;45(11):906-13. doi: 10.1136/bjsports-2011-090192.
5. Patel AV, Bernstein L, Deka A, Feigelson HS, Campbell PT, Gapstur SM, Colditz GA, Thun MJ. Leisure time spent sitting in relation to total mortality in a prospective cohort of US adults. *Am J Epidemiol*. 2010;172(4):419-29. doi: 10.1093/aje/kwq155.
6. Verloigne M, Van Lippevelde W, Maes L, Yildirim M, Chinapaw M, Manios Y, Androutsos O, Kovacs E, Bringolf-Isler B, Brug J, De Bourdeaudhuij I. Levels of physical activity and sedentary time among 10- to 12-year-old boys and girls across 5 European countries using accelerometers: an observational study within the ENERGY-project. *Int J Behav Nutr Phys Act*. 2012;9:34. doi: 10.1186/1479-5868-9-34.
7. Hu FB, Li TY, Colditz GA, Willett WC, Manson JE. Television watching and other sedentary behaviors in relation to risk of obesity and type 2 diabetes mellitus in women. *JAMA*. 2003;289(14):1785-91. doi: 10.1001/jama.289.14.1785.

8. Rezende LFM, Sa TH, Mielke GI, Viscondi JYK, Rey-Lopez JP, Garcia LMT. All-cause mortality attributable to sitting time: analysis of 54 countries worldwide. *Am J Prev Med*. 2016;51(2):253-63. doi: 10.1016/j.amepre.2016.01.022.
9. Katzmarzyk PT, Church TS, Craig CL, Bouchard C. Sitting time and mortality from all causes, cardiovascular disease, and cancer. *Med Sci Sports Exerc*. 2009;41(5):998-1005. doi: 10.1249/MSS.0b013e3181930355.
10. Dunstan DW, Barr EL, Healy GN, Salmon J, Shaw JE, Balkau B, Magliano DJ, Cameron AJ, Zimmet PZ, Owen N. Television viewing time and mortality: the Australian Diabetes, Obesity and Lifestyle Study (AusDiab). *Circulation*. 2010;121(3):384-91. doi: 10.1161/CIRCULATIONAHA.109.894824.
11. Warren TY, Barry V, Hooker SP, Sui X, Church TS, Blair SN. Sedentary behaviors increase risk of cardiovascular disease mortality in men. *Med Sci Sports Exerc*. 2010;42(5):879-85. doi: 10.1249/MSS.0b013e3181c3aa7e.
12. Hu FB, Leitzmann MF, Stampfer MJ, Colditz GA, Willett WC, Rimm EB. Physical activity and television watching in relation to risk for type 2 diabetes mellitus in men. *Arch Intern Med*. 2001;161(12):1542-8. doi: 10.1001/archinte.161.12.1542.
13. Jakes RW, Day NE, Khaw KT, Luben R, Oakes S, Welch A, Bingham S, Wareham NJ. Television viewing and low participation in vigorous recreation are independently associated with obesity and markers of cardiovascular disease risk: EPIC-Norfolk population-based study. *Eur J Clin Nutr*. 2003;57(9):1089-96. doi: 10.1038/sj.ejcn.1601648.
14. Sedentary Behaviour Research Network. What is sedentary behaviour?
<https://www.sedentarybehaviour.org/what-is-sedentary-behaviour/>: Sedentary Behaviour Research Network; 2017 [cited 2019 September 16].

15. Levine JA. Measurement of energy expenditure. *Public Health Nutr.* 2007;8(7a):1123-32. doi: 10.1079/PHN2005800.
16. Sirard JR, Pate RR. Physical activity assessment in children and adolescents. *Sports Med.* 2001;31(6):439-54. doi: 10.2165/00007256-200131060-00004.
17. Pate RR. Physical activity assessment in children and adolescents. *Crit Rev Food Sci Nutr.* 1993;33(4-5):321-6. doi: 10.1080/10408399309527627.
18. Hausswirth C, Bigard AX, Le Chevalier JM. The Cosmed K4 telemetry system as an accurate device for oxygen uptake measurements during exercise. *Int J Sports Med.* 1997;18(6):449-53. doi: 10.1055/s-2007-972662.
19. Perez-Suarez I, Martin-Rincon M, Gonzalez-Henriquez JJ, Fezzardi C, Perez-Regalado S, Galvan-Alvarez V, Juan-Habib JW, Morales-Alamo D, Calbet JAL. Accuracy and precision of the COSMED K5 portable analyser. *Front Phys.* 2018;9:1764. doi: 10.3389/fphys.2018.01764.
20. Vandarakis D, Salacinski AJ, Broeder CE. A comparison of COSMED metabolic systems for the determination of resting metabolic rate. *Res Sports Med.* 2013;21(2):187-94. doi: 10.1080/15438627.2012.757226.
21. Schrack JA, Simonsick EM, Ferrucci L. Comparison of the Cosmed K4b(2) portable metabolic system in measuring steady-state walking energy expenditure. *PLoS One.* 2010;5(2):e9292. doi: 10.1371/journal.pone.0009292.
22. Ainsworth BE, Haskell WL, Herrmann SD, Meckes N, Bassett DR, Jr., Tudor-Locke C, Greer JL, Vezina J, Whitt-Glover MC, Leon AS. 2011 Compendium of Physical Activities: a second update of codes and MET values. *Med Sci Sports Exerc.* 2011;43(8):1575-81. doi: 10.1249/MSS.0b013e31821ece12.

23. Fischer C, Yildirim M, Salmon J, Chinapaw MJM. Comparing different accelerometer cut-points for sedentary time in children. *Pediatr Exerc Sci*. 2012;24(2):220-8.
24. Marcotte RT, Petrucci GJ, Jr., Cox MF, Freedson PS, Staudenmayer JW, Sirard JR. Estimating sedentary time from a hip- and wrist-worn accelerometer. *Med Sci Sports Exerc*. 2020;52(1):225-32. doi: 10.1249/MSS.0000000000002099.
25. Peterson NE, Sirard JR, Kulbok PA, DeBoer MD, Erickson JM. Validation of accelerometer thresholds and inclinometry for measurement of sedentary behavior in young adult university students. *Res Nurs Health*. 2015;38(6):492-9. doi: 10.1002/nur.21694.
26. Carr LJ, Mahar MT. Accuracy of intensity and inclinometer output of three activity monitors for identification of sedentary behavior and light-Intensity activity. *J Obes*. 2012;2012:9. doi: 10.1155/2012/460271.
27. Crouter SE, Flynn JJ, Bassett DR, Jr. Estimating physical activity in youth using a wrist accelerometer. *Med Sci Sports Exerc*. 2015;47(5):944-51. doi: 10.1249/MSS.0000000000000502.
28. Crouter SE, Horton M, Bassett DR, Jr. Use of a two-regression model for estimating energy expenditure in children. *Med Sci Sports Exerc*. 2012;44(6):1177-85. doi: 10.1249/MSS.0b013e3182447825.
29. Hanggi JM, Phillips LR, Rowlands AV. Validation of the GT3X ActiGraph in children and comparison with the GT1M ActiGraph. *J Sci Med Sport*. 2013;16(1):40-4. doi: 10.1016/j.jsams.2012.05.012.
30. Crouter SE, DellaValle DM, Haas JD, Frongillo EA, Bassett DR. Validity of ActiGraph 2-regression model, Matthews cut-points, and NHANES cut-points for assessing free-living physical activity. *J Phys Act Health*. 2013;10(4):504-14. doi: 10.1123/jpah.10.4.504.

31. Aminian S, Hinckson EA. Examining the validity of the activPAL monitor in measuring posture and ambulatory movement in children. *Int J Behav Nutr Phys Act.* 2012;9:119. doi: 10.1186/1479-5868-9-119.
32. Janssen X, Cliff DP, Reilly JJ, Hinkley T, Jones RA, Batterham M, Ekelund U, Brage S, Okely AD. Validation of activPAL defined sedentary time and breaks in sedentary time in 4- to 6-year-olds. *Pediatr Exerc Sci.* 2014;26(1):110-7. doi: 10.1123/pes.2013-0106.
33. Crouter SE, Hibbing PR, LaMunion SR. Use of objective measures to estimate sedentary time in youth. *J Meas Phys Behav.* 2018;1(3):136-42. doi: 10.1123/jmpb.2018-0007.
34. Davies G, Reilly JJ, McGowan AJ, Dall PM, Granat MH, Paton JY. Validity, practical utility, and reliability of the activPAL in preschool children. *Med Sci Sports Exerc.* 2012;44(4):761-8. doi: 10.1249/MSS.0b013e31823b1dc7.
35. Lyden K, Keadle SK, Staudenmayer J, Freedson PS. The activPAL™ accurately classifies activity intensity categories in healthy adults. *Med Sci Sports Exerc.* 2017;49(5):1022-8. doi: 10.1249/MSS.0000000000001177.
36. Dowd KP, Harrington DM, Donnelly AE. Criterion and concurrent validity of the activPAL professional physical activity monitor in adolescent females. *PLoS One.* 2012;7(10):e47633. doi: 10.1371/journal.pone.0047633.
37. Crouter SE, Churilla JR, Bassett DR, Jr. Estimating energy expenditure using accelerometers. *Eur J Appl Physiol.* 2006;98(6):601-12. doi: 10.1007/s00421-006-0307-5.
38. Lyden K, Kozey Keadle SL, Staudenmayer JW, Freedson PS. Validity of two wearable monitors to estimate breaks from sedentary time. *Med Sci Sports Exerc.* 2012;44(11):2243-52. doi: 10.1249/MSS.0b013e318260c477.

39. Kozey-Keadle S, Libertine A, Lyden K, Staudenmayer J, Freedson PS. Validation of wearable monitors for assessing sedentary behavior. *Med Sci Sports Exerc.* 2011;43(8):1561-7. doi: 10.1249/MSS.0b013e31820ce174.
40. Crouter SE, Kuffel E, Haas JD, Frongillo EA, Bassett DR, Jr. Refined two-regression model for the ActiGraph accelerometer. *Med Sci Sports Exerc.* 2010;42(5):1029-37. doi: 10.1249/MSS.0b013e3181c37458.
41. Treuth MS, Schmitz K, Catellier DJ, McMurray RG, Murray DM, Almeida MJ, Going S, Norman JE, Pate R. Defining accelerometer thresholds for activity intensities in adolescent girls. *Med Sci Sports Exerc.* 2004;36(7):1259-66. doi: 10.1249/01.MSS.0000074670.03001.98.
42. Trost SG, Loprinzi PD, Moore R, Pfeiffer KA. Comparison of accelerometer cut points for predicting activity intensity in youth. *Med Sci Sports Exerc.* 2011;43(7):1360-8. doi: 10.1249/MSS.0b013e318206476e.
43. Ridgers ND, Salmon J, Ridley K, O'Connell E, Arundell L, Timperio A. Agreement between activPAL and ActiGraph for assessing children's sedentary time. *Int J Behav Nutr Phys Act.* 2012;9(1):15. doi: 10.1186/1479-5868-9-15.
44. Trost SG, McIver KL, Pate RR. Conducting accelerometer-based activity assessments in field-based research. *Med Sci Sports Exerc.* 2005;37(11):S531-S43. doi: 10.1249/01.mss.0000185657.86065.98.
45. Crouter SE, Horton M, Bassett DR, Jr. Validity of ActiGraph child-specific equations during various physical activities. *Med Sci Sports Exerc.* 2013;45(7):1403-9. doi: 10.1249/MSS.0b013e318285f03b.

