Title

CHAP-Adult: A Reliable and Valid Algorithm to Classify Sitting and Measure Sitting Patterns Using Data From Hip-Worn Accelerometers in Adults Aged 35.

Journal
Journal for the Measurement of Physical Behaviour, 5(4)

Authors
Bellettiere, John
Nakandala, Supun
Tuz-Zahra, Fatima
et al.

Publication Date
2022-12-01

DOI
10.1123/jmpb.2021-0062

Peer reviewed

Published in final edited form as:
J Meas Phys Behav. 2022 December ; 5(4): 215–223. doi:10.1123/jmpb.2021-0062.

CHAP-Adult: A Reliable and Valid Algorithm to Classify Sitting and Measure Sitting Patterns Using Data From Hip-Worn Accelerometers in Adults Aged 35+

John Bellettiere[1], **Supun Nakandala**[2], **Fatima Tuz-Zahra**[1], **Elisabeth A.H. Winkler**[3], **Paul R. Hibbing**[4], **Genevieve N. Healy**[3], **David W. Dunstan**[5,6], **Neville Owen**[5,7], **Mikael Anne Greenwood-Hickman**[8], **Dori E. Rosenberg**[8], **Jingjing Zou**[1], **Jordan A. Carlson**[4,9], **Chongzhi Di**[10], **Lindsay W. Dillon**[1], **Marta M. Jankowska**[11], **Andrea Z. LaCroix**[1], **Nicola D. Ridgers**[12], **Rong Zablocki**[1], **Arun Kumar**[2], **Loki Natarajan**[1]

[1]Herbert Wertheim School of Public Health and Human Longevity Science, University of California San Diego, La Jolla, CA, USA

[2]Department of Computer Science and Engineering, University of California San Diego, La Jolla, CA, USA

[3]School of Public Health, the University of Queensland, Brisbane, QLD, Australia

[4]Center for Children's Healthy Lifestyles & Nutrition, Children's Mercy Hospital, Kansas City, MO, USA

[5]Baker Heart and Diabetes Institute, Melbourne, VIC, Australia

[6]Mary MacKillop Institute for Health Research, Australian Catholic University, Melbourne, VIC, Australia

[7]Centre for Urban Transitions, Swinburne University of Technology, Melbourne, VIC, Australia

[8]Kaiser Permanente Washington Health Research Institute, Seattle, WA, USA

[9]Department of Pediatrics, University of Missouri–Kansas City, Kansas City, MO, USA

[10]Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, WA, USA

[11]Qualcomm Institute/Calit2, University of California San Diego, La Jolla, CA, USA

[12]School of Exercise and Nutrition Sciences, Institute for Physical Activity and Nutrition, Deakin University, Geelong, VIC, Australia

Abstract

Background: Hip-worn accelerometers are commonly used, but data processed using the 100 counts per minute cut point do not accurately measure sitting patterns. We developed and validated a model to accurately classify sitting and sitting patterns using hip-worn accelerometer data from a wide age range of older adults.

---

corresponding author is Bellettiere (jbellettiere@health.ucsd.edu).
Author Contributions: Bellettiere and Nakandala are co-first authors. Kumar and Natarajan are co-senior authors.

Methods: Deep learning models were trained with 30-Hz triaxial hip-worn accelerometer data as inputs and activPAL sitting/nonsitting events as ground truth. Data from 981 adults aged 35–99 years from cohorts in two continents were used to train the model, which we call CHAP-Adult (Convolutional Neural Network Hip Accelerometer Posture-Adult). Validation was conducted among 419 randomly selected adults not included in model training.

Results: Mean errors (activPAL − CHAP-Adult) and 95% limits of agreement were: sedentary time −10.5 (−63.0, 42.0) min/day, breaks in sedentary time 1.9 (−9.2, 12.9) breaks/day, mean bout duration −0.6 (−4.0, 2.7) min, usual bout duration −1.4 (−8.3, 5.4) min, alpha .00 (−.04, .04), and time in ≥30-min bouts −15.1 (−84.3, 54.1) min/day. Respective mean (and absolute) percent errors were: −2.0% (4.0%), −4.7% (12.2%), 4.1% (11.6%), −4.4% (9.6%), 0.0% (1.4%), and 5.4% (9.6%). Pearson's correlations were: .96, .92, .86, .92, .78, and .96. Error was generally consistent across age, gender, and body mass index groups with the largest deviations observed for those with body mass index ≥30 kg/m^2.

