Title: Wrist accelerometer cut-points for classifying sedentary behavior in children.

Christiana M.T. van Loo (corresponding author)[1]

University of Wollongong, NSW, 2522, Australia

E: cmtvl646@uowmail.edu.au

T: +61 2 4239 2274

Anthony D. Okely[1], Marijka J. Batterham[2], Trina Hinkley[3], Ulf Ekelund[4,5], Søren Brage[5], John J. Reilly[6], Stewart G. Trost[7], Rachel A. Jones[1], Xanne Janssen[6], Dylan P. Cliff[1]

1. Early Start Research Institute and Illawarra Health and Medical Research Institute, University of Wollongong, Australia

2. School of Mathematics and Applied Statistics, University of Wollongong, Australia

3. Deakin University, Geelong, Australia, Institute for Physical Activity and Nutrition (IPAN), School of Exercise and Nutrition Sciences

4. Norwegian School of Sports Sciences, Norway

5. MRC Epidemiology Unit, University of Cambridge, United Kingdom

6. University of Strathclyde, School of Psychological Sciences and Health, Scotland

7. Institute of Health and Biomedical Innovation at Queensland Centre for Children's Health Research, School of Exercise and Nutrition Science, Queensland University of Technology

Abstract

Introduction: To examine the validity and accuracy of wrist accelerometers for classifying sedentary behavior (SB) in children.

Methods: Fifty-seven children (5-8y and 9-12y) completed a ~170min protocol including 15 semi-structured activities and transitions. Nine ActiGraph (GT3X+) and two GENEActiv wrist cut-points were evaluated. Direct observation was the criterion measure. The accuracy of wrist cut-points was compared to that achieved by the ActiGraph hip cut-point (≤25 counts/15s) and the thigh-mounted activPAL3™. Analyses included equivalence testing, Bland-Altman procedures and area under the receiver operating curve (ROC-AUC).

Results: The most accurate ActiGraph wrist cut-points (Kim, vector magnitude: ≤3958 counts/60s and vertical axis: ≤1756 counts/60s) demonstrated good classification accuracy (ROC-AUC = 0.85-0.86) and accurately estimated SB time in 5-8y (equivalence p=0.02; mean bias: 4.1%, limits of agreement [LoA]: -20.1-28.4%) and 9-12y (equivalence p<0.01; -2.5%, -27.9-22.9%). Mean bias of SB time estimates from Kim were smaller than ActiGraph hip (5-8y: 15.8%, -5.7-37.2%; 9-12y: 17.8%, -3.9-39.5%) and similar to or smaller than activPAL3™ (5-8y: 12.6%, -39.8-14.7%; 9-12y: -1.4%, -13.9-11.0%), although classification accuracy was similar to ActiGraph hip (ROC-AUC = 0.85) but lower than activPAL3™ (ROC-AUC = 0.92-0.97). Mean bias (5-8y: 6.5%, -16.1-29.1%; 9-12y: 10.5%, -13.6-34.6%) for the most accurate GENEActiv wrist cut-point (Schaefer: ≤0.19g) was smaller than ActiGraph hip, and activPAL3™ in 5-8y, but larger than activPAL3™ in 9-12y. However, SB time estimates from Schaefer were not equivalent to direct observation (equivalence p>0.05) and classification accuracy (ROC-AUC = 0.79-0.80) was lower than for ActiGraph hip and activPAL3™.

Conclusion: The most accurate SB ActiGraph (Kim) and GENEActiv (Schaefer) wrist cut-points can be applied in children with similar confidence as the ActiGraph hip cut-point (≤25 counts/15s), although activPAL3™ was generally more accurate.

Keywords: activity monitor, youth, validation, physical activity, objective measurement, sitting

Introduction

Sedentary behaviors (SB) are defined as any waking behaviors in a sitting or reclining position that require an energy expenditure of ≤1.5 metabolic equivalents (30). Although some studies among children and adolescents suggest that the total volume or pattern of SB is associated with adverse health outcomes, independent of moderate- to vigorous intensity physical activity (MVPA) (7, 8, 24), overall the evidence appears to be inconsistent (6, 11). Accurate measures of SB are essential for both observational and experimental research to further investigate the influence of SB on health outcomes, as well as the prevalence and determinants of SB, and the effectiveness of interventions to reduce SB.

Accelerometry is the method of choice for objectively measuring the amount and patterning of SB in children (32) and various accelerometers are available for placement on different body locations (e.g. hip, wrist or thigh) (17). Hip-mounted accelerometers have commonly been used in children (32), with cut-point approaches typically applied to define SB (17). For example, large population surveys, such as the National Health and Nutrition Examination Study (NHANES) 2003-2004 incorporated hip-worn ActiGraph accelerometers and SB time was estimated using a <100 counts/minute threshold (22). However, concerns about low participant compliance to accelerometry protocols and subsequent data loss have resulted in a shift from hip to wrist placement (14). NHANES 2011-2014 (31) incorporated wrist-worn accelerometers and the data from this study and other initial reports (13, 28) indicate that wrist-placement results in increased wear time due to greater compliance, which in turn leads to greater confidence that the data are representative of daily physical activity and SB. The ActiGraph (ActiGraph LLC, Pensacola Beach, FL) and GENEActiv (ActivInsights Ltd., Cambridge, UK) are accelerometer-based motion sensors typically worn on the hip or wrist. Thresholds or cut-points have been developed for the wrist-worn ActiGraph (5, 9, 19) and GENEActiv (26, 29) to classify SB in children. The wrist cut-points were developed using different age groups, sample sizes and activity protocols, which results in variations in the cut-points used to classify SB. For example, wrist cut-points developed for ActiGraph's vertical axis (VA; *x*-axis) range from 35 counts[c]/5s (9) to 202c/5s (Chandler et al., personal communication, 2016). Using different accelerometer models, placing them at different body locations, and applying different cut-points, results in considerable differences in estimates of SB (17, 28), which makes it difficult to compare outcomes between studies and examine the epidemiology of SB. Therefore, comparison of these assessment methods is needed. Rowlands et al. (2014) compared free-living SB estimates from a GENEActiv (26) signal vector magnitude (SVM) wrist cut-point ($\text{Phillips}_{\text{SVM}}$: right wrist, <6gs; left wrist, <7gs) with the widely used ActiGraph hip cut-point for VA (Evenson: ≤25c/15s) (12) in a sample of free-living 10-12 year-olds (28). This study reported that the outcomes from these monitors were highly correlated, however, sedentary time estimated by $\text{Phillips}_{\text{SVM}}$ was significantly lower (9.6%) than estimates from the ActiGraph hip cut-point. Because the study did not have a criterion measure of SB, the level of error from each measure is unknown. Furthermore, the relative validity of the range of

GENEActiv and ActiGraph wrist cut-points remains unknown, because only one accelerometer model and one cut-point for the wrist were evaluated.

It is also important to evaluate the validity of recent SB wrist cut-points against alternative objective measures to understand the accuracy of newer approaches relative to other options for assessing SB. One alternative method is thigh-mounted accelerometry, such as the activPAL3^{TM} (PAL Technology Ltd., Glasgow, UK) posture detection system, which classifies periods spent sitting/lying, standing or stepping. Because of the monitor's placement on the thigh, it uses the orientation (angle to vertical) of the thigh to accurately estimate SB (34), rather than simply the movement intensity measures used in traditional hip-based cut-point approaches which have difficulties differentiating between standing and sitting (17, 21). Whether or not wrist-based cut-point approaches provide equally accurate estimates of SB relative to alternative approaches such as hip- or thigh-based accelerometry is unclear and requires further investigation. Furthermore, it is important to evaluate the accuracy of the wrist cut-points to detect breaks in SB in order to understand their influence on health outcomes.