46. Mitchell T, Borner K, Finch J, Kerr J, Carlson JA. Using activity monitors to measure sit-to-stand transitions in overweight/obese youth. *Med Sci Sports Exerc.* 2017;49(8):1592-8. doi: 10.1249/MSS.0000000000001266.
47. Ridley K, Ridgers ND, Salmon J. Criterion validity of the activPAL and ActiGraph for assessing children's sitting and standing time in a school classroom setting. *Int J Behav Nutr Phys Act.* 2016;13:75. doi: 10.1186/s12966-016-0402-x.
48. Steeves JA, Bowles HR, McClain JJ, Dodd KW, Brychta RJ, Wang J, Chen KY. Ability of thigh-worn actiGraph and activPAL monitors to classify posture and motion. *Med Sci Sports Exerc.* 2015;47(5):952-9. doi: 10.1249/MSS.0000000000000497.
49. Why activPAL? <http://www.palt.com/why-activpal/2019> [cited 2019 October 3].
50. Grant PM, Ryan CG, Tigbe WW, Granat MH. The validation of a novel activity monitor in the measurement of posture and motion during everyday activities. *Br J Sports Med.* 2006;40(12):992-7. doi: 10.1136/bjism.2006.030262.
51. The World Health Organization. Obesity and overweight <https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight>: The World Health Organization; 2018 [cited 2019].

52. Tremblay MS, Aubert S, Barnes JD, Saunders TJ, Carson V, Latimer-Cheung AE, Chastin SFM, Altenburg TM, Chinapaw MJM, Altenburg TM, Aminian S, Arundell L, Atkin AJ, Aubert S, Barnes J, Barone Gibbs B, Bassett-Gunter R, Belanger K, Biddle S, Biswas A, Carson V, Chaput J-P, Chastin S, Chau J, ChinApaw M, Colley R, Coppinger T, Craven C, Cristi-Montero C, de Assis Teles Santos D, del Pozo Cruz B, del Pozo-Cruz J, Dempsey P, do Carmo Santos Gonçalves RF, Ekelund U, Ellingson L, Ezeugwu V, Fitzsimons C, Florez-Pregonero A, Friel CP, Fröberg A, Giangregorio L, Godin L, Gunnell K, Halloway S, Hinkley T, Hnatiuk J, Husu P, Kadir M, Karagounis LG, Koster A, Lakerveld J, Lamb M, Larouche R, Latimer-Cheung A, LeBlanc AG, Lee E-Y, Lee P, Lopes L, Manns T, Manyanga T, Martin Ginis K, McVeigh J, Meneguci J, Moreira C, Murtagh E, Patterson F, Rodrigues Pereira da Silva D, Pesola AJ, Peterson N, Pettitt C, Pilutti L, Pinto Pereira S, Poitras V, Prince S, Rathod A, Rivière F, Rosenkranz S, Routhier F, Santos R, Saunders T, Smith B, Theou O, Tomasone J, Tremblay M, Tucker P, Umstattd Meyer R, van der Ploeg H, Villalobos T, Viren T, Wallmann-Sperlich B, Wijndaele K, Wondergem R, on behalf of STCPP. Sedentary Behavior Research Network (SBRN) – terminology consensus project process and outcome. *Int J Behav Nutr Phys Act*. 2017;14(1):75. doi: 10.1186/s12966-017-0525-8.
53. Te Velde SJ, De Bourdeaudhuij I, Thorsdottir I, Rasmussen M, Hagstromer M, Klepp KI, Brug J. Patterns in sedentary and exercise behaviors and associations with overweight in 9-14-year-old boys and girls--a cross-sectional study. *BMC Public Health*. 2007;7:16. doi: 10.1186/1471-2458-7-16.
54. Tremblay MS, Colley RC, Saunders TJ, Healy GN, Owen N. Physiological and health implications of a sedentary lifestyle. *Appl Physiol Nutr Metab*. 2010;35(6):725-40. doi: 10.1139/H10-079.

55. Lajous M, Chavarro J, Peterson KE, Hernandez-Prado B, Cruz-Valdez A, Hernandez-Avila M, Lazcano-Ponce E. Screen time and adiposity in adolescents in Mexico. *Public Health Nutr.* 2009;12(10):1938-45. doi: 10.1017/S1368980009004881.
56. Singh GK, Kogan MD, Van Dyck PC, Siahpush M. Racial/ethnic, socioeconomic, and behavioral determinants of childhood and adolescent obesity in the United States: analyzing independent and joint associations. *Ann Epidemiol.* 2008;18(9):682-95. doi: 10.1016/j.annepidem.2008.05.001.
57. Braithwaite I, Stewart AW, Hancox RJ, Beasley R, Murphy R, Mitchell EA, Group IPTS. The worldwide association between television viewing and obesity in children and adolescents: cross sectional study. *PLoS One.* 2013;8(9):e74263. doi: 10.1371/journal.pone.0074263.
58. Matthews CE, Chen KY, Freedson PS, Buchowski MS, Beech BM, Pate RR, Troiano RP. Amount of time spent in sedentary behaviors in the United States, 2003–2004. *Am J Clin Nutr.* 2008;167(7):875-81. doi: 10.1093/aje/kwm390.
59. Centers of Disease Control and Prevention. Childhood obesity facts <https://www.cdc.gov/obesity/data/childhood.html>: Center of Disease Control and Prevention,; 2018 [cited 2019 September 16].
60. Bai Y, Chen S, Laurson KR, Kim Y, Saint-Maurice PF, Welk GJ. The associations of youth physical activity and screen time with fatness and fitness: the 2012 NHANES National Youth Fitness Survey. *PLoS One.* 2016;11(1):e0148038. doi: 10.1371/journal.pone.0148038.
61. Hancox RJ, Milne BJ, Poulton R. Association between child and adolescent television viewing and adult health: a longitudinal birth cohort study. *Lancet.* 2004;364(9430):257-62. doi: 10.1016/S0140-6736(04)16675-0.

62. Thorp AA, Owen N, Neuhaus M, Dunstan DW. Sedentary behaviors and subsequent health outcomes in adults a systematic review of longitudinal studies, 1996-2011. *Am J Prev Med.* 2011;41(2):207-15. doi: 10.1016/j.amepre.2011.05.004.
63. Stamatakis E, Hamer M, Dunstan DW. Screen-based entertainment time, all-cause mortality, and cardiovascular events: population-based study with ongoing mortality and hospital events follow-up. *J Am Coll Cardiol.* 2011;57(3):292-9. doi: 10.1016/j.jacc.2010.05.065.
64. Wijndaele K, Healy GN, Dunstan DW, Barnett AG, Salmon J, Shaw JE, Zimmet PZ, Owen N. Increased cardiometabolic risk is associated with increased TV viewing time. *Med Sci Sports Exerc.* 2010;42(8):1511-8. doi: 10.1249/MSS.0b013e3181d322ac.
65. Healy GN, Dunstan DW, Salmon J, Cerin E, Shaw JE, Zimmet PZ, Owen N. Breaks in sedentary time: beneficial associations with metabolic risk. *Diabetes Care.* 2008;31(4):661-6. doi: 10.2337/dc07-2046.
66. Hamilton MT, Hamilton DG, Zderic TW. Role of low energy expenditure and sitting in obesity, metabolic syndrome, type 2 diabetes, and cardiovascular disease. *Diabetes.* 2007;56(11):2655-67. doi: 10.2337/db07-0882.
67. Carson V, Janssen I. Volume, patterns, and types of sedentary behavior and cardio-metabolic health in children and adolescents: a cross-sectional study. *BMC Public Health.* 2011;11:274. doi: 10.1186/1471-2458-11-274.
68. Cliff DP, Jones RA, Burrows TL, Morgan PJ, Collins CE, Baur LA, Okely AD. Volumes and bouts of sedentary behavior and physical activity: associations with cardiometabolic health in obese children. *Obesity (Silver Spring).* 2014;22(5):E112-8. doi: 10.1002/oby.20698.

69. Cliff DP, Okely AD, Burrows TL, Jones RA, Morgan PJ, Collins CE, Baur LA. Objectively measured sedentary behavior, physical activity, and plasma lipids in overweight and obese children. *Obesity (Silver Spring)*. 2013;21(2):382-5. doi: 10.1002/oby.20005.
70. Ekelund U, Luan J, Sherar LB, Esliger DW, Griew P, Cooper A, International Children's Accelerometry Database C. Moderate to vigorous physical activity and sedentary time and cardiometabolic risk factors in children and adolescents. *JAMA*. 2012;307(7):704-12. doi: 10.1001/jama.2012.156.
71. Carson V, Stone M, Faulkner G. Patterns of sedentary behavior and weight status among children. *Pediatr Exerc Sci*. 2014;26(1):95-102. doi: 10.1123/pes.2013-0061.
72. Paing AC, McMillan KA, Kirk AF, Collier A, Hewitt A, Chastin SFM. The associations of sedentary time and breaks in sedentary time with 24-hour glycaemic control in type 2 diabetes. *Prev Med Rep*. 2018;12:94-100. doi: 10.1016/j.pmedr.2018.09.002.
73. Schofield WN. Predicting basal metabolic rate, new standards and review of previous work. *Hum Nutr Clin Nutr*. 1985;39 Suppl 1:5-41.
74. McMurray RG, Butte NF, Crouter SE, Trost SG, Pfeiffer KA, Bassett DR, Puyau MR, Berrigan D, Watson KB, Fulton JE. Exploring metrics to express energy expenditure of physical activity in youth. *PLoS One*. 2015;10(6):e0130869. doi: 10.1371/journal.pone.0130869.
75. Butte NF, Watson KB, Ridley K, Zakeri IF, McMurray RG, Pfeiffer KA, Crouter SE, Herrmann SD, Bassett DR, Long A, Berhane Z, Trost SG, Ainsworth BE, Berrigan D, Fulton JE. A Youth Compendium of Physical Activities: activity codes and metabolic intensities. *Med Sci Sports Exerc*. 2018;50(2):246-56. doi: 10.1249/MSS.0000000000001430.

76. Bassett DR, Jr., Howley ET, Thompson DL, King GA, Strath SJ, McLaughlin JE, Parr BB. Validity of inspiratory and expiratory methods of measuring gas exchange with a computerized system. *J Appl Physiol.* 2001;91(1):218-24. doi: 10.1152/jappl.2001.91.1.218.
77. Overstreet BS, Bassett DR, Jr., Crouter SE, Rider BC, Parr BB. Portable open-circuit spirometry systems. *J Sports Med Phys Fitness.* 2017;57(3):227-37. doi: 10.23736/S0022-4707.16.06049-7.
78. McLaughlin JE, King GA, Howley ET, Bassett DR, Jr., Ainsworth BE. Validation of the COSMED K4b2 portable metabolic system. *Int J Sports Med.* 2001;22(4):280-4. doi: 10.1055/s-2001-13816.
79. Crouter SE, LaMunion SR, Hibbing PR, Kaplan AS, Bassett DR, Jr. Accuracy of the Cosmed K5 portable calorimeter. *PLoS One.* 2019;14(12):e0226290. doi: 10.1371/journal.pone.0226290.
80. Winkert K, Kirsten J, Dreyhaupt J, Steinacker JM, Treff G. The COSMED K5 in breath-by-breath and mixing chamber mode at low to high intensities. *Med Sci Sports Exerc.* 2019. doi: 10.1249/MSS.0000000000002241.
81. Rosdahl H, Gullstrand L, Salier-Eriksson J, Johansson P, Schantz P. Evaluation of the Oxycon Mobile metabolic system against the Douglas bag method. *Eur J Appl Physiol.* 2010;109(2):159-71. doi: 10.1007/s00421-009-1326-9.
82. Stroud LC, Feiveson AH, Ploutz-Snyder R, De Witt JK, Everett ME, Gernhardt ML. Comparison of metabolic gas analysis between a standard laboratory system and a portable device. *J Sports Sci Med.* 2009;8(3):491-2.