Conclusions: Overall, these strong validation results indicate CHAP-Adult represents a significant advancement in the ambulatory measurement of sitting and sitting patterns using hip-worn accelerometers. Pending external validation, it could be widely applied to data from around the world to extend understanding of the epidemiology and health consequences of sitting.

Keywords

sedentary behavior; activity classification; computational methods; neural networks; validation; machine learning

---

Time spent sitting increases risk of developing major chronic diseases and is associated with multiple adverse health outcomes (Dempsey et al., 2020; Katzmarzyk et al., 2019; Saunders et al., 2020). Sitting is most often operationalized as amount of time accumulated during a day or week. However, emerging epidemiologic and experimental evidence shows that the *temporal pattern* of accumulating sitting time in a comparatively prolonged versus more interrupted manner (henceforth referred to as sitting patterns; Tremblay et al., 2017) is also important to public health inquiry (Bellettiere et al., 2017; Diaz et al., 2016; Dunstan et al., 2012; Hartman et al., 2018; Healy et al., 2011; Owen et al., 2020; Saunders et al., 2018). The mounting evidence has led several national and international public health guideline development committees to call for more research on the topic, both to establish its importance in chronic disease etiology and subsequently to support development of quantitative public health recommendations (Dempsey et al., 2020; Katzmarzyk et al., 2019; Saunders et al., 2020).

Accurate, scalable measurement of sitting time and sitting patterns is critically needed to inform future iterations of public health guidelines. Most prospective cohorts (Wijndaele et al., 2015) have used hip-worn accelerometers and processed the data using the 100 counts per minute (cpm) cut point (Migueles et al., 2017). However, this method has low accuracy for measuring posture and postural transitions (Barreira et al., 2015) and consequently leads to low-quality measures of sitting patterns. The estimated percent error in measuring sitting patterns using the 100 cpm is evident across age groups, and ranges from 40% to over 300% depending on the sedentary metric being used (Bellettiere et al., 2021; Carlson et al., 2019). The degree of misclassification generated by this cut point classification leaves some uncertainty around quantitative recommendations for sitting patterns. A few large cohorts (Stamatakis et al., 2020) use devices worn on the thigh that are optimized and often programmed specifically to detect posture and can accurately measure both sitting time and sitting patterns when compared with direct observation (Lyden et al., 2012). However, data from these studies are limited, and the pool of available information regarding sitting patterns could be rapidly expanded by developing new data processing methods for hip-worn accelerometer data to generate measures of sitting patterns that converge with estimates from thigh-worn devices. The application of these new processing methods to both newly collected hipwornaccelerometer data and the large amount of archival data could produce accurate and homogenous norms and risk estimates related to sitting patterns and health, both of which are needed to generate quantitative public health recommendations.

Our research group has recently demonstrated that deep learning models can be used to improve the utility of hip-worn monitors (Greenwood-Hickman et al., 2021). Using a combined Convolutional Neural Network (CNN) + bidirectional long shortterm memory network (BiLSTM) model architecture in a sample of older adults, we developed a novel method for predicting sitting postures from hip-worn accelerometer data (e.g., ActiGraph) in older adults by training against concurrently thigh-worn activPAL inclinometers. This algorithm, which we named CHAP (**C**NN **H**ip **A**ccelerometer **P**osture), substantially out-performed previous, widely used data processing methods (e.g., 100 cpm cut point based method) in older adults (≥65 years), achieving 93% agreement in predicted minute-level posture (i.e., sitting vs. upright/nonsitting) and 83% sensitivity for detecting sit to stand postural transitions (Greenwood-Hickman et al., 2021). CHAP was trained for use in an older adult population, but norms and risk estimates for sitting patterns and how they relate to health are needed across the age spectrum. Especially since many existing cohort studies that collect accelerometer data include younger, middle-aged, and older adults. Therefore, accurate posture classification is needed across a broader age range. Ideally, a single postural detection algorithm could be used to make accurate predictions for all adults.

Building on CHAP to fashion a procedure suitable for adults of a broader age range, we created the CHAP-Adult algorithm bytraining a new and more generalizable CNN+BiLSTM model using data from adults across a broader age range (35–99 years). We leveraged data from participants of two cohort studies who concurrently wore ActiGraph and activPAL accelerometers for up to 7dayswhile going about their usual (i.e., free living) behavior patterns (N = 1,397). (Healy et al., 2015; Rosenberg et al., 2020) The objective of this study was to assess the validity and reliability of CHAP-Adult, which we did using a separate, randomly selected validation sample of 419 adults. We also assessed reliability and validity separately by subgroups of age, gender, and body mass index (BMI).