To our knowledge, no comprehensive validation studies have been conducted in children in which sedentary wrist cut-points for the ActiGraph or GENEActiv have been evaluated simultaneously during a standardised activity protocol, against a criterion measure and alternative objective measures of SB. Therefore, the aims of this study were to examine the classification accuracy and validity of sedentary wrist cut-points for ActiGraph and GENEActiv, relative to the hip-mounted ActiGraph (Evenson: ≤25c/15s) and the thigh-mounted activPAL3^{TM}, using direct observation as the criterion measure in 5-12 year-olds. Based on evidence that the thigh-mounted activPAL3^{TM} demonstrated acceptable accuracy for classifying SB in school-aged children (34) and that traditional hip-based accelerometers tend to overestimate time spent in SB (17), and the assumption that wrist cut-points might have similar difficulties as hip cut-points in discriminating between standing and sitting, it was hypothesized that the most accurate wrist cut-points would demonstrate similar accuracy as the hip cut-point for assessing SB, but lower accuracy than the thigh-mounted activPAL3™.

Methods

Participants

Fifty-seven children aged 5-12y who were without physical or health conditions that would affect participation in physical activity were recruited as part of an activity monitor validation study. The study was approved by the University of Wollongong Health and Medical Human Research Ethics Committee. Written parental consent and participant assent were obtained prior to participation.

Procedures

Participants were required to visit the laboratory on two occasions. Anthropometric measures were completed during the first visit using standardised procedures while children were wearing light clothing and with shoes removed. BMI (kg/m^2) and weight status were calculated (20). Children completed a protocol of 15 semi-structured activities from sedentary (lying down, TV viewing, handheld e-game, writing/coloring, computer game), light (getting ready for school, standing class activity, slow walk, dancing), and moderate-to-vigorous (tidy up, brisk walk, soccer, basketball, running, locomotor course) intensity (Supplemental Digital Content 1). Activities were equally divided over 2 visits and completed in a structured order of increasing intensity for 5 min, except for lying down (10 min).

At each visit, children were fitted with an ActiGraph GT3X+ on the right hip (midaxilla line at the level of the iliac crest) with an elastic belt, and an ActiGraph GT3X+ and a GENEActiv dorsally on each wrist. The distal and proximal position of the ActiGraph and GENEActiv monitors on each wrist was alternated for each participant to avoid placement effects. An activPAL3™ was placed mid-anteriorly on the right thigh.

Activity monitors

The ActiGraph GT3X+ is a tri-axial accelerometer that measures accelerations ranging in magnitude ±6g. Raw accelerometry data can be stored at a user-specified sample frequency ranging from 30-100Hz. The GENEActiv has a waterproof design and measures tri-axial accelerations ranging in magnitude ±8g at a sample frequency ranging from 10-100Hz. The ActiGraph and GENEActiv were initialised with a sample frequency of 100Hz. Data reduction approaches were performed according to the methods used to develop each cut-point (Table 1), as reported in original calibration studies (5, 9, 12, 19, 26, 29). Raw ActiGraph data were downloaded using ActiLife version 6.12.1. ActiGraph hip and wrist data were converted to counts per 5s (5, 9), 15s (12), or 60s (19) corresponding to the epoch lengths used in their development. Output variables for ActiGraph monitors were VA, which is sensitive to movement only along the longitudinal axis of the lower arm or the dominant plane of the body (hip) and vector magnitude (VM), a 3-dimensional measure of the acceleration which is not sensitive to orientation and direction of movement. Raw GENEActiv wrist data were downloaded and converted into 1s epochs using the GENEActiv software version 2.2 according to methods described by Philips et al. (26), in order to create gravity-subtracted signal vector magnitude (SVMgs) data. Customized software was used to filter the raw GENEActiv data (bandpass filter, cut-off frequencies: 0.2 and 15Hz) in order to remove the gravitational acceleration component as well as high-frequency sensor noise, as described by Schaefer et al. (29). An average gravity-subtracted signal vector magnitude (SVMg) was then calculated for each second using a formula described by the authors.

The activPAL3^{TM} is an activity monitor worn on the thigh that uses tri-axial acceleration data (20Hz) to assess the position and movement of the limb. The activPAL3^{TM} software version 7.2.32 with proprietary algorithms was used to classify tri-axial accelerometry data into periods spent sitting/lying, standing or stepping. Event records created by the software were used to create 1s epoch data files which were used in the analyses to classify periods spent sedentary. The activPAL3^{TM} was initialised with minimum sitting or upright period of 1s.

Direct observation

Direct observation was used as criterion measure to establish the classification accuracy and validity of the cut-points. Children were recorded on video completing the activities as well as during transitions between activities. A single observer coded all videos using Vitessa 0.1 (University of Leuven, Belgium) which generated a time stamp every time a change in posture or intensity was coded by the observer. Subsequently, a second-by-second classification system was generated. Every second following the time stamp inserted by the observer was classified as being the same posture as the one occurring at the time stamp itself until the next time stamp was created, indicating that a change in the child's posture had occurred. In the event of two postures occurring within the same second, this second was duplicated in order to label both postures. Labels for postures were sitting/lying (gluteus muscles resting on ground, feet, legs or any other surface, or lying in prone position), standing (e.g both feet touching the ground, squatting, standing on one foot, kneeling on one or two knees), stepping (e.g moving one leg in front of the other, including stepping with a flight phase, jumping, stepping, sliding/side gallop) and "off screen" for direct observation using 1s epochs. A dichotomous coding system was applied to re-code postures into sedentary (sitting/lying: "1") and non-sedentary (standing, stepping: "0"). Videos of 5 randomly selected participants were analysed twice by the same observer and by a second observer to test inter- and intra-observer reliability. Inter- and intra-observer reliability were examined using Cohen's Kappa and single measure intra-class correlation coefficients (ICC) from two-way mixed effect models (fixed-effects = observer; random effects = participants), using the consistency definition. Cohen's Kappa coefficient for inter-observer reliability was 0.941. Inter-observer ICC was 0.974 (0.974 - 0.974) and intra-observer ICC was 0.963 (0.962 - 0.963).

Data synchronization

Monitors and direct observation were time synchronized using an internal computer clock. Second-by-second direct observation data were synchronized with 1s epoch data from activPAL3™ and GENEActiv. Direct observation and activPAL3™ data files contained events of duplicated seconds when two postures were assigned to the same second. If this was the case for direct observation data, these seconds were duplicated at the corresponding time point for activPAL3™ and GENEActiv output. If this was the case for activPAL3™ data, the seconds were duplicated for direct observation and GENEActiv output. The second-by-second duplicates were not generated for ActiGraph output, because these data were exported in 5s, 15s and 60s epochs. This method was applied for evaluation of classification accuracy and was in line with previous validation studies in preschool children (10, 18). In order to align direct observation with ActiGraph epochs, new time frames were created for direct observation with steps of 5s, 15s and 60s. If >50% of the seconds within an epoch were classified as sedentary, the epoch was coded as sedentary ("1"), if ≤50% of the epoch was classified as sedentary, the epoch was coded as non-sedentary ("0"). The synchronized direct observation and accelerometry data were excluded when direct observation epochs were coded as “off screen”. For estimates of time spent in different postures, codes of duplicated seconds for either direct observation (0.02% of total direct observation data) or accelerometer (0.04% of total activPAL3™ data) were assigned 0.5sec, in order to avoid artificially inflating the total time observed. The absolute number of SB breaks for each method was defined as the number of transitions from SB to non-SB.

Statistical analyses

Prior to analyses, the total sample was divided into two age groups (5-8y, n=25 and 9-12y, n=32) because of the potential that younger and older children might engage in SB differently (17). Analyses included equivalence testing, Bland-Altman procedures and calculating sensitivity, specificity and area under the receiver operating curve (ROC-AUC) to evaluate and compare the accuracy and validity of different SB cut-points for wrist mounted ActiGraph and GENEActiv accelerometers, hip-worn ActiGraph accelerometer and activPAL3™. The equivalence of estimated sedentary time from different activity monitors, sites and cut-points and direct observation was examined at the group level of measurement using the 95% equivalence test. In order to reject the null-hypothesis of the equivalence test, the 90% confidence interval (CI) of time spent sedentary predicted by the monitors should fall entirely within the predefined equivalence region of ±10% (2). The 90% CIs of the estimated sedentary time were bootstrapped, because the sample sizes of the age groups were relatively small and, therefore, not all data were normally distributed. Agreement and systematic bias for estimated sedentary time were evaluated at the individual level using Bland-Altman procedures (17). For the ROC analyses, classification accuracy was rated as excellent (ROC-AUC ≥ 0.90), good (ROC-AUC = 0.80-0.89), fair (ROC-AUC = 0.70-0.79) or poor (ROC-AUC < 0.70) (23). The difference between the absolute number of SB breaks estimated by the monitors and direct observation was tested using paired sample *t*-tests.