83. Pinnington HC, Wong P, Tay J, Green D, Dawson B. The level of accuracy and agreement in measures of FEO₂, FECO₂ and VE between the Cosmed K4b2 portable, respiratory gas analysis system and a metabolic cart. *J Sci Med Sport*. 2001;4(3):324-35.
84. Littlewood RA, White MS, Bell KL, Davies PS, Cleghorn GJ, Grote R. Comparison of the Cosmed K4b(2) and the Deltatrac II metabolic cart in measuring resting energy expenditure in adults. *Clin Nutr*. 2002;21(6):491-7.
85. McNaughton LR, Sherman R, Roberts S, Bentley DJ. Portable gas analyser Cosmed K4b2 compared to a laboratory based mass spectrometer system. *J Sport Med Phys Fit*. 2005;45(3):315-23.
86. Duffield R, Dawson B, Pinnington HC, Wong P. Accuracy and reliability of a Cosmed K4b2 portable gas analysis system. *J Sci Med Sport*. 2004;7(1):11-22.
87. Guidetti L, Meucci M, Bolletta F, Emerenziani GP, Gallotta MC, Baldari C. Validity, reliability and minimum detectable change of COSMED K5 portable gas exchange system in breath-by-breath mode. *PLoS One*. 2018;13(12):e0209925. doi: 10.1371/journal.pone.0209925.
88. Beaver WL, Lamarra N, Wasserman K. Breath-by-breath measurement of true alveolar gas exchange. *J Appl Physiol Respir Environ Exerc Physiol*. 1981;51(6):1662-75. doi: 10.1152/jappl.1981.51.6.1662.
89. Roecker K, Prettin S, Sorichter S. Gas exchange measurements with high temporal resolution: the breath-by-breath approach. *Int J Sports Med*. 2005;26 Suppl 1:S11-8. doi: 10.1055/s-2004-830506.
90. Matthews CE. Calibration of accelerometer output for adults. *Med Sci Sports Exerc*. 2005;37(11):S512-S22. doi: 10.1249/01.mss.0000185659.11982.3d.

91. Beijst C, Schep G, Breda E, Wijn PF, Pul C. Accuracy and precision of CPET equipment: a comparison of breath-by-breath and mixing chamber systems. *J Med Eng Technol*. 2013;37(1):35-42. doi: 10.3109/03091902.2012.733057.
92. Proctor DN, Beck KC. Delay time adjustments to minimize errors in breath-by-breath measurement of VO₂ during exercise. *J Appl Physiol*. 1996;81(6):2495-9. doi: 10.1152/jappl.1996.81.6.2495.
93. Darter BJ, Rodriguez KM, Wilken JM. Test-retest reliability and minimum detectable change using the K4b2: oxygen consumption, gait efficiency, and heart rate for healthy adults during submaximal walking. *Res Q Exerc Sport*. 2013;84(2):223-31. doi: 10.1080/02701367.2013.784720.
94. Thomas SS, Buckon CE, Schwartz MH, Russman BS, Sussman MD, Aiona MD. Variability and minimum detectable change for walking energy efficiency variables in children with cerebral palsy. *Dev Med Child Neurol*. 2009;51(8):615-21. doi: 10.1111/j.1469-8749.2008.03214.x.
95. Puhl J, Greaves K, Hoyt M, Baranowski T. Children's Activity Rating Scale (CARS): description and calibration. *Res Q Exerc Sport*. 1990;61(1):26-36. doi: 10.1080/02701367.1990.10607475.
96. Melanna FC, Greg JP, Robert TM, Brittany RM, John S, Patty SF, John RS. A novel video-based direct observation system for assessing physical activity and sedentary behavior in children and young adults. *J Meas Phys Behav*. 2020;3(1):50-7. doi: 10.1123/jmpb.2019-0015.
97. McKenzie T. System for Observing Play and Leisure Activity in Youth (SOPLAY). 2002.

98. McKenzie TL, Cohen DA, Sehgal A, Williamson S, Golinelli D. System for Observing Play and Recreation in Communities (SOPARC): reliability and feasibility measures. *J Phys Act Health*. 2006;3(s1):S208-S22. doi: 10.1123/jpah.3.s1.s208.
99. Lyden K, Petruski N, Staudenmayer J, Freedson P. Direct observation is a valid criterion for estimating physical activity and sedentary behavior. *J Phys Act Health*. 2014;11(4):860-3. doi: 10.1123/jpah.2012-0290.
100. Welch WA, Swartz AM, Cho CC, Strath SJ. Accuracy of direct observation to assess physical activity in older adults. *J Aging Phys Act*. 2016;24(4):583-90. doi: 10.1123/japa.2015-0216.
101. Carmines EG, Zeller RA. Reliability and validity assessment. Beverly Hills, Calif.: Sage Publications; 1979. 70 p. p.
102. Bassett DR, Jr., Rowlands A, Trost SG. Calibration and validation of wearable monitors. *Med Sci Sports Exerc*. 2012;44(1 Suppl 1):S32-8. doi: 10.1249/MSS.0b013e3182399cf7.
103. Klesges RC, Eck LH, Mellon MW, Fulliton W, Somes GW, Hanson CL. The accuracy of self-reports of physical activity. *Med Sci Sports Exerc*. 1990;22(5):690-7.
104. Douwes M, de Kraker H, Blatter BM. Validity of two methods to assess computer use: self-report by questionnaire and computer use software. *Int J Ind Ergon*. 2007;37(5):425-31. doi: <https://doi.org/10.1016/j.ergon.2007.01.002>.
105. Rosenberg DE, Norman GJ, Wagner N, Patrick K, Calfas KJ, Sallis JF. Reliability and validity of the Sedentary Behavior Questionnaire (SBQ) for adults. *J Phys Act Health*. 2010;7(6):697-705.

106. Rosenberg DE, Bull FC, Marshall AL, Sallis JF, Bauman AE. Assessment of sedentary behavior with the International Physical Activity Questionnaire. *J Phys Act Health*. 2008;5 Suppl 1:S30-44. doi: 10.1123/jpah.5.s1.s30.
107. Jefferis BJ, Sartini C, Ash S, Lennon LT, Wannamethee SG, Whincup PH. Validity of questionnaire-based assessment of sedentary behaviour and physical activity in a population-based cohort of older men; comparisons with objectively measured physical activity data. *Int J Behav Nutr Phys Act*. 2016;13:14. doi: 10.1186/s12966-016-0338-1.
108. Gardiner PA, Clark BK, Healy GN, Eakin EG, Winkler EA, Owen N. Measuring older adults' sedentary time: reliability, validity, and responsiveness. *Med Sci Sports Exerc*. 2011;43(11):2127-33. doi: 10.1249/MSS.0b013e31821b94f7.
109. Clark BK, Healy GN, Winkler EA, Gardiner PA, Sugiyama T, Dunstan DW, Matthews CE, Owen N. Relationship of television time with accelerometer-derived sedentary time: NHANES. *Med Sci Sports Exerc*. 2011;43(5):822-8. doi: 10.1249/MSS.0b013e3182019510.
110. Van Cauwenberg J, Van Holle V, De Bourdeaudhuij I, Owen N, Deforche B. Older adults' reporting of specific sedentary behaviors: validity and reliability. *BMC Public Health*. 2014;14:734. doi: 10.1186/1471-2458-14-734.
111. Gennuso KP, Matthews CE, Colbert LH. Reliability and validity of 2 self-report measures to assess sedentary behavior in older adults. *J Phys Act Health*. 2015;12(5):727-32. doi: 10.1123/jpah.2013-0546.
112. Visser M, Koster A. Development of a questionnaire to assess sedentary time in older persons--a comparative study using accelerometry. *BMC Geriatr*. 2013;13:80. doi: 10.1186/1471-2318-13-80.

113. Matthews CE, Keadle SK, Sampson J, Lyden K, Bowles HR, Moore SC, Libertine A, Freedson PS, Fowke JH. Validation of a previous-day recall measure of active and sedentary behaviors. *Med Sci Sports Exerc.* 2013;45(8):1629-38. doi: 10.1249/MSS.0b013e3182897690.
114. O'Neill B, McDonough SM, Wilson JJ, Bradbury I, Hayes K, Kirk A, Kent L, Cosgrove D, Bradley JM, Tully MA. Comparing accelerometer, pedometer and a questionnaire for measuring physical activity in bronchiectasis: a validity and feasibility study. *Respir Res.* 2017;18(1):16. doi: 10.1186/s12931-016-0497-2.
115. Weston AT, Petosa R, Pate RR. Validation of an instrument for measurement of physical activity in youth. *Med Sci Sports Exerc.* 1997;29(1):138-43. doi: 10.1097/00005768-199701000-00020.
116. Donahoe K, Macdonald DJ, Tremblay MS, Saunders TJ. Validation of PiezoRx pedometer derived sedentary time. *Int J Exerc Sci.* 2018;11(7):552-60.
117. Prince SA, LeBlanc AG, Colley RC, Saunders TJ. Measurement of sedentary behaviour in population health surveys: a review and recommendations. *PeerJ.* 2017;5:e4130. doi: 10.7717/peerj.4130.
118. Marshall SJ, Gorely T, Biddle SJ. A descriptive epidemiology of screen-based media use in youth: a review and critique. *J Adolesc.* 2006;29(3):333-49. doi: 10.1016/j.adolescence.2005.08.016.
119. Bryant MJ, Lucove JC, Evenson KR, Marshall S. Measurement of television viewing in children and adolescents: a systematic review. *Obes Rev.* 2007;8(3):197-209. doi: 10.1111/j.1467-789X.2006.00295.x.

120. Dwyer GM, Hardy LL, Peat JK, Baur LA. The validity and reliability of a home environment preschool-age physical activity questionnaire (Pre-PAQ). *Int J Behav Nutr Phys Act.* 2011;8:86. doi: 10.1186/1479-5868-8-86.
121. Salmon J, Campbell KJ, Crawford DA. Television viewing habits associated with obesity risk factors: a survey of Melbourne schoolchildren. *Med J Aust.* 2006;184(2):64-7.
122. Leatherdale S.T. LRE, Faulkner G. Reliability and validity of the physical activity and sedentary behaviour measures in the COMPASS study. COMPASS Technical Report Series [Internet]. 2014 February 29, 2020; 2(1).
123. Gabriel K, McClain J, Schmid K, Storti K, Ainsworth B. Reliability and convergent validity of the past-week Modifiable Activity Questionnaire. *Public Health Nutr.* 2010;14:1-8. doi: 10.1017/S1368980010002612.
124. Salmon J, Owen N, Crawford D, Bauman A, Sallis JF. Physical activity and sedentary behavior: a population-based study of barriers, enjoyment, and preference. *Health Psychol.* 2003;22(2):178-88. doi: 10.1037//0278-6133.22.2.178.
125. Lynch BM, Friedenreich CM, Khandwala F, Liu A, Nicholas J, Csizmadi I. Development and testing of a past year measure of sedentary behavior: the SIT-Q. *BMC Public Health.* 2014;14:899. doi: 10.1186/1471-2458-14-899.
126. Yamamura C, Tanaka S, Futami J, Oka J, Ishikawa-Takata K, Kashiwazaki H. Activity diary method for predicting energy expenditure as evaluated by a whole-body indirect human calorimeter. *J Nutr Sci Vitaminol (Tokyo).* 2003;49(4):262-9. doi: 10.3177/jnsv.49.262.
127. Mensah K, Maire A, Oppert JM, Dugas J, Charreire H, Weber C, Simon C, Nazare JA. Assessment of sedentary behaviors and transport-related activities by questionnaire: a validation study. *BMC Public Health.* 2016;16:753. doi: 10.1186/s12889-016-3412-3.