Methods

Participants

Data for this study come from two distinct studies among adults who concurrently wore an activPAL on their thigh and an ActiGraph GT3X+ on their hip: the Australian Diabetes, Obesity, and Lifestyle study (AusDiab) and the Adult Changes in Thought (ACT) study.

The AusDiab is a population-based epidemiologic cohort that initially enrolled adults aged 25 years and older throughout Australia. The original complex survey sampling methods were previously published (Dunstan et al., 2002). In brief, three data collection points (1999–2000, 2004–2005, and 2011–2012) were undertaken with each assessment including questionnaires and biomedical assessments conducted at a local testing site. Of the 11,247 adults who completed the baseline assessment in 1999–2000, 4,562 (all ≥ 35 years) attended one of 46 testing centers across Australia in the 2011–2012 follow-up (Tanamas et al., 2013). From this group, a subsample of 1,014 ambulatory community-living participants were invited to join an ancillary activity monitor study, described in detail elsewhere (Healy et al., 2015). Participants were asked to wear the activPAL3™ (thigh) and the ActiGraph GT3X+ (hip) for seven consecutive days, with the ActiGraph removed for sleeping and the activPAL worn continuously. Of those approached, 77% (n = 782) provided informed written consent and participated. Protocols for the study were approved by the Alfred Health Human Ethics Committee (project no. 39/11).

The ACT began in 1994 as an ongoing longitudinal cohort study investigating risk factors for the development of dementia and a wide range of cognitive and noncognitive factors of healthy aging. ACT recruited adults aged over 65 years randomly sampled from the King County membership of Group Health Cooperative of Puget Sound (now Kaiser Permanente Washington). As of 2005, ACT continually replaced participants who have died, were lost to follow-up, or who were diagnosed with dementia, in order to maintain a consistent enrollment of approximately 2,000 participants. Study participants attend biennial assessment visits. Beginning with biennial visits in April 2016, ACT participants were invited to participate in an activity monitor substudy (ACT-AM). Those who consented were invited to wear an ActiGraph GT3X+ accelerometer on the hip (ActiGraph LLC), an activPAL micro3 on the thigh (PAL Technologies), or, if willing, both continuously for seven consecutive days. Participants were excluded from ACT-AM if they were wheelchair bound, or receiving hospice or care for a critical illness, or resided in a nursing home, or, if memory problems became evident during testing. In total, 1,211 ACT participants provided written informed consent to participate in ACT-AM and wore at least one device. Details on the ACT-AM cohort were published elsewhere (Rosenberg et al., 2020) and protocols for the study were approved by the Kaiser Permanente Washington Institutional Review Board.

Data Processing

ActiGraph data originally collected at 30 Hz from both studies were obtained so that identical data processing methods could be applied. The 30-Hz data were subsequently converted to 1-min epochs using ActiLife software. ActiGraph nonwear was detected using the commonly used (Migueles et al., 2017) Choi algorithm applied to vector magnitude cpm using a 90-min window, 30-min stream-frame, and 2-min tolerance (Choi et al., 2011, 2012). For each night of device wear, participants completed sleep logs and the resulting data were used to identify in-bedandout-of-bedtime. TheactivPAL data were processed using PALbatch (version 7.2.32; PAL Technologies, original VANE algorithm) with the default 10-s minimum sitting/upright period and converted to event level files. The activPAL “sitting/lying” events are referred to hereinafter as “sitting.” ActiGraph and activPAL data were then plotted together with sleep log and nonwear data using heat maps (Greenwood-Hickman et al., 2021) that were then visually inspected by trained analysts.

Inclusion Criteria for This Training and Validation Study

Figure 1 shows a STROBE diagram detailing participant allocation into the training and validation subsamples. Analyses excluded any time either device was not worn or the person was in bed, and no exclusions were made based on minimum wear time. A total of 51 (6.5%) participants from AusDiab and 231 (19.1%) from ACT did not have concurrently worn devices and were excluded from the analyses. Because ACT participants could choose to wear a single device, this exclusion rate is higher for ACT participants. Data were visually inspected for all AusDiab and ACT participants who wore both devices. In some cases, there was evidence that the devices gradually fell out of sync with each other in a phenomenon called time drift (Steel et al., 2019), possibly due to malfunction or different internal clocks (see the supplemental digital content in Greenwood-Hickman et al., 2021, for an example and description of drift). This occurred in 38 AusDiab participants and 271 ACT participants, with differential time drift rates possibly due to different internal clocks used in the activPAL3™ and the activPAL micro. To ensure that our prediction algorithm used the most accurate and optimal criterion data, we excluded samples with time drift for training and for validation.