Results

Descriptive characteristics of participants are presented in Table 2. All participants completed the protocol and had valid activPAL3^{TM} and ActiGraph wrist and hip data. For one of the visits, video data were unavailable for 3 children (age 5, 9 and 10y) and GENEActiv wrist data were unavailable for 3 different children (all 9-12y). Out of the remaining 250,854 1s epochs from 5-8y and 296,134 epochs from 9-12y, 27,983 epochs and 23,513 epochs of direct observation were coded as "off screen" and excluded from analyses, respectively, leaving 222,872 (88.8%) valid epochs for 5-8y and 272,622 (92.1%) valid epochs for 9-12y. Mean direct observation time for 5-8y was 167.2 ± 21.9 min, of which 78.0 ± 11.8 min was coded as SB. Mean direct observation time for 9-12y was 154.2 ± 35.6 min, of which 69.5 ± 18.4 min was coded as SB. Results are presented for the non-dominant wrist (unless stated otherwise), because placement on this wrist was recommended by the physical activity monitor protocol (4) released by the National Health and Nutrition Examination Survey and previous studies have used the non-dominant wrist for the development of wrist cut-points (5, 16, 29). Results for the dominant wrist are presented in Supplemental Digital Content.

Validation of ActiGraph wrist cut-points

Figures 1 (5-8y) and 2 (9-12y) present the 95% equivalence tests for accelerometry-based estimated time spent in SB from wrist-worn ActiGraph and GENEActiv cut-points, the hip-worn ActiGraph cut-point and activPAL3^{TM}, as well as the equivalence region of direct observation. At the group level, estimates of SB time from Kim et al.'s ActiGraph VM wrist cut-point (Kim_{VM}) were equivalent to direct observation (p=0.02) in 5-8y, and estimates from the VA cut-point (Kim_{VA}) approached equivalence (p=0.08). Mean bias for estimated SB time from Kim_{VM} was 4.1% (limits of agreement [LoA]: -20.1% – 28.4%) (Table 3), whereas

Kim_{VA} underestimated SB time by 6.5% (LoA: -33.1% – 20.2%). In 9-12y, $Crouter_{VA/ROC}$ and Kim_{VA} were equivalent to direct observation ($p<0.01$) and $Crouter_{VM/ROC}$ approached equivalence ($p=0.05$). These cut-points underestimated SB time by 1.7% (LoA: -25.9% – 22.5%), 2.5% (LoA: -27.9% – 22.9%) and 5.3% (LoA: -27.9% – 22.9%), respectively. Estimates of SB time from other ActiGraph wrist cut-points were not equivalent to direct observation in either age group. The mean bias varied from 7.2% ($Crouter_{VA/ROC}$) to 20.5% ($Chandler_{VA/2016}$) in 5-8y and from 10.9% ($Crouter_{VA/REG}$) to 29.6% ($Chandler_{VA/2016}$) in 9-12y. Good classification accuracy (Table 4) was found for Kim_{VA} (both age groups: ROC-AUC = 0.86) and Kim_{VM} (5-8y: ROC-AUC = 0.85; 9-12y: ROC-AUC = 0.82). Classification accuracy for other ActiGraph wrist cut-points was fair (5-8y: ROC-AUC = 0.77-0.79, 9-12y: ROC-AUC = 0.72-0.75). At the individual level (Table 3), LoAs for all cut-points, including the most accurate ActiGraph wrist cut-points, were relatively wide (range = $Chandler_{VA/2016}$ in 5-8y: 0.0% – 41.0%; to $Chandler_{VA/2016}$ in 9-12y: -6.6% – 65.9%), which indicated large random error. No systematic bias (Table 3) was found for any of the ActiGraph wrist cut-points ($p>0.05$). Findings of the equivalence test, classification accuracy and Bland-Altman analyses for ActiGraph wrist cut-points for the dominant wrist (Supplemental Digital Content 2, 3 and 4) were consistent with findings for the non-dominant wrist. Compared to direct observation, the absolute number of breaks were overestimated by all ActiGraph cut-points in both age groups for both wrists (5-8y: mean difference range = 2.4-160.8, all $p<0.05$; 9-12y: mean difference range = 1.8-138.6, all $p<0.05$), except from Kim_{VM} for the non-dominant wrist (5-8y: mean difference = 1.4±5.7, $p=0.24$; 9-12y: mean difference = 1.8, $p=0.05$) (Supplemental Digital Content 5). Mean differences with direct observation were larger for wrist cut-points developed with 5sec epochs (5-8y: 154.4±4.1, 9-12y: 129.9±5.2) compared to cut-points developed with 60sec epochs (5-8y: 2.9±1.2, 9-12y: 2.5±0.8).

Validation of GENEActiv wrist cut-points

Estimates of SB time from GENEActiv wrist cut-points $Phillips_{SVM}$ and $Schaefer_{SVM}$ for the non-dominant wrist were not equivalent to direct observation (Figures 1 and 2). $Phillips_{SVM}$ and $Schaefer_{SVM}$ overestimated SB time in 5-8y by 16.8% (LoA: -3.9% – 29.6%) and 9.6% (LoA: -13.8% – 33.0%), respectively, and in 9-12y by 17.8% (LoA: -11.6% – 47.3%) and 12.6% (LoA: -12.3% – 37.6%), respectively (Table 3). Although estimates from the GENEActiv wrist cut-points for the dominant wrist were also not equivalent to direct observation in both age groups, the cut-points performed slightly better for this wrist when estimating SB time at the group level (Supplemental Digital Content 4). For the dominant wrist, $Phillips_{SVM}$ and $Schaefer_{SVM}$ overestimated SB time in 5-8y by 8.1% (LoA: -24.0% – 40.1%) and 6.5% (LoA: -16.1% – 29.1%), respectively, and in 9-12y by 8.2% (LoA: -18.6% – 35.0%) and 10.5% (LoA: -13.6% – 34.6%), respectively (Supplemental Digital Content 2). Classification accuracy for all GENEActiv wrist cut-points were fair to good in both age groups and for both wrists (ROC-AUC = 0.79-0.80). At the individual level, the LoA was smallest for $Phillips_{SVM}$ (-3.9% – 29.6%), although all other LoAs for GENEActiv cut-points were relatively wide, which indicated large random error (Table 3 and Supplemental Digital Content 2). No systematic bias was found for any of the GENEActiv wrist cut-points ($p>0.05$). All GENEActiv wrist cut-points overestimated the absolute number of breaks compared to direct observation in both age groups (5-8y: mean difference range = 354.8-468.8, all $p<0.01$; 9-12y: mean difference range = 313.2-398.1, all $p<0.01$) (Supplemental Digital Content 5). Mean differences with direct observation were larger for the GENEActiv wrist cut-points developed with 1sec epochs, compared to the ActiGraph cut-points developed with both 5sec epochs and 60sec epochs.

Comparison of validity of wrist cut-points against ActiGraph hip cut-point and activPAL3™

In 5-8y, estimates of SB time by activPAL3™ (12.6% [LoA: -39.8% – 14.7%]) and the hip-worn ActiGraph (15.8% [LoA: -5.7% – 37.2%]) were not equivalent to direct observation, and the most accurate ActiGraph wrist cut-points (Kim_{VA} and Kim_{VM}), GENEActiv wrist cut-points for the dominant wrist and $Schaefer_{SVM}$ for the non-dominant wrist had smaller mean biases. Despite these differences, LoAs for the ActiGraph and GENEActiv wrist cut-points were similarly wide to activPAL3™ and the hip-worn ActiGraph. In contrast to the group level findings, classification accuracy for the Kim cut-points were significantly lower than activPAL3™ (ROC-AUC = 0.92, 95%CI = 0.92-0.93), but similar to the hip-worn ActiGraph (ROC-AUC = 0.85, 95%CI = 0.84-0.85) in 5-8y. Classification accuracy of both GENEActiv wrist cut-points for the non-dominant and dominant wrist was significantly lower than activPAL3™ and the hip-worn ActiGraph.