128. Pulsford RM, Cortina-Borja M, Rich C, Kinnafick FE, Dezateux C, Griffiths LJ. ActiGraph accelerometer-defined boundaries for sedentary behaviour and physical activity intensities in 7-year-old children. *PLoS One*. 2011;6(8). doi: ARTN e21822 10.1371/journal.pone.0021822.
129. Sirard JR, Trost SG, Pfeiffer KA, Dowda M, Pate RR. Calibration and evaluation of an objective measure of physical activity in preschool children. *J Phys Act Health*. 2005;2(3):345.
130. Reilly JJ, Coyle J, Kelly L, Burke G, Grant S, Paton JY. An objective method for measurement of sedentary behavior in 3- to 4-year olds. *Obes Res*. 2003;11(10):1155-8. doi: 10.1038/oby.2003.158.
131. Bassett DR, Troiano RP, McClain JJ, Wolff DL. Accelerometer-based physical activity: total volume per day and standardized measures. *Med Sci Sports Exerc*. 2015;47(4):833-8. doi: 10.1249/MSS.0000000000000468.
132. Chen KY, Janz KF, Zhu W, Brychta RJ. Redefining the roles of sensors in objective physical activity monitoring. *Med Sci Sports Exerc*. 2012;44(1 Suppl 1):S13-23. doi: 10.1249/MSS.0b013e3182399bc8.
133. Troiano RP, McClain JJ, Brychta RJ, Chen KY. Evolution of accelerometer methods for physical activity research. *Br J Sports Med*. 2014;48(13):1019-23. doi: 10.1136/bjsports-2014-093546.
134. Broderick JM, Ryan J, O'Donnell DM, Hussey J. A guide to assessing physical activity using accelerometry in cancer patients. *Support Care Cancer*. 2014;22(4):1121-30. doi: 10.1007/s00520-013-2102-2.
135. Oliver M, Schluter PJ, Healy GN, Tautolo el S, Schofield G, Rush E. Associations between breaks in sedentary time and body size in Pacific mothers and their children: findings

- from the Pacific Islands Families study. *J Phys Act Health*. 2013;10(8):1166-74. doi: 10.1123/jpah.10.8.1166.
136. Ojiambo R, Cuthill R, Budd H, Konstabel K, Casajus JA, Gonzalez-Aguero A, Anjila E, Reilly JJ, Easton C, Pitsiladis YP, Consortium I. Impact of methodological decisions on accelerometer outcome variables in young children. *Int J Obes (Lond)*. 2011;35 Suppl 1:S98-103. doi: 10.1038/ijo.2011.40.
137. Van Loo CM, Okely AD, Batterham MJ, Hinkley T, Ekelund U, Brage S, Reilly JJ, Trost SG, Jones RA, Janssen X, Cliff DP. Wrist accelerometer cut points for classifying sedentary behavior in children. *Med Sci Sports Exerc*. 2017;49(4):813-22. doi: 10.1249/MSS.0000000000001158.
138. Edwardson CL, Winkler EAH, Bodicoat DH, Yates T, Davies MJ, Dunstan DW, Healy GN. Considerations when using the activPAL monitor in field-based research with adult populations. *J Sport Health Sci*. 2017;6(2):162-78. doi: 10.1016/j.jshs.2016.02.002.
139. Evenson KR, Catellier DJ, Gill K, Ondrak KS, McMurray RG. Calibration of two objective measures of physical activity for children. *J Sports Sci*. 2008;26(14):1557-65. doi: 10.1080/02640410802334196.
140. Crouter SE, Oody JF, Bassett DR, Jr. Estimating physical activity in youth using an ankle accelerometer. *J Sports Sci Med*. 2018;36(19):2265-71. doi: 10.1080/02640414.2018.1449091.
141. Crouter SE, Clowers KG, Bassett DR. A novel method for using accelerometer data to predict energy expenditure. *J Appl Physiol*. 2006;100(4):1324-31. doi: 10.1152/jappphysiol.00818.2005.

142. Aguilar-Farias N, Brown WJ, Peeters GM. ActiGraph GT3X+ cut-points for identifying sedentary behaviour in older adults in free-living environments. *J Sci Med Sport*. 2014;17(3):293-9. doi: 10.1016/j.jsams.2013.07.002.
143. Lyden K, Keadle SK, Staudenmayer J, Freedson PS. A method to estimate free-living active and sedentary behavior from an accelerometer. *Med Sci Sports Exerc*. 2014;46(2):386-97. doi: 10.1249/MSS.0b013e3182a42a2d.
144. Hildebrand M, Hansen BH, van Hees VT, Ekelund U. Evaluation of raw acceleration sedentary thresholds in children and adults. *Scand J Med Sci Sports*. 2017;27(12):1814-23. doi: 10.1111/sms.12795.
145. Koster A, Shiroma EJ, Caserotti P, Matthews CE, Chen KY, Glynn NW, Harris TB. Comparison of sedentary estimates between activPAL and hip- and wrist-worn ActiGraph. *Med Sci Sports Exerc*. 2016;48(8):1514-22. doi: 10.1249/MSS.0000000000000924.
146. Bakrania K, Yates T, Rowlands AV, Esliger DW, Bunnewell S, Sanders J, Davies M, Khunti K, Edwardson CL. Intensity thresholds on raw acceleration data: Euclidean Norm Minus One (ENMO) and Mean Amplitude Deviation (MAD) approaches. *PLoS One*. 2016;11(10):e0164045. doi: 10.1371/journal.pone.0164045.
147. Rowlands AV, Yates T, Olds TS, Davies M, Khunti K, Edwardson CL. Wrist-worn accelerometer-brand independent posture classification. *Med Sci Sports Exerc*. 2016;48(4):748-54. doi: 10.1249/MSS.0000000000000813.
148. Staudenmayer J, He S, Hickey A, Sasaki J, Freedson P. Methods to estimate aspects of physical activity and sedentary behavior from high-frequency wrist accelerometer measurements. *J Appl Physiol*. 2015;119(4):396-403. doi: 10.1152/jappphysiol.00026.2015.

149. An HS, Kim Y, Lee JM. Accuracy of inclinometer functions of the activPAL and ActiGraph GT3X+: A focus on physical activity. *Gait Posture*. 2017;51:174-80. doi: 10.1016/j.gaitpost.2016.10.014.
150. Harrington DM, Welk GJ, Donnelly AE. Validation of MET estimates and step measurement using the activPAL physical activity logger. *J Sports Sci*. 2011;29(6):627-33. doi: 10.1080/02640414.2010.549499.
151. McKenzie TL, Smith NJ. Studies of physical education in the United States using SOFIT: a review. *Res Q Exerc Sport*. 2017;88(4):492-502. doi: 10.1080/02701367.2017.1376028.
152. Kim Y, Lee JM, Peters BP, Gaesser GA, Welk GJ. Examination of different accelerometer cut-points for assessing sedentary behaviors in children. *PLoS One*. 2014;9(4):e90630. doi: 10.1371/journal.pone.0090630.
153. Chandler JL, Brazendale K, Beets MW, Mealing BA. Classification of physical activity intensities using a wrist-worn accelerometer in 8-12-year-old children. *Pediatr Obes*. 2016;11(2):120-7. doi: 10.1111/ijpo.12033.
154. Utter J, Neumark-Sztainer D, Jeffery R, Story M. Couch potatoes or french fries: are sedentary behaviors associated with body mass index, physical activity, and dietary behaviors among adolescents? *J Am Diet Assoc*. 2003;103(10):1298-305. doi: 10.1016/s0002-8223(03)01079-4.
155. Tucker LA, Tucker JM. Television viewing and obesity in 300 women: evaluation of the pathways of energy intake and physical activity. *Obesity (Silver Spring)*. 2011;19(10):1950-6. doi: 10.1038/oby.2011.184.

156. Guo C, Zhou Q, Zhang D, Qin P, Li Q, Tian G, Liu D, Chen X, Liu L, Liu F, Cheng C, Qie R, Han M, Huang S, Wu X, Zhao Y, Ren Y, Zhang M, Liu Y, Hu D. Association of total sedentary behaviour and television viewing with risk of overweight/obesity, type 2 diabetes and hypertension: A dose-response meta-analysis. *Diabetes Obes Metab.* 2020;22(1):79-90. doi: 10.1111/dom.13867.
157. Hu J, Ding N, Yang L, Ma Y, Gao M, Wen D. Association between television viewing and early childhood overweight and obesity: a pair-matched case-control study in China. *BMC Pediatr.* 2019;19(1):184. doi: 10.1186/s12887-019-1557-9.
158. Tremblay MS, LeBlanc AG, Kho ME, Saunders TJ, Larouche R, Colley RC, Goldfield G, Connor Gorber S. Systematic review of sedentary behaviour and health indicators in school-aged children and youth. *Int J Behav Nutr Phys Act.* 2011;8:98. doi: 10.1186/1479-5868-8-98.
159. Salmon J, Tremblay MS, Marshall SJ, Hume C. Health risks, correlates, and interventions to reduce sedentary behavior in young people. *Am J Prev Med.* 2011;41(2):197-206. doi: 10.1016/j.amepre.2011.05.001.
160. Gorely T, Biddle SJ, Marshall SJ, Cameron N. The prevalence of leisure time sedentary behaviour and physical activity in adolescent boys: an ecological momentary assessment approach. *Int J Pediatr Obes.* 2009;4(4):289-98. doi: 10.3109/17477160902811181.
161. Lou DW. Sedentary behaviors and youth: current trends and the impact on health Active Living Research2014 [cited 2019 October 3]. Available from: https://activelivingresearch.org/sites/activelivingresearch.org/files/ALR_Brief_SedentaryBehavior_Jan2014.pdf.

162. Myers A, Gibbons C, Butler E, Dalton M, Buckland N, Blundell J, Finlayson G. A novel integrative procedure for identifying and integrating three-dimensions of objectively measured free-living sedentary behaviour. *BMC Public Health*. 2017;17(1):979. doi: 10.1186/s12889-017-4994-0.
163. Davies G, Reilly JJ, Paton JY. Objective measurement of posture and posture transitions in the pre-school child. *Physiol Meas*. 2012;33(11):1913-21. doi: 10.1088/0967-3334/33/11/1913.
164. Hibbing PR, Bassett DR, Coe DP, LaMunion SR, Crouter SE. Youth metabolic equivalents differ depending on operational definitions. *Med Sci Sports Exerc*. 2020. doi: 10.1249/MSS.0000000000002299.
165. LaMunion SR, Blythe AL, Hibbing PR, Kaplan AS, Clendenin BJ, Crouter SE. Use of consumer monitors for estimating energy expenditure in youth. *Appl Physiol Nutr Metab*. 2020;45(2):161-8. doi: 10.1139/apnm-2019-0129.
166. Treuth MS, Schmitz K, Catellier DJ, McMurray RG, Murray DM, Almeida MJ, Going S, Norman JE, Pate R. Defining accelerometer thresholds for activity intensities in adolescent girls. *Med Sci Sports Exerc*. 2004;36(7):1259-66.
167. Du Y, Liu B, Sun Y, Snetselaar LG, Wallace RB, Bao W. Trends in adherence to the Physical Activity Guidelines for Americans for aerobic activity and time spent on sedentary behavior among US adults, 2007 to 2016. *JAMA Netw Open*. 2019;2(7):e197597. doi: 10.1001/jamanetworkopen.2019.7597.