Data from the resulting 688 AusDiab and 709 ACT participants were then randomly assigned within each cohort to the for-training data set (70%) or the for-validation data set (30%). To ensure model accuracy was not biased by overfitting that occurs when the model is trained and validated using the same data, the team who trained the CHAP-Adult algorithm used only the for-training data set and had no access to the for-validation data set.

Sitting Time, Prolonged Sitting Time, Breaks in Sitting, and Sitting Accumulation

Total sitting time was measured as the average number of minutes per day sitting. There is no agreed-upon metric for measuring sitting patterns, so we present validation results for postural transitions, prolonged sitting time, and three common pattern indicators based on the duration of sitting bouts. Time spent in ≥30-min bouts (herein termed *prolonged sitting*) was computed as the average minutes per day spent in bouts ≥30 min. The number of breaks in sitting (sometimes termed sit–stand transitions, or interruptions) is approximately equal to the number of sitting bouts and the number of nonsitting bouts. For convenience, in this study we have used the term “number of breaks” and report the average number of sitting bouts per day. Mean sitting bout duration was computed as the arithmetic mean of all sitting bout durations in minutes. Using procedures outlined by Chastin and Granat (2010) and Chastin et al. (2015), we calculated usual bout duration (in minutes) via nonlinear regression, and alpha (see Supplementary Methods [available online] for equations). Alpha is a unitless parameter that indicates how steep the distribution of bout duration is, which is assumed to be power–law (see Supplemental Figure 1 in Bellettiere et al., 2017). Unlike mean and usual bout duration, higher values of alpha indicate a more interrupted sitting pattern. All measures were based on strict bouts that contained only sitting and did not include modified bouts with a tolerance for upright behaviors.

The CNN+BiLSTM Model and Training

Details of the machine learning architecture and training procedures have been previously published (Greenwood-Hickman et al., 2021; Nakandala et al., 2021). Briefly, the input data (raw triaxial 30-Hz ActiGraph data) and the ground truth data (activPAL event-level data classified as sitting or not sitting using majority rules for each epoch) during awake wear time were split into 10-s nonoverlapping windows. Features in the ActiGraph data that could differentiate sitting from not sitting were automatically learned using a CNN model. These features were then input into a BiLSTM to learn the features required to recognize temporal patterns that could further distinguish sitting from not sitting. This network was specifically designed to identify the timing of transitioning from sitting to not sitting. The CHAP-Adult algorithm was trained by first separating each person in the for-training data set into two randomly selected groups, a training group (n = 782) and a test group (n = 196). The for-validation data set was not used during this process. During training, data from the training group were fed into the CNN+ BiLSTM model, classifying each 10-s window of ActiGraph data as sitting or not sitting. Next, these predicted classifications were compared with the true labels of the time-matched activPAL data and the learnable parameters of the CNN+BiLSTM model were updated to maximize the accuracy of the predicted classifications using neural network training methods. This process was repeated for several iterations until the model achieved a good classification accuracy. Model selection was then performed to find the best model configuration. Several models with different hyperparameter (i.e., nonlearnable model parameters) values were trained on the training group and the model that had the highest accuracy on the test set was chosen, and ultimately dubbed CHAP-Adult. Finally, CHAP-Adult was applied to the for-validation sample and then the processed data were used to compute the five sitting pattern metrics.

Statistical Methods

Participant characteristics were summarized for each cohort and for the for-training and for-validation groups and compared between groups with two-sample t tests (continuous variables) or chi-square tests (categorical variables). Validation statistics were calculated and reported only from the for-validation data set. To assess how well CHAP-Adult predicted each instance of sitting versus not sitting at the 10-s epoch, sensitivity, specificity, positive predictive value (PPV), and negative predictive value were computed separately for each individual, then summarized using boxplots (Figure 2). To assess how well CHAP-Adult captured the timing of postural transitions, for each individual separately, we used the transition pairing method (Hibbing et al., 2020) to pair activPAL-labeled transitions and CHAP-Adult-labeled transitions, with a 1-min lag time tolerance. From these matches sensitivity (fraction of activPAL transitions matched to a CHAP-Adult transition) and PPV (fraction of CHAP-Adult transitions matched to an activPAL transition) were calculated and summarized using boxplots (Figure 2). Agreement between activPAL and CHAP-Adult measures was assessed for sitting time, prolonged sitting time, breaks in sitting, and sitting accumulation by computing mean error with 95% limits of agreement, percent error, absolute mean error, mean absolute percent error (MAPE), Pearson correlation coefficient, and the Lin's (1989) concordance correlation coefficient. Agreement was assessed for overall and also reported within subgroups based on age (35–49, 50–64, and ≥65 years), sex (male and female), and BMI (<30 and ≥30 kg/m^2).