In 9-12y, estimates of SB time by activPAL3™ were equivalent to DO (-1.4% [LoA: -13.95 - 11.0%]) (p<0.01), which was also the case for the most accurate ActiGraph wrist cut-points ($Crouter_{VA/ROC}$ and Kim_{VA}). However, mean biases were larger and estimates of SB time were not equivalent to direct observation for the hip-worn ActiGraph (17.8% [LoA: -3.9% - 39.5%]), and GENEActiv cut-points for either wrist in 9-12y. LoAs for the ActiGraph and GENEActiv wrist cut-points were wider than activPAL3™, but similar to ActiGraph on the hip in 9-12y. The most accurate ActiGraph wrist cut-point (Kim_{VA}) exhibited lower classification accuracy than activPAL3™ (ROC-AUC = 0.97, 95%CI = 0.97-0.97), but was similar to the hip-worn ActiGraph (ROC-AUC = 0.85, 95%CI = 0.84-0.85) in 9-12y. Classification accuracy of the GENEActiv cut-points for both wrists was lower than activPAL3™ and the hip-worn ActiGraph, in 9-12y.

Mean differences with direct observation for SB breaks were larger for most ActiGraph and both GENEActiv wrist cut-points compared to the activPAL3™ (5-8y: 8.5±6.0, $p<0.01$; 9-12: 3.2±3.1, $p<0.01$) and the hip-worn ActiGraph (5-8y: 33.2±13.7, $p<0.01$; 9-12: 29.3±10.9, $p<0.01$) in both age groups, except for the Kim_{VM} cut-points where the differences were smaller.

Discussion

This study examined the accuracy and validity of ActiGraph and GENEActiv wrist cut-points for classifying SB in 5-12 year-old children. The ActiGraph wrist cut-points Kim_{VM} and Kim_{VA} accurately estimated SB time in 5-8y and 9-12y, respectively, at the group level, and exhibited good classification accuracy. These cut-points provided more accurate estimates of SB time compared to the Evenson ActiGraph hip cut-point (≤25c/15s). Although GENEActiv wrist cut-points appeared to provide more accurate group-level estimates of SB time than the ActiGraph hip cut-point for 5-8y and 9-12y, these cut-points over-estimated SB time, and classification accuracy was significantly lower than for the ActiGraph hip cut-point and activPAL3™ in both age groups. Excluding an overestimation of SB time in 5-8y, activPAL3™ exhibited greater accuracy than the ActiGraph and GENEActiv wrist cut-points and the ActiGraph hip cut-point. Overall, the most accurate ActiGraph and GENEActiv wrist cut-points estimated SB with similar accuracy as the ActiGraph hip cut-point, although the accuracy of the thigh-mounted activPAL3™ was generally higher. The KIM_{VM} cut-point estimated the absolute number of breaks in SB more accurately than the ActiGraph hip cut-point and activPAL3™ in both age groups, whereas the other ActiGraph and GENEActiv wrist cut-points showed larger overestimations. To our knowledge, no previous studies have simultaneously evaluated the relative validity of multiple ActiGraph or GENEActiv wrist cut-points developed in different studies among children. Crouter et al. (9) cross-validated their

ActiGraph wrist cut-points using indirect calorimetry in an independent sample of 11-14 year-olds who completed 2h of unstructured physical activity. The authors reported that the errors for estimated SB time were small (-8.6% – 2.5%) and not significantly different from the criterion measure. However, traditional analyses that fail to reject the null hypothesis of similarity do not necessarily demonstrate that the cut-points meet an acceptable level of accuracy (2). Therefore, testing the equivalence could be beneficial when examining the clinical significance of potential errors. In our study, mean bias for estimated SB time from Crouter et al.'s cut-points were slightly larger, ranging from -7.2% to 11.5% in 5-8y and -1.7% to 16.8% in 9-12y. Equivalence testing indicated that only $\text{Crouter}_{\text{VA/ROC}}$ in 9-12y was equivalent to direct observation, although the classification accuracy for Crouter et al.'s cut-points across both age groups was only fair (ROC-AUC = 0.73 – 0.79). This suggests that, although errors may appear small, they might still be meaningful and misclassification of SB and non-SB may cancel each other out. Other methodological differences between our study and that of Crouter et al. (9), such as the younger age range of participants in our study could have contributed to the differences in findings, because younger and older children potentially engage in and move between sedentary and non-sedentary behaviors differently (17). Furthermore, the use of different criterion measures might have also contributed to the differences in measurement errors. (17)

Kim et al. (19) used a protocol of 12 randomly selected semi-structured activities to develop ActiGraph wrist cut-points (Kim_{VA} and Kim_{VM}) in a sub-sample of 7-13 year-olds (n = 49), and also provided results for the Evenson ActiGraph hip cut-point (≤25c/15s, n = 125) against which wrist cut-points could be compared. Although ROC-AUC values were not reported for the hip-worn ActiGraph, sensitivity (Se: true positive rate) for the wrist cut-points (Se: 93.0 – 94.3%) was similar to the hip cut-point (Se = 93.7%), whereas specificity (Sp: true negative rate) for the wrist cut-points (Sp: 79.9 – 83.5%) was lower than the hip cut-point (Sp = 92.5%) for classifying SB, suggesting that the hip-worn ActiGraph was slightly more accurate for classifying non-SB activities. However, the current study found that the classification accuracy for Kim et al.'s ActiGraph wrist cut-points and the ActiGraph hip cut-point was similar in both age groups. Cut-point approaches for hip-mounted monitors cannot reliably distinguish between standing still and SB, because SB is classified based on lack of movement, resulting in non-SB activities with minimal lower body movement being misclassified as SB. Because our study included transitions between activities, which likely involved standing with minimal movement, as well as a standing "classroom activity", the likelihood of misclassifying non-SB as SB by the hip-worn ActiGraph was higher than in Kim et al.'s (19) protocol. In contrast, Kim et al. (19) indicated that most instances of misclassification of non-SB by the hip monitor occurred during a hand weight exercise involving minimal trunk and lower body movement. As such, our findings suggest that wrist cut-points may have similar limitations to hip cut-points in misclassifying standing still as SB.

In relation to wrist GENEActiv SB cut-points, Rowlands et al. (28) compared $\text{Phillips}_{\text{SVM}}$ for the non-dominant wrist with the ActiGraph hip cut-point (Evenson: ≤25c/15s) in a sample of free-living 10-12 year-olds and reported that estimates of habitual SB time were 9.6% lower for the GENEActiv wrist cut-point compared to the ActiGraph hip cut-point, however, we found that the estimates of these cut-points were similar. The difference in study designs may have contributed to these contrasting findings. However, our results showed larger misclassification of SB by $\text{Phillips}_{\text{SVM}}$ compared to the hip-worn ActiGraph, and therefore precision for classifying SB and estimates at the individual level might be lower than group-level estimates.

Although some cut-points in the current study appear to provide reasonably accurate estimates of SB time, the ROC-AUC values indicate that classification accuracy was only categorised as fair or good. For example, group level estimates of SB time from Kim_{VM} and Kim_{VA} were equivalent or almost equivalent to direct observation and mean biases were smaller than that observed for the hip-worn ActiGraph and activPAL3^{TM}, however ROC-AUC values were lower than activPAL3^{TM} and similar to the ActiGraph hip cut-point. In 9-12y, the cut-points $Crouter_{VA/ROC}$ and Kim_{VA} were equivalent to DO and estimates of SB time were more accurate than the hip-worn ActiGraph and similar to activPAL3^{TM}. However, although classification accuracy for Kim_{VA} was good, classification accuracy for $Crouter_{VA/ROC}$ was only fair and lower than both activPAL3^{TM} and the hip-worn ActiGraph. A possible explanation is that SB as estimated by wrist cut-points was misclassified as non-SB in some activities. For instance, the highest percentage of misclassified SB epochs (AG: 0.4%-7.3%, GA: 1.4%-5.7%) was found during the coloring activity in 5-8y, which requires the child to use the hand, and so wrist monitors might record counts high enough to be misclassified as non-SB. In contrast, standing still while writing on a white board resulted in the highest percentage of misclassified epochs during non-SB activities for the non-dominant hand (5-8y: AG, 6.7%-9.7%, GA: 8.1%-8.6%; 9-12y: AG, 6.1%-9.0%, GA: 7.7%-8.3%), because the wrist monitors recorded low activity counts on this hand and misclassified epochs during the task as SB. Misclassification of SB and non-SB for wrist cut-points may cancel each other out, resulting in seemingly accurate group-level estimates of SB time. Hip-placed monitors on the other hand seem to overestimate SB time at the group level, due to the misclassification of standing still as SB. The results of this study suggest that, while hip-based cut-points that typically misclassify standing still as SB, wrist cut-points exhibit some misclassification of non-SB as SB and vice-versa. Progress on alternative approaches, such as those utilising machine learning (15, 27, 33) is therefore required, but until such strategies are widely available, the use of the most accurate ActiGraph and GENEActiv wrist cut-points for estimating SB is recommended.