APPENDICES

APPENDIX A: TABLES

Table 1. Description of the activities.

Domain	Activity	Description
Sedentary	Lying	Lying on back arms by the sides in a quiet room without stimuli or further activity.
	Reclining	Sitting in a reclined posture in an office chair with no further activity.
	Browsing Internet	Using a desktop computer to perform self-selected computer activities.
	Computer Gaming	Using a desktop computer to play self-selected computer games.
	Book Reading	Sitting at a tabletop or desk reading a book.
Household	Dusting	Using a rag to clean various surfaces (e.g. desks and counters) in the laboratory.
	Sweeping	Cleaning the floor from paper shreds using a broom and a dustpan.
Walking/running	Slow Overground	Walking at a slow, self-selected pace around the perimeter of a basketball or tennis court.
	Walking	
	Brisk Overground	Walking at a brisk, self-selected pace around the perimeter of a basketball or tennis court.
	Walking	
	Stair Climbing	Ascending and descending a flight of stairs at a self-selected pace.

Table 2. Continued

Walking/running	Overground Running	Running at self-selected pace around the perimeter of a basketball or tennis court.
	Catch	Playing catch with a football.
Sports/exercise	Soccer	Playing soccer (dribbling, kicking and passing the ball, etc.)
	Basketball	Playing basketball (dribbling, bouncing, and passing the ball, throwing baskets, etc.)
	Stationary Cycling	Pedaling a stationary bicycle at a self-selected speed and resistance (minimum of 25 W).
	Jumping jacks	Performing jumping jacks at a self-selected pace.

Table 3. Descriptive characteristics of the participants.

	Males (n = 31)	Females (n = 34)	Overall (N = 65)
Age (yrs; mean $\pm$ SD)	12.4 $\pm$ 2.9	12.8 $\pm$ 3.6	12.6 $\pm$ 3.2
Height (cm; mean $\pm$ SD)	154.0 $\pm$ 19.0	152.2 $\pm$ 15.8	153.1 $\pm$ 17.3
Weight (kg; mean $\pm$ SD)	44.9 $\pm$ 16.5	47.6 $\pm$ 17.3	46.3 $\pm$ 16.8
BMI Classification [n (%)]			
Underweight (5 th percentile)	1 (3.2)	1 (2.9)	2 (3.1)
Normal (5 th to <85 th percentile)	28 (90.3)	26 (76.5)	54 (83.1)
Overweight (85 th to <95 th percentile)	1 (3.2)	5 (14.7)	6 (9.2)
Obese ($\geq$ 95 th percentile)	1 (3.2)	2 (5.9)	3 (4.6)
Body Fat (%; mean $\pm$ SD)	17.9 $\pm$ 4.5	25.6 $\pm$ 7.2	21.9 $\pm$ 7.2
Ethnicity [n (%)]			
Hispanic	1 (3.2)	3 (8.8)	4 (6.2)
Race [n (%)]			
African American	2 (6.5)	4 (11.8)	6 (9.2)
Caucasian	29 (93.5)	30 (88.2)	59 (90.8)

Table 4. Comparison of the measured total sedentary time by the criterion measures.

Measure	Sedentary Time (min) mean $\pm$ SD	Reference Measure: Combined Approach		Reference Measure: DO	
		MAPE (%)	Percent Difference (%)	MAPE (%)	Percent Difference (%)
DO+IC	52.4 $\pm$ 8.4	-	-	-	-
DO	56.9 $\pm$ 6.6*	14.7	-10.2	-	-
IC	49.9 $\pm$ 6.5*#	10.1	3.6	12.9	12.2

Note. Percent difference = [(criterion minus estimate) / criterion $\times$ 100]; DO, direct observation; IC, Cosmed K4b² indirect

calorimeter; DO+IC, combined approach; MAPE mean absolute percent error. Sedentary time from combined approach based on direct observation sedentary posture and K4b² energy expenditure $\leq$ 1.5 youth-METs together.

*Significantly different from DO+IC, $p \leq 0.05$.

#Significantly different from DO, $p \leq 0.05$.

Table 5. Comparison of the measured total sedentary time by DO and the estimated by the objective measures.

		Reference Measure: DO		
	Measure	Total Sedentary Time (min), mean $\pm$ SD	MAPE (%)	Percent Difference (%)
Posture	DO	56.9 $\pm$ 6.6	-	-
	AP _{left}	56.6 $\pm$ 8.7	12.0	-0.2
	AP _{right}	60.8 $\pm$ 6.9*	12.4	-7.7
	AG Inclinometer	89.7 $\pm$ 12.4*	59.3	-59.3
Counts	VA <25 counts/10 sec	63.1 $\pm$ 6.4*	13.6	-11.8
	VM <75 counts/10 sec	55.4 $\pm$ 6.2	9.2	1.9
	VA <100 counts/60 sec	53.8 $\pm$ 6.9*	10.8	4.8
	VA <150 counts/60 sec	55.1 $\pm$ 8.0	11.4	2.7

Note. Percent difference = [(criterion minus estimate) / criterion x 100]; DO, direct observation; AP_{left}, activPAL worn on the left thigh; AP_{right}, activPAL worn on the right thigh; VA, vertical axis; VM, AG Inclinometer, ActiGraph GT9X inclinometer function. MAPE, mean absolute percent error.

*Significantly different from DO, $p \leq 0.05$.

Table 6. Comparison of the measured total sedentary time by IC and the estimated by the objective measures.

		Reference Measure: IC		
	Measure	Total Sedentary Time (min), mean $\pm$ SD	MAPE (%)	Mean Percent Difference (%)
Posture	IC	49.9 $\pm$ 6.5	-	-
	AP _{left}	56.6 $\pm$ 8.7*	17.4	-14.5
	AP _{right}	60.8 $\pm$ 6.9*	23.2	-23.0
	AG Inclinometer	89.7 $\pm$ 12.4*	82.6	-82.6
Counts	VA <25 counts/10 sec	63.1 $\pm$ 6.4*	28.4	-28.0
	VM <75 counts/10 sec	55.4 $\pm$ 6.2*	13.6	-12.0
	VA <100 counts/60 sec	53.8 $\pm$ 6.9*	12.7	-8.8
	VA <150 counts/60 sec	55.1 $\pm$ 8.0*	14.7	-11.4

Note. Percent difference = [(criterion minus estimate) / criterion x 100]; IC, Cosmed K4b² indirect calorimeter; AP_{left}, activPAL worn on the left thigh; AP_{right}, activPAL worn on the right thigh; VA, vertical axis; VM, AG Inclinometer, ActiGraph GT9X inclinometer function. MAPE, mean absolute percent error.

*Significantly different from IC, $p \leq 0.05$.

Table 7. Comparison of the measured total sedentary time by the combined approach and the estimated by the objective measures.

		Reference Measure: DO+IC		
	Measure	Total Sedentary Time (min), mean $\pm$ SD	MAPE (%)	Mean Percent Difference (%)
Posture	DO+IC	52.4 $\pm$ 8.4	-	-
	AP _{left}	56.6 $\pm$ 8.7*	10.2	-8.4
	AP _{right}	60.8 $\pm$ 6.9*	19.5	-17.7
	AG Inclinometer	89.7 $\pm$ 12.4*	75.3	-74.6
Counts	VA <25 counts/10 sec	63.1 $\pm$ 6.4*	23.5	-22.3
	VM <75 counts/10 sec	55.4 $\pm$ 6.2*	12.6	-7.3
	VA <100 counts/60 sec	53.8 $\pm$ 6.9	12.7	-4.3
	VA <150 counts/60 sec	55.1 $\pm$ 8.0*	13.6	-6.7

Note. Percent difference = [(criterion minus estimate) / criterion x 100]; DO+IC, combined approach (DO, direct observation + IC, Cosmed K4b² indirect calorimeter); MAPE mean absolute percent error. Sedentary time from Combined Approach based on DO sedentary posture and IC energy expenditure $\leq$ 1.5 youth-METs together.

*Significantly different from DO+IC, $p \leq 0.05$.

APPENDIX B: FIGURES

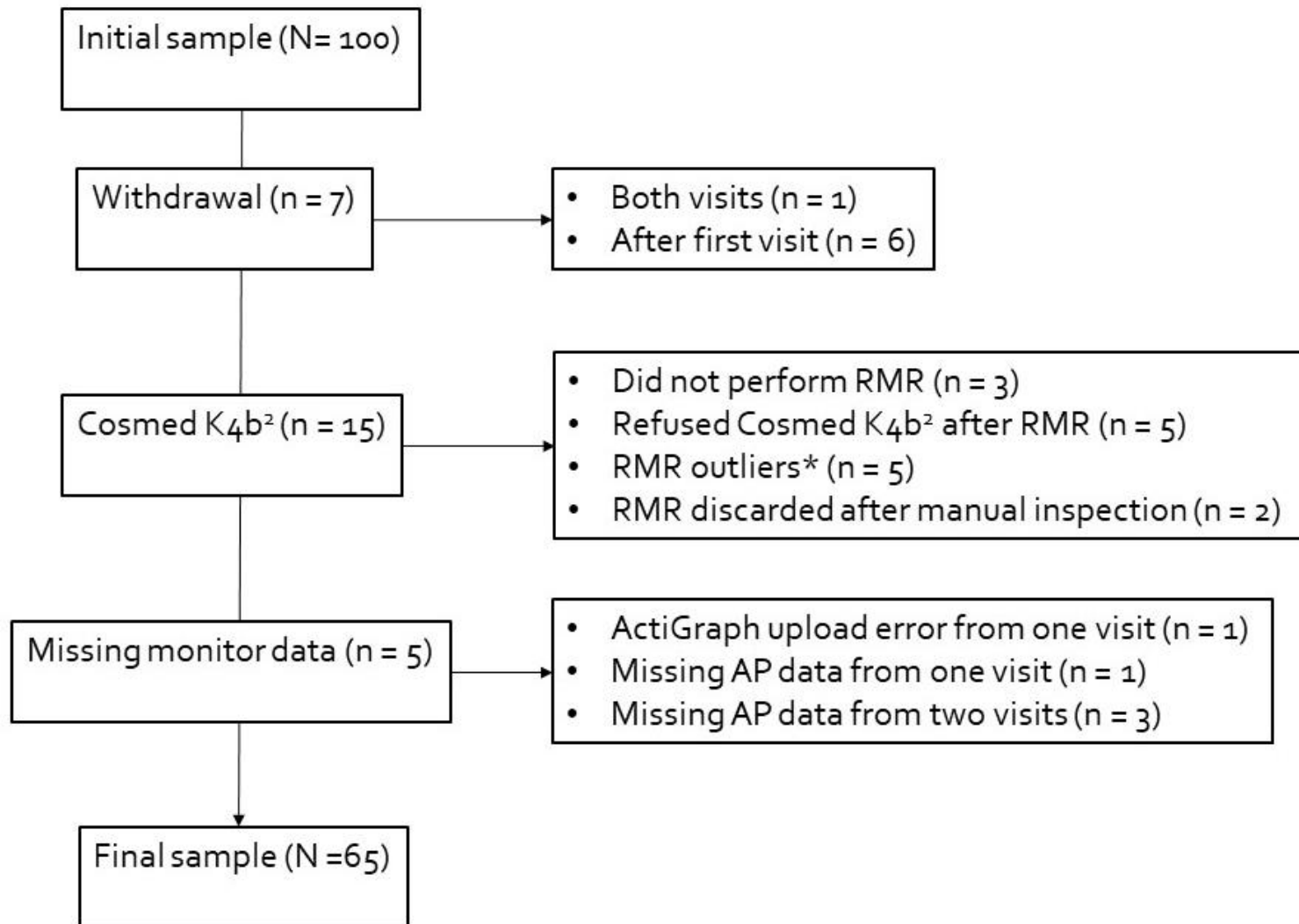


Figure 1. Flow chart to describe the process of reaching the final sample.