Results

Overall, ages ranged from 35 to 90+ years old in the combined ACT and AusDiab data set. Most participants were women (58.5% of ACT and 56.3% of AusDiab), and approximately 23% had a BMI ≥30 kg/m^2 (Table 1). There were notably more adults with high school education or less in AusDiab (29.6%) than in ACT (8.7%). There was no evidence of randomization failure or imbalanced data sets, with no meaningful differences on measured variables between the for-training and for-validation data sets.

When applied to the for-validation data set, the CHAP-Adult algorithm correctly classified as sitting an average (*SD*) of 95.3% (4.7%) of the activPAL sitting time (sensitivity), while the algorithm correctly classified 89.8% (8.1%) of activPAL not sitting (specificity). On average 93.5% (6.0%) of the predicted sitting instances and 92.2% (8.4%) of predicted nonsitting instances were correct (PPV and negative predictive value respectively). To accurately measure sitting patterns, it is critical to get the timing of postural transitions correct as well as the total number. An average of 74.4% (10.2%) of the activPAL transitions were detected within a 1-min window by CHAP-Adult (sensitivity) and 77.6% (12.4%) of the transitions that CHAP-Adult predicted were true postural transitions recorded on the activPAL within a 1-min window (PPV).

The estimates of sitting, prolonged sitting, breaks in sitting, and sitting patterns produced using CHAP-Adult correlated strongly with those derived from the activPAL, with Pearson's correlations (r) ranging between .78 and .96. Correlations were lower for alpha (r = .78) and the number of breaks per day (r = .86) than for total sitting time (r = .96), time spent in prolonged (≥30 min) sitting bouts (r = .96), usual bout duration (r = .92), and mean bout duration (r = .92). Concordance correlation coefficients were nearly identical to the Pearson's correlations.

There was good agreement in sitting time, prolonged sitting time, breaks in sitting, and sitting patterns (Table 2), with MAPE <5% for total sitting time and alpha, between 5% and 10% for prolonged sitting time and usual bout duration, and only slightly higher than 10% for mean bout duration and postural transitions. Mean differences (activPAL − CHAP-Adult) were also modest, ranging from 0% to 5.4% of the activPAL value, and indicating on average a small amount of overestimation of total sitting time, mean bout duration, usual bout duration, and prolonged sitting time and underestimation of postural transitions.

Validation results in each population subgroup (Table 3 and Supplementary Table S1 [available online]) resembled those for the overall for-validation sample with consistently high correlations and reasonably modest error that was similar though not identical in each subgroup. For example, BMI-stratified Pearson's correlations were: total sitting time, $r_{<30\ \mathrm{kg/m}}2$ = .96 and $r_{\geq 30\ \mathrm{kg/m}}2$ = .95; breaks in sitting time, $r_{<30\ \mathrm{kg/m}}2$ = .88 and $r_{\geq 30\ \mathrm{kg/m}}2$ = .79; and mean bout duration, $r_{<30\ \mathrm{kg/m}}2$ = .95 and $r_{\geq 30\ \mathrm{kg/m}}2$ = .88, with MAPEs ranging from 3.7% to 16.7% across the different sedentary pattern metrics. Importantly, across age categories, Pearson's correlations were ≥.95 for total sitting time, ≥.82 for breaks in sitting time, and ≥.9 for mean bout duration, with MAPEs ranging from 3.5% to 13.9% for these sedentary pattern metrics across age categories. Error across age categories was larger for usual bout duration $MAPE_{35–49}$ = 5.8%, $MAPE_{50–64}$ = 6.4%, and $MAPE_{\geq 65}$ = 11.6% and for prolonged sitting $MAPE_{35–49}$ = 7.3%, $MAPE_{50–64}$ = 8.6%, and $MAPE_{\geq 65}$ = 10.4%, although remaining relatively low with Pearson's correlations ranging between .90 and .99.