ActiGraph wrist cut-points developed with 60s epochs seemed to perform better for estimating SB time at the group level and the absolute number of SB breaks, and exhibited higher classification accuracy and compared to cut-points developed with 5s or 1s epochs. This could be explained by a higher number of data points when using shorter epochs, resulting in a higher chance of misclassification. The lower classification accuracy with shorter epochs might have contributed to the lower performance of the GENEActiv wrist cut-points as they were developed with 1 s data. This is in contrast to the common use of short epochs for accurately capturing sporadic and intermittent bursts of high–intensity physical activity in children (3). Previous studies have evaluated the effect of epoch length in free-living school-aged children using ActiGraph hip data and showed that time spent in SB decreases when longer epochs are applied (1, 25). A possible explanation is that very short periods (e.g. 1-5s) of standing relatively still might be fairly common in children, resulting in non-SB being misclassified as SB using short epochs. In contrast, when using 60s epochs, standing still would need to occur for almost all of a 60s period for this to be misclassified as SB, and it is possible that this is less common than short periods of standing still among children. Although most ActiGraph wrist cut-points designed for 5s epochs over-estimated SB in our analyses, $Crouter_{VA/ROC}$ and $Crouter_{VM/ROC}$ under-estimated SB in 5-8y and exhibited similar accuracy as those for 60s epochs in 9-12y, and so the combination of epoch and cut-point is likely to be important. Nevertheless, our findings indicate that the most accurate SB wrist cut-points were designed for 60s epochs, which has implications for field-based applications. In studies of free-living children, estimates of both SB and physical activity are often desirable. If data are reduced using short epochs such as 5s to estimate physical activity, the most accurate SB cut-points for 5s epochs could be applied, such as Crouter et al.'s $Crouter_{VA/ROC}$ or $Crouter_{VM/ROC}$ (9) for ActiGraph and $Phillips_{SVM}$ (26) or Schaefer et al.'s (29) for GENEActiv. Although these cut-points exhibited lower classification accuracy than the most accurate 60s wrist cut-points and the ActiGraph hip cut-point, group-level estimates of SB time were more accurate than the ActiGraph hip cut-point.

A unique strength of the study was that several currently available wrist cut-points for ActiGraph and GENEActiv were evaluated simultaneously, against a criterion measure and common alternative objective measures of SB. Another strength was that data from the entire activity protocol in our study were analysed including transitions between activities, with the aim to also include data of behaviors outside of structured activities. Additionally, the wide age range of the sample allowed for analyses across two age groups. However, because the study protocol predominantly included structured activities completed in a laboratory setting, the findings should be confirmed under free-living conditions.

In summary, the use of the most accurate ActiGraph and GENEActiv wrist-based activity monitor cut-points for estimating SB can be applied in free-living children with similar confidence as the hip-based ActiGraph cut-point (≤25c/15s), although alternative approaches may be needed to achieve the generally higher accuracy of thigh-based approaches such as activPAL3^{TM}.

Acknowledgements

We would like to thank all children and their parents for their participation. We also thank Melinda Smith for her assistance with recruitment and data collection and Woranart Maneenin for video analyses. This study was funded by the National Heart Foundation of Australia (G11S5975). DPC is supported by an Australian Research Council Discovery Early Career Researcher Award (DE140101588). ADO is supported by a National Heart Foundation of Australia Career Development Fellowship (CR11S 6099). TH is funded by a National Health and Medical Research Council Early Career Fellowship (APP1070571). The work of UE and SB is funded by the UK Medical Research Council (MC_UU_12015/3). ST is supported by the National Health and Medical Research Council Centre of Research Excellence on Sitting Time and Chronic Disease Prevention (APP1057608).

Conflict of Interest

The authors have no conflict of interest to declare. The results of the present study do not constitute endorsement by the American College of Sports Medicine. The results are presented clearly, honestly, and without fabrication, falsification, or inappropriate data manipulation.

References

1. Aibar A, Chanal J. Physical Education: The Effect of Epoch Lengths on Children's Physical Activity in a Structured Context. *PloS one*. 2015;10(4). doi: 10.1371/journal.pone.0121238
2. Batterham M, Van Loo C, Charlton K, Cliff D, Okely A. Improved interpretation of studies comparing methods of dietary assessment: Combining equivalence testing with the limits of agreement. *Br J Nutr* 2016; 115(7):1273-1280.
3. Cain KL, Sallis JF, Conway TL, Van Dyck D, Calhoon L. Using accelerometers in youth physical activity studies: a review of methods. *J Phys Act Health*. 2013;10(3):437-50.
4. Center for Disease Control and Prevention [internet]. Physical activity Monitor (PAM) Procedures Manual; [cited 2016 May 19]. Available from: http://www.cdc.gov/nchs/data/nhanes/nhanes_11_12/Physical_Activity_Monitor_Manual.pdf
5. Chandler J, Brazendale K, Beets M, Mealing B. Classification of physical activity intensities using a wrist-worn accelerometer in 8–12-year-old children. *Pediatr Obes*. 2015; 11(2):120-127.
6. Cliff D, Hesketh K, Vella S et al. Objectively measured sedentary behaviour and health and development in children and adolescents: systematic review and meta-analysis. *Obes Rev*. 2015;17(4):330-344.
7. Cliff DP, Jones RA, Burrows TL et al. Volumes and bouts of sedentary behavior and physical activity: associations with cardiometabolic health in obese children. *Obesity*. 2014;22(5):112-8.

8. Cliff DP, Okely AD, Burrows TL et al. Objectively measured sedentary behavior, physical activity, and plasma lipids in overweight and obese children. *Obesity*. 2013;21(2):382-5.

9. Crouter SE, Flynn JI, Bassett Jr DR. Estimating physical activity in youth using a wrist accelerometer. *Med Sci Sports Exerc*. 2015;47(5):944-51.

10. 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.

11. Ekelund U, Luan Ja, Sherar LB et al. Moderate to vigorous physical activity and sedentary time and cardiometabolic risk factors in children and adolescents. *Jama*. 2012;307(7):704-12.

12. 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.

13. Fairclough SJ, Noonan R, Rowlands AV, van Hees V, Knowles Z, Boddy LM. Wear Compliance and Activity in Children Wearing Wrist and Hip-Mounted Accelerometers. *Med Sci Sports Exerc*. 2016;48(2):245-253.

14. Freedson PS, John D. Comment on estimating activity and sedentary behavior from an accelerometer on the hip and wrist. *Med Sci Sports Exerc*. 2013;45(5):962-3.

15. Hagenbuchner M, Cliff DP, Trost SG, Van Tuc N, Peoples GE. Prediction of activity type in preschool children using machine learning techniques. *J Sci Med Sport*. 2015;18(4):426-31.

16. Hildebrand M, Van Hees VT, Hansen BH, Ekelund U. Age-group comparability of raw accelerometer output from wrist-and hip-worn monitors. *Med Sci Sports Exerc*. 2014;46(9):1816-24.

17. Janssen X, Cliff DP. Issues related to measuring and interpreting objectively measured sedentary behavior data. *Meas Phys Educ Exerc Sci*. 2015;19(3):116-24.