*>2SD from the mean for those <12 years old and for those ≥12 years old.

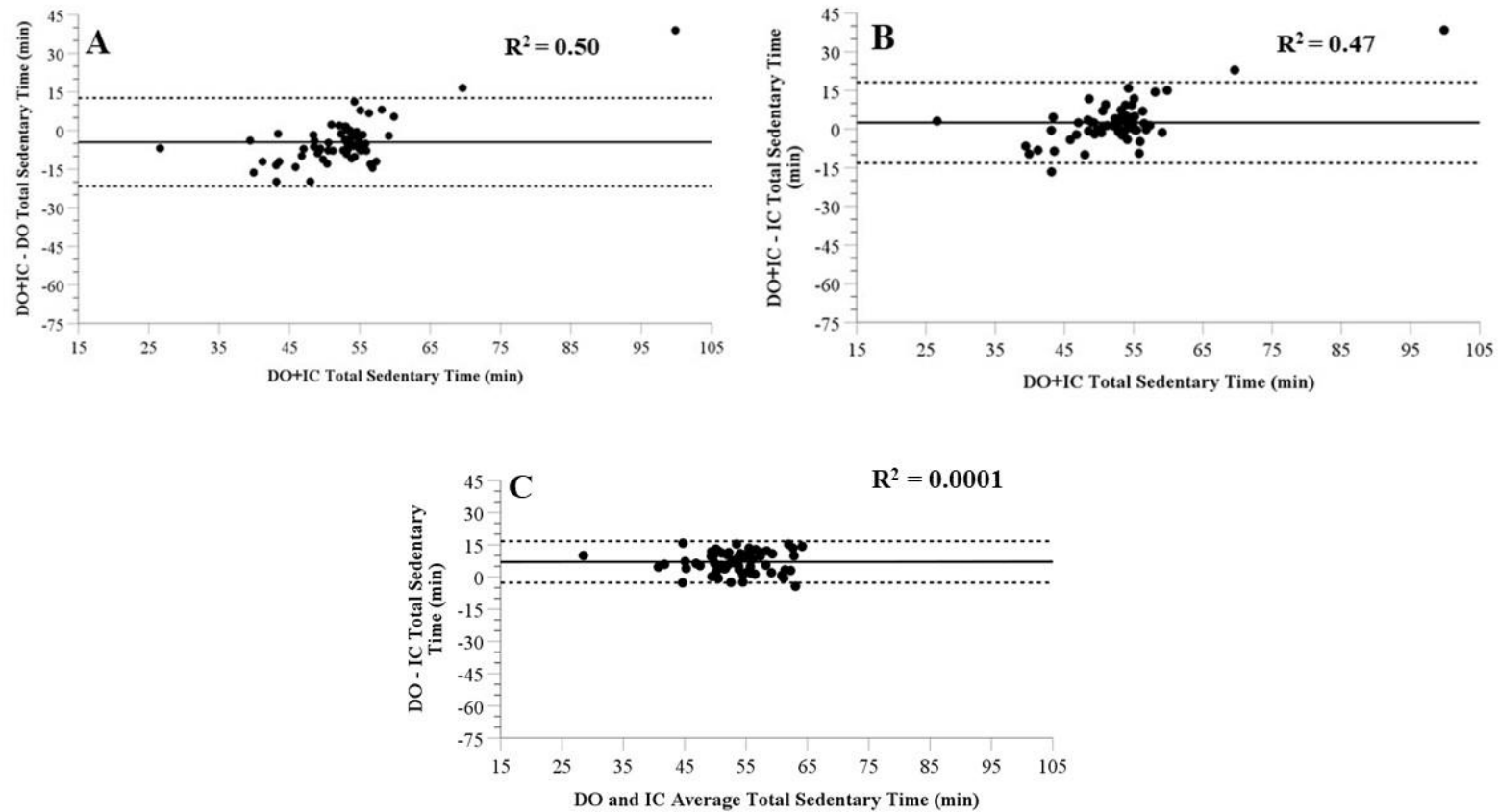


Figure 2. Bland-Altman plots depicting error scores (measured minus estimation) against the combined approach for the time spent in SB during the two-day activity protocol using A) direct observation (DO), B) indirect calorimetry (IC), i.e. the Cosmed K4b², and C) DO and IC vs. combined approach (DO+IC). The solid line represents the mean and the dashed line the 95% prediction interval.

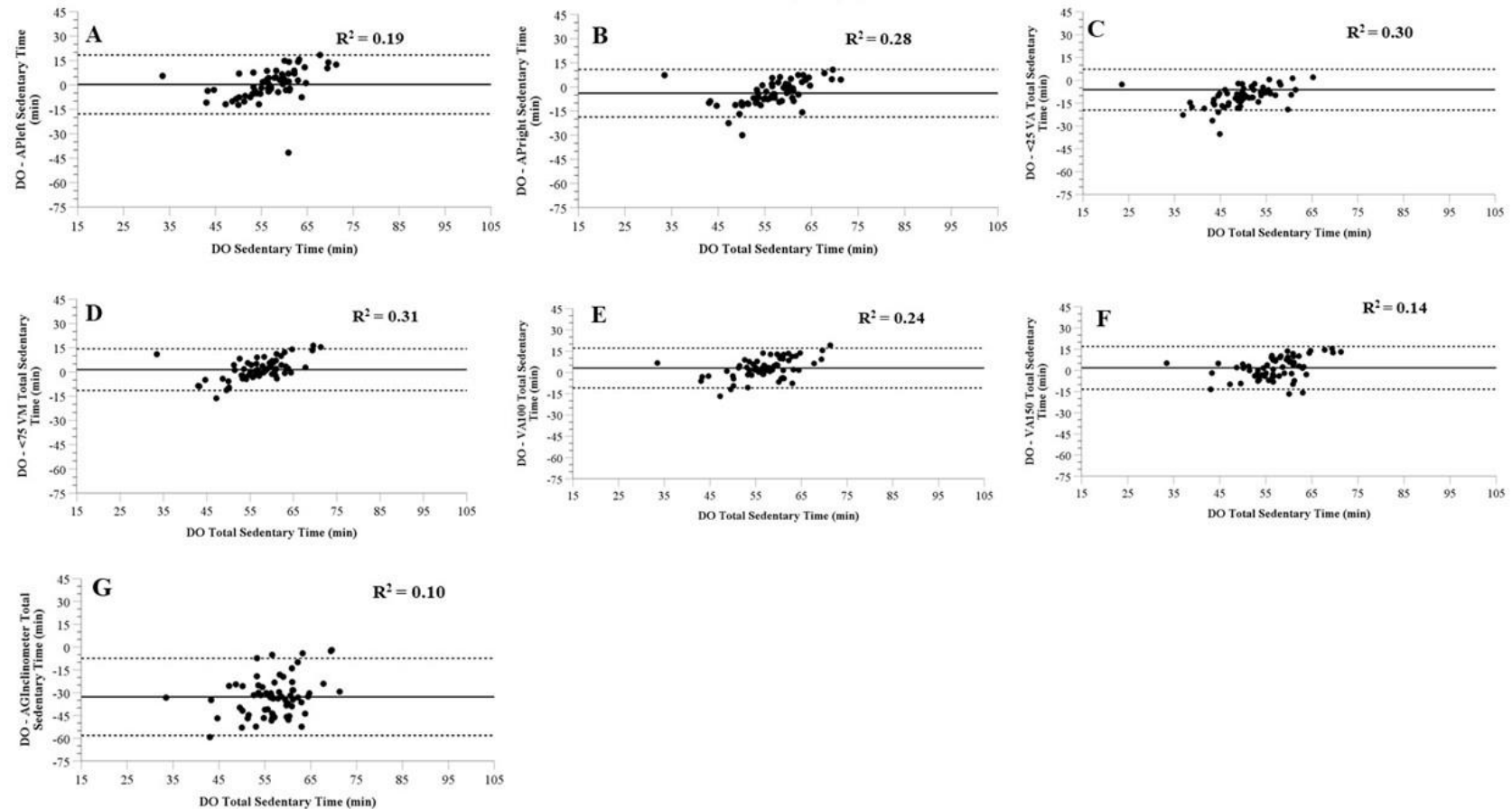


Figure 3. Bland-Altman plots depicting error scores (measured minus estimation) against DO for the time spent in SB during the two-day activity protocol using A) the activPAL on the left thigh (AP_{left}), B) the activPAL on the right thigh (AP_{right}), C) vertical axis (VA) <25 counts/10 sec, D) vector magnitude (VM) <75 counts/10 sec, E) VA <100 counts/60 sec, F) VA <150 counts/60 sec, and G) the ActiGraph inclinometer function (AG_{inclinometer}). The solid line represents the mean and the dashed line the 95% prediction interval.