Discussion

Our findings demonstrate the CHAP-Adult algorithm can provide accurate classification of sitting and sitting patterns for men and women over a wide range of adult ages, expanding on our previous successful work with the CNN and BiLSTM model architecture in older adults (Greenwood-Hickman et al., 2021). The algorithm performed well at distinguishing sitting from not sitting (mean balanced accuracy = 93% ± 5%) and detected 74% ± 10% of postural transitions. Consequently, the measures derived using this algorithm displayed good validity not only for total sitting time (MAPE = 4%) but also for breaks in sitting time (MAPE = 11.6%), prolonged sitting time (MAPE = 9.6%), and sitting accumulation patterns (MAPE = 1%–12%), which have historically been more difficult to capture accurately. Agreement metrics were consistent across age, gender, and BMI subgroups, with mean error approximately three times lower than when metrics are computed using the common 100 cpm data processing method (discussed below). Overall, CHAP-Adult, a single algorithm that can be used to detect sitting behaviors across a wide range of adult age groups, represents a significant advancement in the ambulatory measurement of sitting and sitting patterns using hip-worn monitors. If validity in an external adult population can also be demonstrated, it would be suitable for extracting valid measures from the numerous historic studies that have collected raw triaxial hip-worn acceleration data (Wijndaele et al., 2015), and could extend our understanding of the epidemiology and health consequences of sitting.

Comparison With Extant Literature

Sedentary behavior, by definition, is waking time spent sitting or reclining with energy expenditure below 1.5 metabolic equivalents (Tremblay et al., 2017). While hip-worn ActiGraph devices cannot accurately distinguish posture, estimates of total time per day spent sedentary using the most common 100 cpm data processing method are typically close to the estimates of sitting time derived from direct observation or activPAL—with mean errors from several studies across a wide age range of adults ranging from 5% to 20% (Aguilar-Farías et al., 2014; Bellettiere et al., 2021; Koster et al., 2016; Kuster et al., 2020; Lyden et al., 2012). This mean error can be further reduced to 2%–3% by increasing the cut point used to classify time as sedentary to 200 cpm (Aguilar-Farías et al., 2014; Koster et al., 2016), or by implementing either the Kuster algorithm (4% error; Kuster et al., 2020), the original CHAP (0% error; Greenwood-Hickman et al., 2021), or CHAP-Adult (2% error).The similar error rates mean that many of these data processing methods would provide similar results, on average, for measuring total sitting time.

Measuring sitting patterns, on the other hand, has been notoriously challenging. Primary because it is predicated on the accurate identification of breaks between sitting or reclining and being upright, which happen rapidly and relatively infrequently in real life. Accordingly, hip-worn accelerometer data processed using the 100 cpm method have historically overestimated *the frequency of sit–stand transitions* (also known as breaks in sedentary behavior) in younger (Barreira et al., 2015; Lyden et al., 2017), middle-aged (Kuster et al., 2020), and older (Bellettiere et al., 2021) adults with mean error ranging from 63% to 99%. There are two primary sources of error using hip-worn accelerometers processed using the 100 cpm cut point: (a) long sitting bouts are artificially broken by seated movement such as fidgeting and wiggling; and (b) standing without ambulation gets misclassified as sitting, and then any subsequent stepping while already standing is incorrectly registered as a "sitting break." Both sources of error stem from an inability to distinguish posture when relying solely on the accelerometer signal from vertical acceleration at the hip. To more accurately assess posture using data from hip-worn devices, Kuster et al. (2020) trained a random forest machine learning algorithm (using ground truth data from the activPAL) using 26 predefined signal features from hip-worn ActiGraph data as inputs (Kuster et al., 2020). This algorithm reduced error in the number of sit– stand transitions to just 18%, thereby improving the accuracy of sitting pattern measurement. Notably, the deep learning characteristics of CNN, adopted by our CHAP models, automatically determine signal features, thus obviating the need to define a priori features. Furthermore, adding BiLSTM models to the CNN architecture leverages the concept that when predicting human behavior (i.e., here, transitions from one posture to the next), modeling the behavior that occurred directly before and after the behavior aligns the prediction process with theoretical models of behavior.(Hovell et al., 2009) This also effectively leverages the within-person correlated data to improve prediction accuracy. This advancement resulted in CHAP-Adult reducing error in the number of breaks to 4%. Importantly, the low error for estimating the number of breaks using CHAP-Adult was consistent across age, gender, and BMI, demonstrating consistent performance across subgroups of adults.