18. Janssen X, Cliff DP, Reilly JJ et al. Validation of activPALTM defined sedentary time and breaks in sedentary time in 4-to 6-year olds. *Pediatr Exerc Sci*. 2014;26(1):110-7.

19. Kim Y, Lee J-M, Peters BP, Gaesser GA, Welk GJ. Examination of different accelerometer cut-points for assessing sedentary behaviors in children. *PloS one*. 2014;9(4). doi: 10.1371/journal.pone.0090630

20. Kuczmarski RJ, Ogden CL, Guo SS et al. 2000 CDC Growth Charts for the United States: methods and development. *Vital Health Stat 11* 2002; 246:1-190.

21. Lubans DR, Hesketh K, Cliff D et al. A systematic review of the validity and reliability of sedentary behaviour measures used with children and adolescents. *Obes Rev*. 2011;12(10):781-99.

22. Matthews CE, Chen KY, Freedson PS et al. Amount of time spent in sedentary behaviors in the United States, 2003–2004. *Am J Epidemiol*. 2008;167(7):875-81.

23. Metz CE. Basic principles of ROC analysis. *Semin Nuel Med.* 1978;8(4):283-98.

24. Mitchell J, Pate R, Beets M, Nader P. Time spent in sedentary behavior and changes in childhood BMI: a longitudinal study from ages 9 to 15 years. *International Journal of Obesity*. 2012;37(1):54-60.

25. Nettlefold L, Naylor P, Warburton DE, Bredin SS, Race D, McKay HA. The Influence of Epoch Length on Physical Activity Patterns Varies by Child's Activity Level. *Res Q Exerc Sport*. 2016;87(1):110-23.

26. Phillips LR, Parfitt G, Rowlands AV. Calibration of the GENEA accelerometer for assessment of physical activity intensity in children. *J Sci Med Sport*. 2013;16(2):124-8.

27. Rowlands AV, Olds TS, Hillsdon M et al. Assessing Sedentary Behavior with the GENEActiv: Introducing the Sedentary Sphere. *Med Sci Sports Exerc*. 2014;46(6):1235-1247.

28. Rowlands AV, Rennie K, Kozarski R et al. Children's physical activity assessed with wrist-and hip-worn accelerometers. *Med Sci Sports Exerc*. 2014; 46(6):2308-16.

29. Schaefer CA, Nigg CR, Hill JO, Brink LA, Browning RC. Establishing and evaluating wrist cutpoints for the GENEActiv accelerometer in youth. *Med Sci Sports Exerc*. 2014;46(4):826.

30. Sedentary Behavior Research Network. Letter to the editor: standardized use of the terms "sedentary" and "sedentary behaviours". *Appl Physiol Nutr Metab*. 2012;37:540-2.

31. 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.

32. Trost SG. State of the Art Reviews: Measurement of Physical Activity in Children and Adolescents. *Am J Lifestyle Med*. 2007;1(4):299-314.

33. Trost SG, Wong W-K, Pfeiffer KA, Zheng Y. Artificial neural networks to predict activity type and energy expenditure in youth. *Med Sci Sports Exerc*. 2012;44(9):1801.

34. Van Loo CMT, Okely AD, Batterham MJ et al. Validation of thigh-based accelerometer estimates of postural allocation in 5-12 year-olds. *J Sci Med Sport* [internet]. 2016 [cited 2016 Aug 18]. Available from: http://www.sciencedirect.com/science/article/pii/S1440244016301517. doi: 10.1016/j.jsams.2016.08.008.

Supplemental Digital Content

Supplemental Digital Content 1.docx

Supplemental Digital Content 2.docx

Supplemental Digital Content 3.docx

Supplemental Digital Content 4.tif

Supplemental Digital Content 5.docx

Figure 1. 95% equivalence test for accelerometry-based estimated time spent in sedentary behaviors in 5-8 year-olds.

Legend Figure 1: Times estimated by wrist-worn ActiGraph and wrist-worn GENEActiv cut-points are equivalent to direct observation if 90% confidence intervals lie entirely within the equivalence region of direct observation.VA: vertical axis; VM: vector magnitude; SVM: gravity-subtracted signal vector magnitude; ROC: developed using receiver operating curve analysis; Regression: developed using regression analysis.

Figure 2. 95% equivalence test for accelerometry-based estimated time spent in sedentary behaviors in 9-12 year-olds.

Legend Figure 2: Times estimated by wrist-worn ActiGraph and wrist-worn GENEActiv cut-points are equivalent to direct observation if 90% confidence intervals lie entirely within the equivalence region of direct observation. VA: vertical axis; VM: vector magnitude; SVM:

gravity-subtracted signal vector magnitude; ROC: developed using receiver operating curve analysis; Regression: developed using regression analysis.

Supplemental Digital Content 4. 95% equivalence test for accelerometry-based estimated time spent in sedentary behaviors for the dominant wrist in a) 5-8 year-olds and b) 9-12 year-olds.

Legend Supplemental Digital Content 4: Times estimated by wrist-worn ActiGraph and wrist-worn GENEActiv cut-points are equivalent to direct observation if 90% confidence intervals lie entirely within the equivalence region of direct observation. VA: vertical axis; VM: vector magnitude; SVM: gravity-subtracted signal vector magnitude; ROC: developed using receiver operating curve analysis; Regression: developed using regression analysis.

Table 1 Sedentary wrist cut-points

Monitor	Author	Outcome variable	Abbreviation	Sample	Activities	Cut-point
ActiGraph	Chandler et al. (4)	Vertical axis	Chandler$_{VA/2015}$	n = 45 Range = 8-12y Mean age = 9.0y 49% boys, 51% girls	Resting, enrichment, walking, playground, splash pad, swimming, endurance run	<161c/5s
		Vector Magnitude	Chandler$_{VM}$			<305c/5s
	Chandler et al. (personal communication)	Vertical axis	Chandler$_{VA/2016}$	n = 167 (calibration: n = 100) Range = 5-11y Mean age = 8.0y 58% boys, 42% girls	Reading books, playing/sorting cards, cutting and pasting from magazines, playing board games, eating a snack, playing games on a tablet, watching TV, and writing with a pencil, walking	<202c/5s

Crouter et al. (8)	Vertical axis	$Crouter_{VA/ROC}$	n = 181 Range = 8-15y Mean age = 12.0y 53.6% boys, 46.4% girls	One out of four structured activity routines including free-living activities such as: resting, reading, watching TV, walking, running, computer games, cleaning, playing wall ball, soccer	≤35c/5s
		$Crouter_{VA/REG}$			≤105c/5s
	Vector Magnitude	$Crouter_{VM/ROC}$			≤100c/5s
		$Crouter_{VM/REG}$			≤275c/5s
Kim et al. (21)	Vertical axis	Kim_{VA}	n = 49 Range = 7-13y Mean age = 10.1y 40.8% boys, 59.2% girls	Set of 12 activities such as: reading, watching TV, walking, running, playing catch, basketball, stationary cycling	≤1756c/60s
	Vector Magnitude	Kim_{VM}			≤3958c/60s

GENEActiv	Phillips et al. (30)	SVMgs	$Phillips_{SVM}$	n = 44 Range = 8-14y Mean age = 10.9y 40.9% boys, 59.1% girls	Lying supine, seated DVD viewing, active computer games (boxing), using a Nintendo Wii, slow walking, brisk walking, slow running and a medium run	Right: <6gs, left: <7gs
	Schaefer et al. (35)	SVMg	$Schaefer_{SVM}$	n = 24 children Range = 6-11y Mean age = 9.2y 54.2% boys, 45.8% girls	Resting, colouring, Lego® building, Wii Sports® games, treadmill walking, jogging, running	≤0.19g

Notes Table 1: VA: vertical axis; VM: vector magnitude; c: counts; s: seconds; SVMg/gs: gravity-subtracted signal vector magnitude; g: gravity; gs: g · seconds; ROC: developed using receiver operating curve analysis; Regression: developed using regression analysis

Table 2. Participant characteristics

	5-8y (n=25)	**9-12y (n=32)**	**Total (n=57)**
Age (y)	7.0 ± 1.2	10.9 ± 1.2	9.2 ± 2.3
Sex			
Boys (*n*)	11 (44.0%)	17 (53.1%)	28 (49.1%)
Girls (*n*)	14 (56.0%)	15 (46.9%)	29 (50.9%)
Height (cm)	123.0 ± 8.9	146.0 ± 9.2	135.9 ± 14.6
Body mass (kg)	24.1 ± 4.0	39.4 ± 9.9	32.7 ± 10.9
BMI percentile	52.8 ± 24.3	53.5 ± 31.9	53.2 ± 28.6
Overweight (*n*)	2 (8.0%)	5 (15.6%)	7 (12.3%)
Obese (*n*)	-	2 (6.6%)	2 (3.5%)
Race			
Caucasian (*n*)	24 (96.0%)	30 (93.8%)	54 (94.7%)
Asian (*n*)	1 (4.0%)	2 (6.2%)	3 (5.3%)

Characteristics of the participants are presented as mean ± SD, distributions of the sample are presented in numbers (*n*) and percentages.