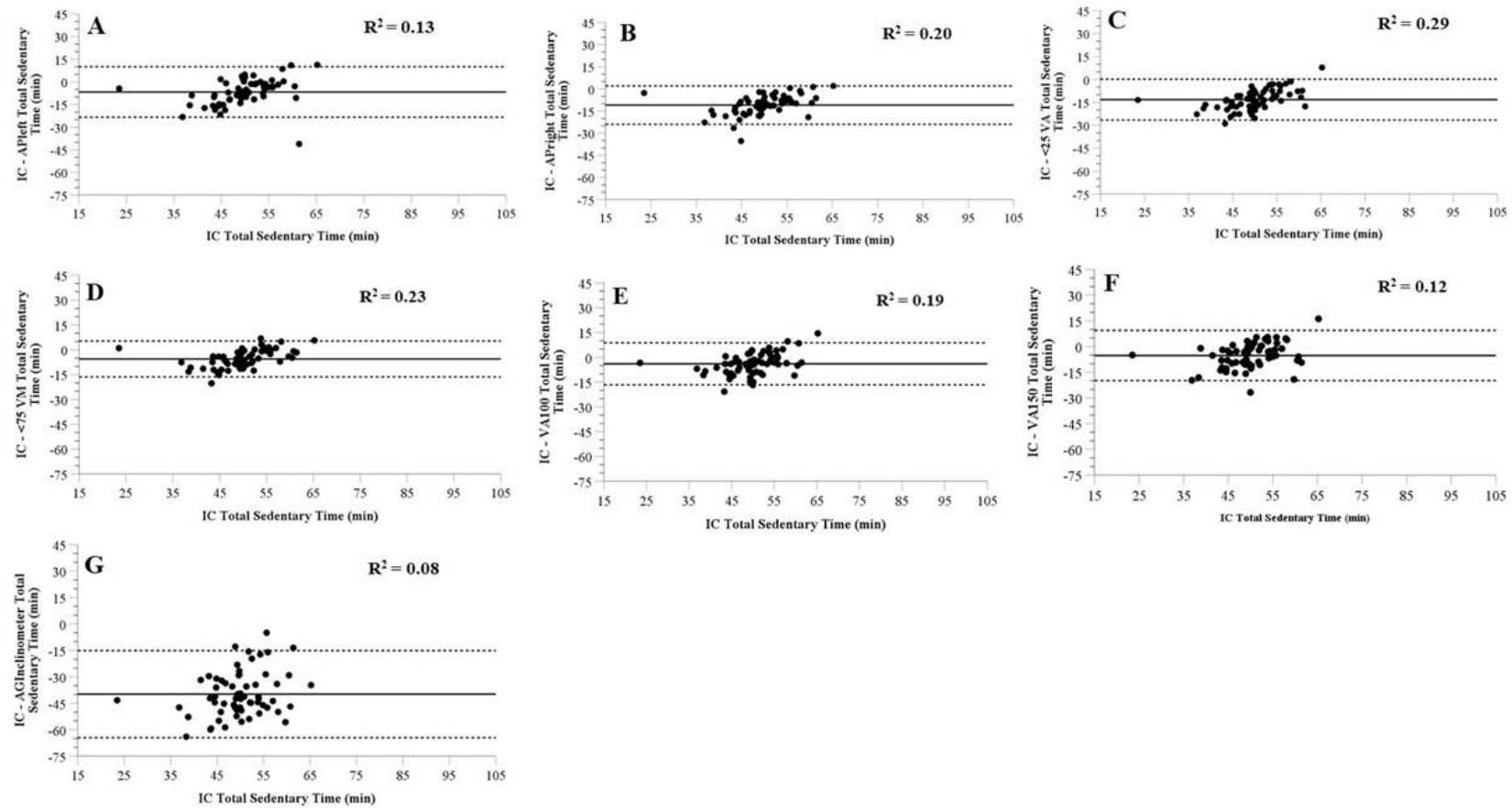


Figure 4. Bland-Altman plots depicting error scores (measured minus estimation) for the time spent in SB during the two-day activity protocol using indirect calorimetry (IC), i.e. the Cosmed K4b² and A) the activPAL on the left thigh (AP_{left}), B) the activPAL on the right thigh (AP_{right}), C) vertical axis (VA) <25 counts/10 sec, D) vector magnitude (VM) <75 counts/10 sec, E) VA <100 counts/60 sec, F) VA <150 counts/60 sec, and G) the ActiGraph inclinometer function (AG_{inclinometer}). The solid line represents the mean and the dashed line the 95% prediction interval.

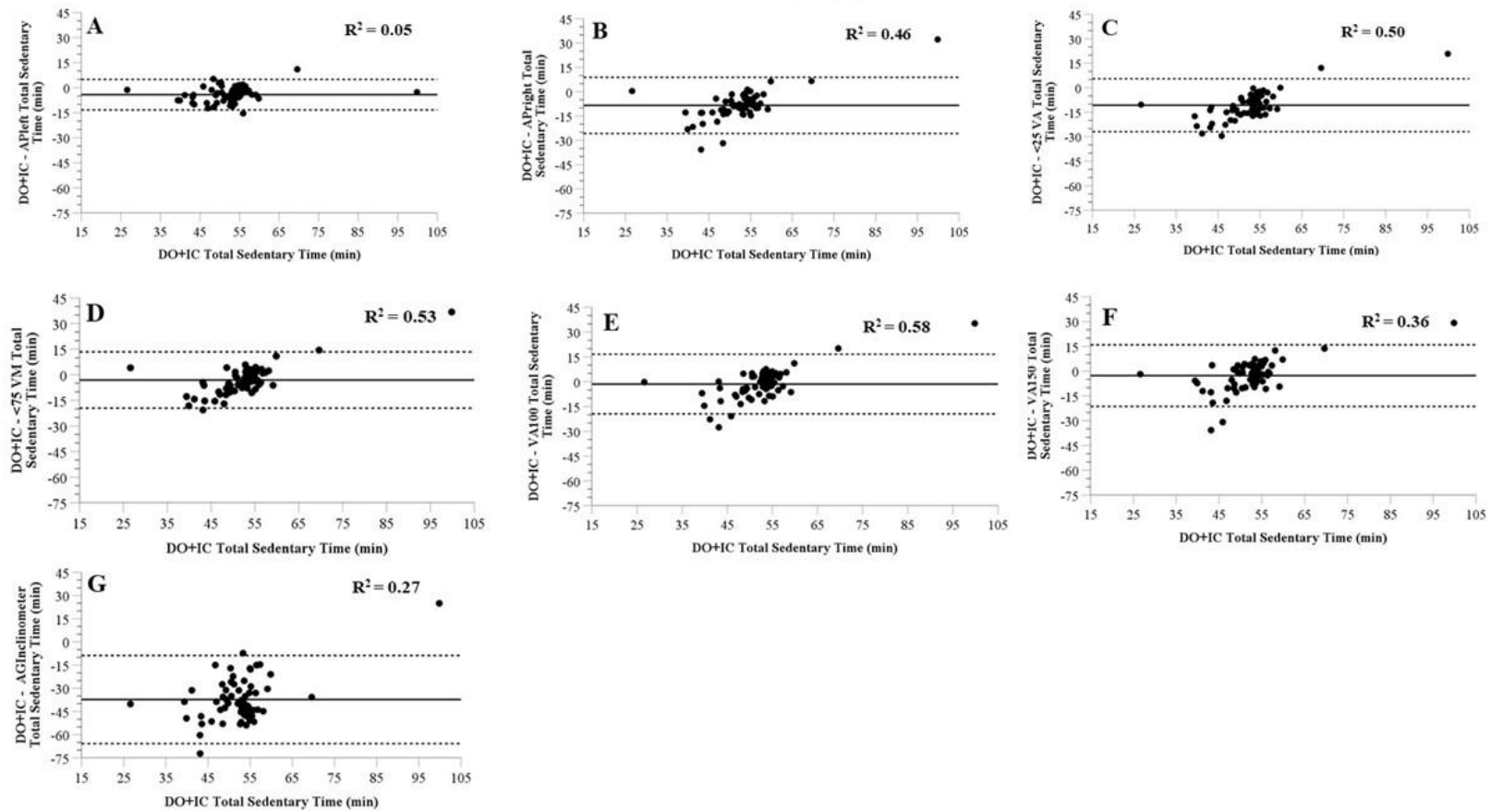


Figure 5. Bland-Altman plots depicting error scores (measured minus estimation) for the time spent in SB during the two-day activity protocol using the combined approach (DO+IC) and A) the activPAL on the left thigh (AP_{left}), B) the activPAL on the right thigh (AP_{right}), C) vertical axis (VA) <25 counts/10 sec, D) vector magnitude (VM) <75 counts/10 sec, E) VA <100 counts/60 sec, F) VA <150 counts/60 sec, and G) the ActiGraph inclinometer function (AG_{inclinometer}). The solid line represents the mean and the dashed line the 95% prediction interval.

APPENDIX C: RECRUITMENT FLYER

Seeking Research Participants

Study Details

- » Perform common daily activities, including household chores and sports, while wearing activity monitors
- » Two visits (~2.5 hours total)
- » Compensation up to \$75
- » Excellent opportunity for research and technology exposure!



Who can Participate?

- » Boys and girls ages 6-18 years old
- » Active or non-active youth eligible
- » All testing will take place at the University of Tennessee campus



**Are you Interested?
Contact us below**

Email: APLStudy@utk.edu

Phone: (865) 974-5651

APPENDIX D: INFORMED CONSENT

Informed Consent Statement for Parent or Guardian

Title of Research: Novel Approaches for Predicting Unstructured Short Periods of Physical Activities in Youth: Study 1 – Development of New Techniques

Principal Investigator: Dr. Scott Crouter

Location: Applied Physiology Laboratory, 1914 Andy Holt Ave., University of Tennessee, Knoxville, TN 37996

PURPOSE

You are being asked to give permission for your child to participate in a research project that will examine new methods to measure the amount of physical activity youth obtain using small devices called accelerometers. These devices are attached to a belt which is worn around the waist, wrist, or ankle similar to a pedometer. Your child will be eligible if he or she is: 1) 6 to 18 years of age, and 2) you and your child can read and speak English. Your child will not be eligible if: 1) he or she has a serious health problem or injury that would prevent them from participating safely in regular physical activity, 2) he or she has a Pacemaker or other internal medical device.

The study will take place at the Applied Physiology Laboratory in the Health, Physical Education, and Recreation (HPER) building at The University of Tennessee. The lead researcher is Scott Crouter, PhD, Assistant Professor in the Department of Kinesiology, Recreation, and Sport Studies. Please read this form and feel free to ask questions. If you have further questions, Dr. Crouter will discuss them with you. You can contact him by telephone (617-287-7509) or email (scrouter@utk.edu).

PARTICIPANT'S INVOLVEMENT IN THIS STUDY

Here is what the study involves:

- 1) You will be asked to fill out a health history questionnaire pertaining to your child's health status to determine eligibility for the study.
- 2) Your child will be asked to visit the Applied Physiology Laboratory in the HPER building on 2 separate days for approximately 100 minutes on visit 1 and 60 minutes on visit 2, for a total time commitment of 160 minutes (2 hrs 40 minutes).
 - a. During visit 1 your child will have his/her height, weight, percent body fat measured, resting metabolic rate measured, and will perform 8 activities randomly selected from a list of 16 activities.
 - b. During visit 2 your child will perform the remaining 8 activities, from the list of 16 that they did not perform during visit 1. The 2 testing days will occur within 2 weeks of each other.
- 3) The percent body fat will be measured using bio-electrical impedance (BIA). BIA is a non-invasive technique that involves standing on a scale while also holding on to a handlebar. During this time an undetectable and harmless electrical current is passed through the body.
- 4) The resting metabolic rate test will take 35 minutes. This test will involve having your child rest in a lying position for 25 minutes and it will take 10 minutes to connect the equipment and prepare for the test. During this time he/she will have 3 accelerometers placed on their waist, 3 on each wrist, 2 on each ankle, and 1 on each shoe using an elastic belt. Another device will also be attached to each thigh using a sticky pad. Your child will also be attached to a Cosmed portable system, which will measure the air that they breathe. This device is light weight and straps onto the body. It also requires the use of a facemask, over the nose and mouth. While wearing the facemask your child will be able to breathe normally and talk, and the mask does not block or restrict airflow.

IRB NUMBER: UTK IRB-15-02487-FB
IRB APPROVAL DATE: 08/16/2018
IRB EXPIRATION DATE: 09/13/2019

_____ Participant's Initials

- 5) Your child will perform 16 different activities. During each visit 8 activities will be performed which include:
- 1) self-paced slow walk, 2) self-paced brisk walk, 3) self-paced running, 4) self-paced stair climbing, 5) shooting a basketball, 6) playing catch with a ball, 7) playing soccer, 8) sitting in a reclined position, 9) sitting upright while reading, 10) sitting upright while searching the internet, 11) sitting upright while playing computer games, 12) lying, 13) cycling at a moderate pace, 14) performing jumping jacks, 15) sweeping, and 16) dusting
- a. While performing the activities your child will be wearing the elastic belts with the accelerometers and the Cosmed metabolic cart as described above. Each visit your child will receive a list of 8 activities from which to choose from and they will be instructed to perform each activity for two times; once for between 60-90 seconds and once for between 4-5 minutes. A trained research assistant will work with your child to ensure they perform the activities correctly.
 - b. While performing the activities a trained research assistant will also video record your child so that exact transition times between activities can be obtained. This is an essential part of this project for development of better tracking methods in real-world situations.

RISKS

There are minimal risks associated with participation in the research study. Since the study involves children engaging in physical activities, there is the possibility of minor injuries such as cuts, bruises, and muscle soreness, and a slight risk of sprains. The risk is similar to that involved in a physical education class or a community sport program. There is also a risk of skin irritation from wearing the sticky pad on the thigh, which will be minimized by cleaning the area after it is removed. If irritation does occur during testing it will be removed immediately. Additionally, there is also a potential for social and psychological risks associated with wearing the Cosmed K4b² and being followed during the activities. Lastly, there is a risk of the identity of your child being known through the use of the video recording.