The Public Health Implications of CHAP-Adult

The U.S. Physical Activity Guideline Advisory Committee and the World Health Organization Sedentary Behavior Guideline Committee recognized heterogeneity across studies in the measurement of bout lengths and breaks in sedentary time as an "important area of future research" and a reason for the lack of evidence regarding how bouts and breaks in sedentary behavior are related to health outcomes (Dempsey et al., 2020; Katzmarzyk et al., 2019). CHAP-Adult was intentionally designed to fill these gaps in two ways: first by being trained and validated among community-living adults from two continents across a wide age range (35–99 years); and second by being designed for application to a very common accelerometer and wear protocol (hip-worn ActiGraph GT3X+) for free-living physical behavior measurement (Wijndaele et al., 2015). With CHAP-Adult, many large epidemiologic cohorts that use *either* hip-worn GT3X+ or thigh-worn activPAL accelerometers can produce estimates of sitting patterns that are in high agreement. This broad applicability will enable consistent and accurate estimates that will help the field of sedentary behavior epidemiology generate norms for sitting patterns. It will also facilitate generation of estimates of the health risks associated with sitting patterns from various, heterogenous cohorts around the world—estimates that can more easily be harmonized for metaanalytic purposes since they would be generated using the same pretrained model (CHAP-Adult). Finally, it is our hope that this widespread availability of accurate data on sitting patterns will spur rigorous investigation of how sitting patterns and total sitting volume are jointly and independently associated with health. Ultimately, the new data generated using CHAP-Adult can therefore help the field progress toward specific quantitative (time based) recommendations that include a threshold for limiting sitting time and recommendations regarding how sitting can be broken up throughout the day.

Strengths and Limitations

CHAP-Adult was trained using 30-Hz data from the triaxial ActiGraph GT3X+ as inputs and activPAL-classified postures as the ground truth, both set up strengths and limitations. Use of triaxial data preclude the application of this pretrained algorithm to data collected using uniaxial accelerometers, limiting applicability, although new deep learning algorithms could be developed for uniaxial accelerometer data streams, provided a large volume of raw acceleration data as well as data from a concurrent ground truth assessment were also available. We also are unsure how well CHAP-Adult will work with triaxial data collected using other devices that collect raw acceleration data (e.g., GENEActiv and Axivity), but we suspect it would perform well as long as a similar wear protocol was followed. CHAP-Adult was not designed for use with data from wrist-worn accelerometers; new algorithms are needed for wrist-based devices. While the activPAL is highly accurate for measuring sitting and sit–stand transitions compared with direct observation (Lyden et al., 2012), it is not without error. Thus, we have shown convergent and concurrent validity, but future work should also look at criterion validity. However, the error in using this ambulatory monitor must be viewed against the benefits of including a large and diverse sample from the United States and Australia, and measuring physical behavior during free-living environments for an average of 7 days. This greatly enhances the utility and therefore the potential impact of CHAP-Adult. Other limitations are that we could not assess specificity and negative predictive value for transitions because the class is highly imbalanced and the validation metrics would be artificially inflated. This algorithm was also developed using data on adults ≥35 years of age, and future similar studies among children and younger adults are warranted. Analyses did not consider clustering or stratification from the multiple cohorts or within cohorts (e.g., AusDiab stratified multistage sampling), however, since this validation is concerned with pairwise comparisons, the impact on the validation metrics is likely negligible.

Conclusions

The CHAP-Adult is a pretrained deep learning algorithm that leverages CNN and BiLSTM architecture to classify sitting and postural transitions in adults over a wide age range. Our validation results demonstrate that CHAP-Adult produces estimates of sitting time and sitting patterns that are highly convergent with the activ-PAL, and does so within subgroups based on age, gender, and BMI. Additional external validation is needed to increase confidence in these findings. Ultimately, we will measure the success of CHAP-Adult by how often it is applied to data around the world. Application to data collected in many established and emerging national and international cohort studies will more strongly position the field of sedentary behavior epidemiology to establish scientifically sound population and subgroup sitting norms and determine how sitting and sitting patterns may be causally related to adverse health outcomes. Better identification of sitting patterns will also enable more accurate classification of the types of activities used to break up sitting (Blankenship et al., 2021), which is an emerging area of public health inquiry. Ultimately, widespread use of CHAP-Adult could generate data that contribute to national and international quantitative (as distinct from current broadly qualitative) recommendations for sitting and sitting patterns.