Table 3 Agreement analysis of accelerometry-based estimations of sedentary behavior compared to direct observation.

	Cut-point	Mean bias (%)	95% LoA	Slope p-value
ActiGraph wrist (vertical axis)	$Crouter_{VA/ROC}$			
	5-8y	7.2	-19.4 - 33.9	0.367
	9-12y	1.7*	-22.5 - 25.9	0.677
	$Crouter_{VA/REG}$			
	5-8y	-7.6	-30.4 - 15.2	0.673
	9-12y	-10.9	-33.1 - 11.3	0.770
	$Chandler_{VA/2015}$			
	5-8y	-15.4	-36.5 - 5.6	0.975
	9-12y	-19.0	-42.1 - 4.1	0.726
	$Chandler_{VA/2016}$			
	5-8y	-20.5	-41.0 - 0.0	0.966
	9-12y	-29.6	-65.9 - 6.6	0.306
	Kim_{VA}			
	5-8y	6.5	-20.2 - 33.1	0.718
	9-12y	2.5*	-22.9 - 27.9	0.892
ActiGraph wrist (vector magnitude)	$Crouter_{VM/ROC}$			
	5-8y	11.5	-16.8 - 39.8	0.323
	9-12y	5.3	-22.5 - 33.2	0.752
	$Crouter_{VM/REG}$			
	5-8y	-11.0	-35.2 - 13.1	0.436
	9-12y	-16.8	-44.6 - 10.9	0.563

	$Chandler_{VM}$			
	5-8y	-14.4	-38.5 - 9.7	0.401
	9-12y	-20.8	-49.8 - 8.1	0.542
	Kim_{VM}			
	5-8y	-4.1*	-28.4 - 20.1	0.522
	9-12y	-13.3	-43.7 - 17.1	0.454
GENEActiv wrist (signal vector magnitude)	$Phillips_{SVM}$			
	5-8y	-16.8	-29.6 - 3.9	0.744
	9-12y	-17.8	-47.3 - 11.6	0.737
	$Schaefer_{SVM}$			
	5-8y	-9.6	-33.0 - 13.8	0.957
	9-12y	-12.6	-37.6 - 12.3	0.898
activPAL3^{TM}	5-8y	12.6	-14.7 - 39.8	0.122
	9-12y	1.4*	-11.0 - 13.9	0.442
ActiGraph hip (vertical axis)	5-8y	-15.8	-37.2 - 5.7	0.204
	9-12y	-17.8	-39.5 - 3.9	0.260

Notes Table 3: LoA: limits of agreement; VA: vertical axis; VM: vector magnitude; SVM: gravity-subtracted signal vector magnitude; c: counts; s: seconds; g: gravity; gs: g · seconds. Mean bias was calculated as: measured SB time – estimated SB time; a positive value indicates underestimation; a negative value indicates overestimation. *Significantly equivalent to direct observation ($p < 0.05$).

Table 4 Classification accuracy of accelerometry-based estimations of sedentary behavior.

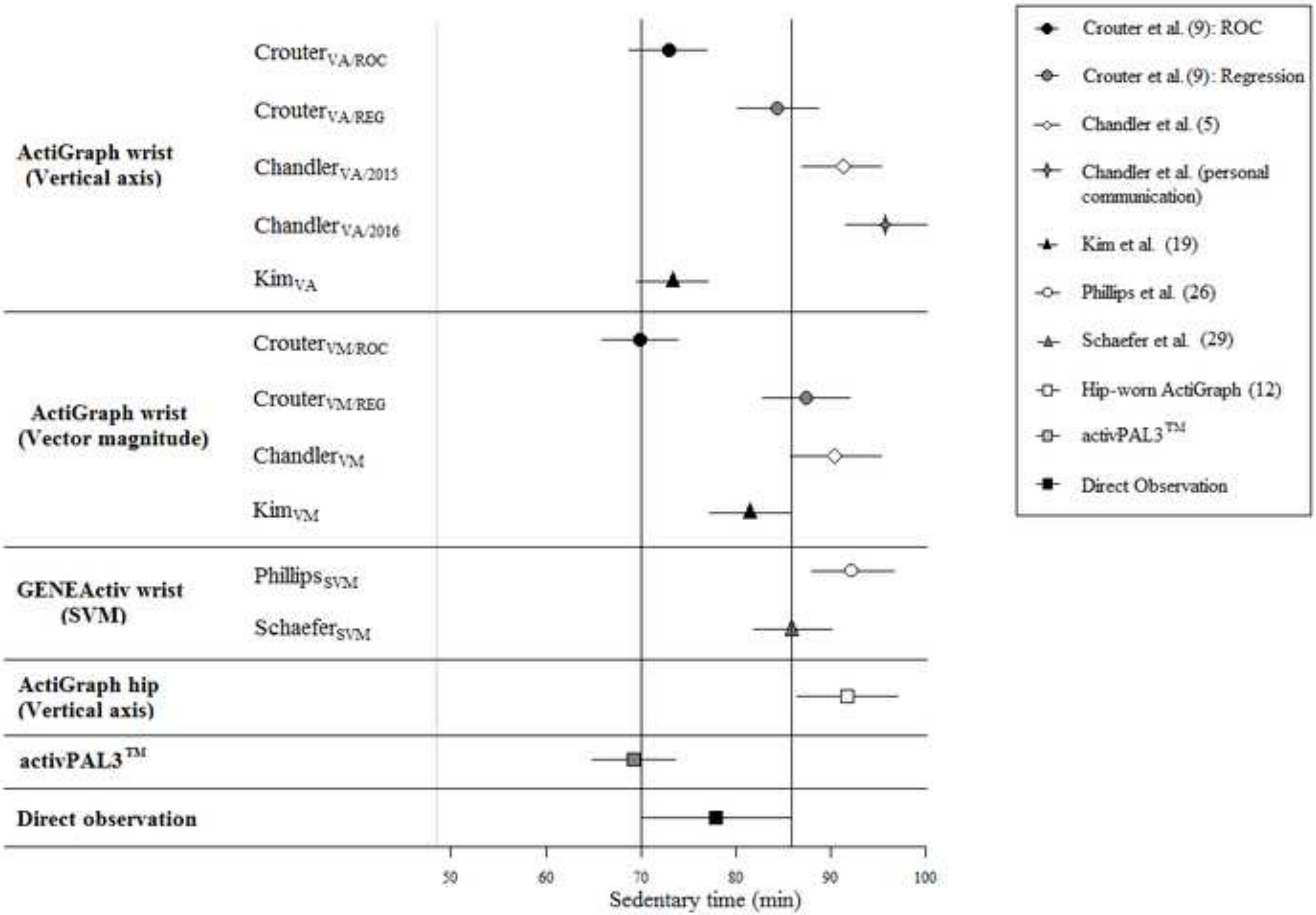
	Cut-point	Se %	95% CI	Sp %	95% CI	ROC-AUC	95% CI
ActiGraph wrist (vertical axis)	$Crouter_{VA/ROC}$						
	5-8y	82.0	81.5 - 82.5	73.6	73.0 - 74.1	0.78	0.77 - 0.78
	9-12y	72.1	71.7 - 72.6	76.5	76.0 - 77.0	0.74	0.74 - 0.75
	$Crouter_{VA/REG}$						
	5-8y	81.9	81.4 - 82.4	76.3	75.8 - 76.8	0.79	0.79 - 0.80
	9-12y	83.3	82.8 - 83.7	66.5	66.0 - 67.0	0.75	0.75 - 0.75
	$Chandler_{VA/2015}$						
	5-8y	86.2	85.7 - 86.6	72.2	71.7 - 72.7	0.79	0.79 - 0.80
	9-12y	87.0	86.6 - 87.4	62.1	61.6 - 62.6	0.75	0.74 - 0.75
	$Chandler_{VA/2016}$						
	5-8y	89.0	88.6 - 89.4	68.8	68.2 - 69.3	0.79	0.79 - 0.79
	9-12y	89.4	89.0 - 89.8	58.8	57.5 - 58.5	0.74	0.73 - 0.74
	Kim_{VA}						
	5-8y	87.8	86.2 - 89.3	83.7	81.8 - 85.4	0.86	0.85 - 0.87
	9-12y	89.5	88.0 - 90.8	83.2	81.5 - 84.8	0.86	0.85 - 0.87
ActiGraph wrist (vector magnitude)	$Crouter_{VM/ROC}$						
	5-8y	83.2	82.7 - 83.6	71.0	70.4 - 71.6	0.77	0.77 - 0.78
	9-12y	73.0	72.5 - 73.4	73.6	73.0 - 74.1	0.73	0.73 - 0.74
	$Crouter_{VM/REG}$						
	5-8y	83.2	82.7 - 83.7	73.6	73.1 - 74.1	0.78	0.78 - 0.79
	9-12y	83.5	83.1 - 84.0	62.3	61.8 - 62.8	0.73	0.73 - 0.73