BENEFITS

The benefits of the proposed research include improved methods for measuring free-living physical activity. This has implications for intervention researchers and those looking at the health of children and adolescents. Participants will learn about their general health characteristics (e.g. BMI) and strategies to become more active.

CONFIDENTIALITY

Information and records included in this study will be kept confidential. Data will be stored in a secure location and will only be made available to the people conducting the study, unless you specifically give permission in writing to do otherwise. The video recordings will be kept on a secure password protected server and only investigators involved with this study will have access to these files. The results of the study will be published, but no reference will be made in oral or written reports that could link participants to the study.

The data files will also be sent to study investigators at the University of Massachusetts Boston and the University of California Riverside. Only non-identifiable data will be sent to these investigators and they will not receive the videos of your child, thus they would not be able to identify who your child is. The Office of Human Research Protections (OHRP) may also review the study records at any time.

COMPENSATION

For participation in the study, your child will receive a \$30 gift card for completion of visit 1 and a \$45 gift card for completion of visit 2, for a total of \$75 in gift cards if both days are completed.

IRB NUMBER: UTK IRB-15-02487-FB
IRB APPROVAL DATE: 08/16/2018
IRB EXPIRATION DATE: 09/13/2019

_____ **Participant's Initials**

EMERGENCY MEDICAL TREATMENT

The University of Tennessee does not "automatically" reimburse subjects for medical claims or other compensation. If physical injury is suffered in the course of research, or for more information, please notify the investigator in charge, Dr. Scott Crouter, at (865) 974-1272.

CONTACT INFORMATION

If you have questions at any time about the study or the procedures (or you experience adverse effects as a result of participating in this study), you should immediately contact the principal investigator, Dr. Scott Crouter, 334 HPER Building, The University of Tennessee, Knoxville, TN 37996, (865) 974-1272. If you have questions concerning your rights as a participant, contact the Office of Research Compliance Officer at (865) 974-7697.

PARTICIPATION

Your participation in this study is voluntary; you may decline to participate without penalty. If you decide to participate, you may withdraw from the study at any time without penalty and without loss of benefits to which you are otherwise entitled. If you withdraw from the study before data collection is completed your data will be returned to you or destroyed.

STATEMENT OF CONSENT

I have read the above information, and I have received a copy of this form. I consent for my child to participate in this study. I also certify that I am 18 years of age or older.

Printed name of your child _____

Parent or Legal Guardian signature _____ Date _____

Investigator's signature _____ Date _____

STATEMENT OF VIDEO CONSENT

I consent for my child to be videotaped while completing the activity described above. Not signing, does not exclude your child from the study and they can still participate, however they will not be videotaped.

Parent or Legal Guardian signature _____ Date _____

I consent for my child's videotape to be used for training purposes in the future. This might include use for demonstration in an undergraduate or graduate kinesiology course or training for future studies where investigators are using the same procedures. By not signing below, your child can still be videotaped for use in the current study, but there file will not be used for future training or educational purposes.

Parent or Legal Guardian signature _____ Date _____

IRB NUMBER: UTK IRB-15-02487-FB
IRB APPROVAL DATE: 08/16/2018
IRB EXPIRATION DATE: 09/13/2019

APPENDIX E: INFORMED ASSENT

Children's Assent form

Title of Research: Novel Approaches for Predicting Unstructured Short Periods of Physical Activities in Youth: Study 1 – Development of New Techniques

Principal Investigator: Dr. Scott Crouter

Location: Applied Physiology Laboratory, 1914 Andy Holt Ave., University of Tennessee, Knoxville, TN 37996

PURPOSE

You are being asked to participate in a research project that will use small devices called accelerometers to measure how much you move. These devices are attached to a belt and can be worn around the waist, hip, or ankle. The study will take place at the Applied Physiology Laboratory in the Health, Physical Education, and Recreation (HPER) building at The University of Tennessee. The lead researcher is Scott Crouter, PhD, Assistant Professor in the Department of Kinesiology, Recreation, and Sport Studies. Please read this form and feel free to ask questions. If you have further questions, Dr. Crouter or a research assistant will discuss them with you. You can contact him by telephone (617-287-7509) or email (scrouter@utk.edu).

PARTICIPANT'S INVOLVEMENT IN THIS STUDY

If you decide that you want to be in this study, this is what you will be asked to do:

- 1) Visit the Applied Physiology Laboratory in the HPER building on 2 different days for approximately 1.5 hours on visit 1 and 1 hour on visit 2.
- 2) Have your height, weight, and percent body fat measured, which requires you to stand barefoot on a scale.
 - a. During your 2 visits you will be asked to do the following tests:
 - b. Resting metabolic rate (visit 1): During this test you will be asked to rest quietly while lying on a table. During the test you will have 3 small devices placed on your waist, 3 on each wrist, 2 on each ankle, and 1 on each shoe using an elastic belt. Another device will also be attached to each thigh using a sticky pad. You will also have devices placed on your shoes and You will also be asked to wear a mask over your nose and mouth which will be connected to a device that is worn on your chest. The mask over your face will not affect how you breathe and you will be able to talk. This test will take 35 minutes to complete.
- 3) Perform 16 different activities (visits 1 and 2): During each visit 8 activities from the list below will be performed:
 - 1) walk slowly, 2) walk quickly, 3) run as fast or as slow as you want, 4) walk up and down stairs, 5) play basketball, 6) play catch with a ball, 7) playing soccer, 8) sit in a reclined position, 9) sit in a chair and read, 10) sit in a chair and play on the internet, 11) sit in a chair and play computer games, 12) lay on your back, 13) indoor cycling, 14) jumping jacks, 15) sweeping a floor, and 16) dusting a table
 - a. While performing the activities you will be asked to wear the elastic belts facemask as you will do for the resting metabolic measurement test. You will do each activity two for either 60-90 seconds or 4-5 minutes. A research assistant will work with you to perform each activity.
 - b. While doing the activities a research assistant will also video record what you are doing.

CAN ANYTHING BAD HAPPEN TO ME?

Although there are very few risks to taking part in this study, you could get hurt just like with any type of physical activity such as what you do in PE class. Dr. Crouter and the other research team members will help you to avoid being hurt or upset. There is also the risk that your skin will become itchy while wearing the sticky pad on your thigh.

IRB NUMBER: UTK IRB-15-02487-FB
IRB APPROVAL DATE: 08/16/2018
IRB EXPIRATION DATE: 09/13/2019

_____ Participant's Initials

CAN ANYTHING GOOD HAPPEN TO ME?

By participating in this study, you will learn about other physical activities that will be good for your health while also having fun.

DO I HAVE OTHER CHOICES?

You can choose not to be in this study if you don't want to.

WILL ANYONE KNOW I AM IN THE STUDY?

We won't tell anyone you took part in this study. Other people who are in the Applied Physiology Laboratory will know that you are being tested but we will not tell them any of your personal information. When we are done with the study, we will write a report about what we found out. We won't use your name in the report.

WHAT WILL YOU RECEIVE?

You will receive a \$30 gift card for completing visit 1 and a \$45 gift card for completing visit 2. If you complete both days of testing you will get a total of \$75 in gift cards.

WHAT IF I DO NOT WANT TO DO THIS?

You don't have to be in this study. It's up to you. If you say yes now, but you change your mind later, that's okay too. All you have to do is tell us.

Before you tell us your decision about being a part of this study; be sure to ask Dr. Crouter or one of his research assistants to tell you more about anything you don't understand.

STATEMENT OF CONSENT

I have read the above information, and I have received a copy of this form. By signing this form I give assent to participate in this study.

Participant's signature _____ Date _____

Investigator's signature _____ Date _____

STATEMENT OF VIDEO CONSENT

I give assent to be videotaped while completing the activity described above. Not signing, does not exclude you from the study and you can still participate, however you will not be videotaped.

Participant's signature _____ Date _____

I give assent for my videotape to be used for training purposes in the future. This might include use for demonstration in an undergraduate or graduate kinesiology course or training for future studies where investigators are using the same procedures. By not signing below, you can still be videotaped for use in the current study, but there file will not be used for future training or educational purposes.

Participant's signature _____ Date _____

IRB NUMBER: UTK IRB-15-02487-FB
IRB APPROVAL DATE: 08/16/2018
IRB EXPIRATION DATE: 09/13/2019

APPENDIX F: HEALTH HISTORY QUESTIONNAIRE

HEALTH HISTORY QUESTIONNAIRE

Today's Date: _____

Name of parent/legal guardian completing the form: _____

ABOUT THE PARTICIPANT

Please answer the following questions about your child.

Name: _____

Address: _____

City: _____ Zip Code: _____

Parent's Phone: _____ Date of Birth (month/day/year): _____

Age: _____ Gender: ____ M ____ F

Current School (if summer, School entering in Fall): _____

Current Grade (if summer, Grade entering in the Fall): _____

How do you identify your child?

- ____ Asian, Non-Hispanic
- ____ Asian, Hispanic
- ____ Black/African American, Non-Hispanic
- ____ Black/African American, Hispanic
- ____ Native Hawaiian/Pacific Islander, Non-Hispanic
- ____ Native Hawaiian/Pacific Islander, Hispanic
- ____ Native American/Alaskan, Non-Hispanic
- ____ Native American/Alaskan, Hispanic
- ____ White, Non-Hispanic
- ____ White, Hispanic

Is this child Multiracial? ____ Y ____ N

Who resides in the household with your child?

____ Both parents ____ Single Mother ____ Single Father ____ Grandparent(s) ____ Other

Please Turn Over

Health History

Has your child ever been diagnosed with any of the follow conditions? If yes, please explain.

	NO	YES	Current Condition	Explanation
Heart Disease				
Diabetes/High Blood Sugar				
High Blood Pressure				
Seizure				
Asthma or Other Lung Condition				
Cancer				
Musculoskeletal Condition or Injury				
Other Serious Illness				
Allergy				

Please describe any additional medical conditions that may affect your child's participation in physical activity.

Is your child taking any medication (including prescription and non-prescription? If yes, please state below.

YES _____ NO _____

Name of Medication

Reason for Taking

For How Long?

Emergency Contact

Name: _____

Relationship: _____ Phone: Work: _____

Home: _____

VITA

Minttu Karoliina Hukka was born in March 31, 1995 to Tero and Mari Hukka in Kotka, Finland. She was the valedictorian of her graduating class in the Kotkan Lyseon Lukio high school in May 2014. Upon graduation, she accepted an NCAA Division I track and field scholarship in Lamar University, in Beaumont TX. After her first year as dietetics major and a scholar student-athlete, she transferred to Boise State University in Boise, ID where she competed for three years for the Broncos and earned both academic and athletic All-American honors before graduating with a Bachelor of Science in Kinesiology, Exercise Science with Summa Cum Laude and the Boise State Athletics Academic Excellence Award in May 2018. In Fall of 2018, Minttu enrolled to the Exercise Physiology Master's program at The University of Tennessee, Knoxville, where she also finished her last year of NCAA eligibility with the Lady Volunteers. In Tennessee, Minttu became interested in studying the wearable devices and monitoring of physical activity and exercise and worked as a research assistant for Dr. Scott E. Crouter. Minttu graduated with a Master of Science in Kinesiology, Exercise Physiology with Summa Cum Laude in May 2020. She is currently looking for career opportunities in the field of applied exercise physiology.