	Chandler$_{VM}$						
	5-8y	84.8	84.3 - 85.3	71.5	71.0 - 72.1	0.78	0.78 - 0.79
	9-12y	84.8	84.4 - 85.3	59.6	59.1 - 60.2	0.72	0.72 - 0.73
	Kim$_{VM}$						
	5-8y	93.6	92.3 - 94.7	77.0	74.9 - 79.0	0.85	0.84 - 0.86
	9-12y	93.5	92.3 - 94.5	71.3	69.3 - 73.2	0.82	0.81 - 0.83
GENEActiv wrist (signal vector magnitude)	Phillips$_{SVM}$						
	5-8y	87.5	87.4 - 87.7	72.9	72.7 - 73.0	0.80	0.80 - 0.80
	9-12y	86.8	86.7 - 87.0	73.3	73.1 - 73.4	0.80	0.80 - 0.80
	Schaefer$_{SVM}$						
	5-8y	82.6	82.4 - 82.7	75.4	75.2 - 75.6	0.79	0.79 - 0.79
	9-12y	83.6	83.4 - 83.7	75.1	74.9 - 75.2	0.79	0.79 - 0.79
activPAL3™	5-8y	97.9	97.8 - 98.0	87.0	86.9 - 87.2	0.92	0.92 - 0.93
	9-12y	97.7	97.6 - 97.8	95.9	95.8 - 96.0	0.97	0.97 - 0.97
ActiGraph hip (vertical axis)	5-8y	92.7	92.1 - 93.3	76.3	75.4 - 77.2	0.85	0.84 - 0.85
	9-12y	93.6	93.0 - 94.1	75.9	75.0 - 76.7	0.85	0.84 - 0.85

Notes Table 4: Se: sensitivity; Sp: specificity; CI: confidence intervals; ROC-AUC: area under the receiver operating curve; VA: vertical axis; VM: vector magnitude; SVM: gravity-subtracted signal vector magnitude; c: counts; s: seconds; g: gravity; gs: g · seconds.

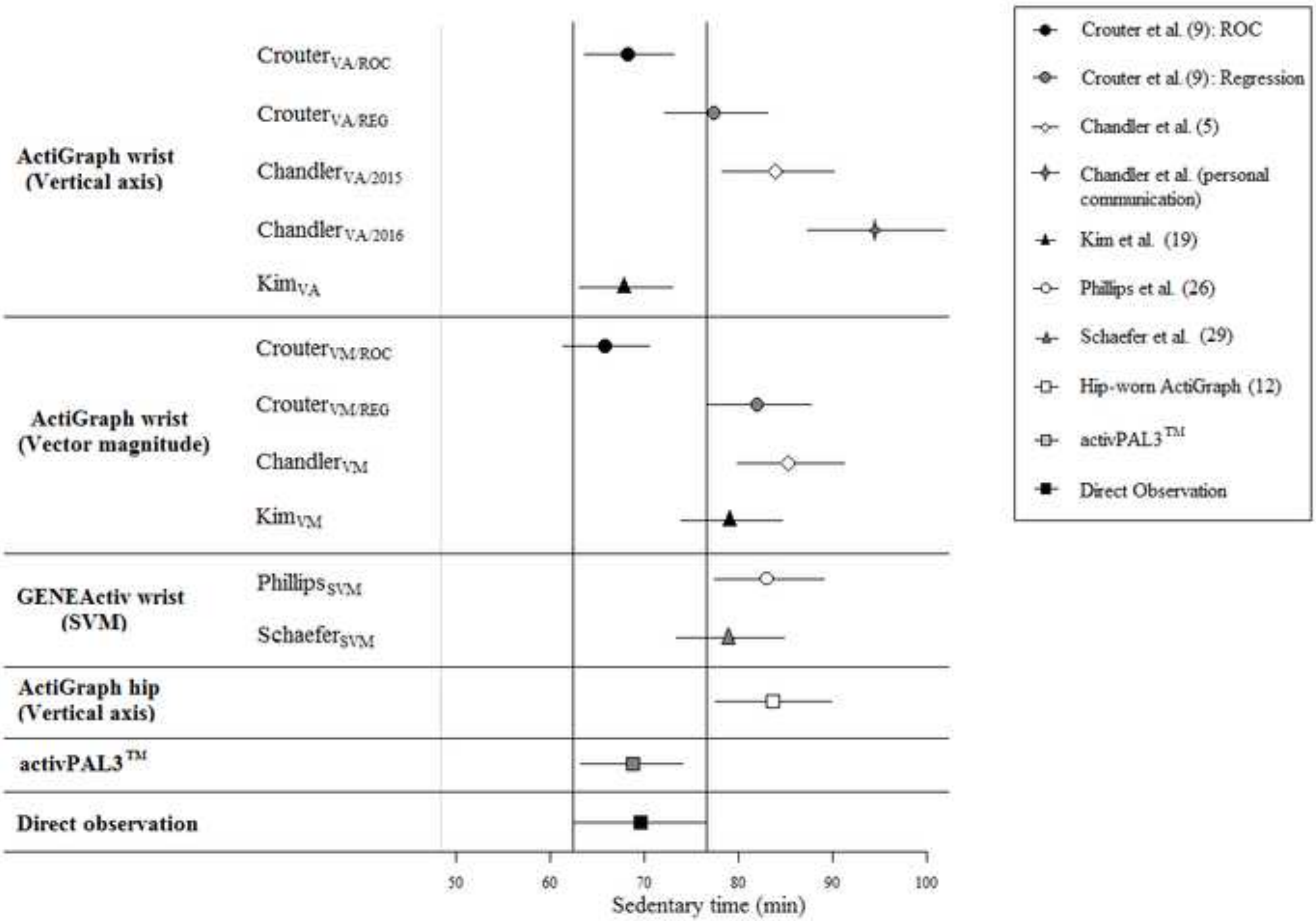
Figure

Click here to download Figure Figure1_5-8y.revisions.tif



Figure

Click here to download Figure Figure2_9-12y.revisions.tif



Click here to access/download
Supplemental Data File (.doc, .tif, pdf, etc.)
Supplemental Digital Content 3.docx

Click here to access/download

Supplemental Data File (.doc, .tif, pdf, etc.)

Supplemental Digital Content 4.revisions.tif

Click here to access/download
Supplemental Data File (.doc, .tif, pdf, etc.)
Supplemental Digital Content 5.docx