Open-sourced code and detailed vignettes for implementing CHAP-Adult are available at the following website: https://adalabucsd.github.io/DeepPostures/.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by grant number U01AG006781 and U19AG066567 from the National Institute on Aging and R01DK114945 from the National Institute of Diabetes and Digestive and Kidney Diseases. It was also supported by a Tobacco-Related Disease Research Program New Investigator Award (31KT1501); Principal Investigator: Bellettiere. It was also supported in part by a Hellman Fellowship, an NSF CAREER Award under award number 1942724, and a gift from VMware. Healy was supported by an National Health and Medical Research Council—Medical Research Future Fund (NHMRC–MRFF) Emerging Leadership Investigator Grant Emerging Leadership Investigator Grant (#1193815). The work also had funding support from the National Health and Medical Research Council (NHMRC grants 233200 and 1007544); Australian Government Department of Health and Aging; Abbott Australasia Pty Ltd.; Alphapharm Pty Ltd.; Amgen Australia; AstraZeneca; Bristol-Myers Squibb; City Health Center-Diabetes Service, Canberra; Department of Health and Community Services—Northern Territory; Department of Health and Human Services—Tasmania; Department of Health—New South Wales; Department of Health—Western Australia; Department of Health—South Australia; Department of Human Services—Victoria; Diabetes Australia; Diabetes Australia Northern Territory; Eli Lilly Australia; Estate of the Late Edward Wilson; GlaxoSmithKline; Jack Brockhoff Foundation; Janssen-Cilag; Kidney Health Australia; Marian & FH Flack Trust; Menzies Research Institute; Merck Sharp & Dohme; Novartis Pharmaceuticals; Novo Nordisk Pharmaceuticals; Pfizer Pty Ltd.; Pratt Foundation, Queensland Health; Roche Diagnostics Australia; Royal Prince Alfred Hospital, Sydney; Sanofi Aventis; sanofi-synthelabo; and the Victorian Government's Operational Infrastructure Support (OIS) Program. The funding sources had no involvement in the study or in the production of the manuscript.

References

Aguilar-Farías N, Brown WJ, & Peeters GMEEG. (2014). ActiGraph GT3X+ cut-points for identifying sedentary behaviour in older adults in free-living environments. Journal of Science and Medicine in Sport, 17(3), 293–299. 10.1016/j.jsams.2013.07.002 [PubMed: 23932934]

Barreira TV, Zderic TW, Schuna JM, Hamilton MT, & Tudor-Locke C. (2015). Free-living activity counts-derived breaks in sedentary time: Are they real transitions from sitting to standing? Gait & Posture, 42(1), 70–72. 10.1016/j.gaitpost.2015.04.008 [PubMed: 25953504]

Bellettiere J, Tuz-Zahra F, Carlson JA, Ridgers ND, Liles S, Greenwood-Hickman MA, ::: Natarajan L. (2021). Agreement of sedentary behavior metrics derived from hip- and thigh-worn accelerometers among older adults: With implications for studying physical and cognitive health. Journal for the Measurement of Physical Behaviour, 4(1), 79–88. 10.1123/jmpb.2020-0036 [PubMed: 34708190]

Bellettiere J, Winkler EAH, Chastin SFM, Kerr J, Owen N, Dunstan DW, & Healy GN. (2017). Associations of sitting accumulation patterns with cardio-metabolic risk biomarkers in Australian adults. PLoS One, 12(6), 1–17. 10.1371/journal.pone.0180119

Blankenship J, Winkler E, Healy G, Dempsey P, Bellettiere J, Owen N, & Dunstan D. (2021). Descriptive epidemiology of interruptions to free-living sitting time in middle-age and older adults. Medicine & Science in Sports & Exercise, 53(12), 2503–2511. 10.1249/mss.0000000000002750 [PubMed: 34310494]

Carlson JA, Bellettiere J, Kerr J, Salmon J, Timperio A, Verswijveren SJJM, & Ridgers ND. (2019). Day-level sedentary pattern estimates derived from hip-worn accelerometer cut-points in

8–12year-olds: Do they reflect postural transitions? Journal of Sports Sciences, 37(16), 1899–1909. 10.1080/02640414.2019.1605646 [PubMed: 31002287]

Chastin SFM, & Granat MH. (2010). Methods for objective measure, quantification and analysis of sedentary behaviour and inactivity. Gait & Posture, 31(1), 82–86. [PubMed: 19854651]

Chastin SFM, Winkler EAH, Eakin EG, Gardiner PA, Dunstan DW, Owen N, & Healy GN. (2015). Sensitivity to change of objectively-derived measures of sedentary behavior. Measurement in Physical Education and Exercise Science, 19(3), 138–147. 10.1080/1091367X.2015.1050592

Choi L, Liu Z, Matthews CE, & Buchowski MS. (2011). Validation of accelerometer wear and nonwear time classification algorithm. Medicine & Science in Sports & Exercise, 43(2), 357–364. 10.1249/MSS.0b013e3181ed61a3.Validation [PubMed: 20581716]

Choi L, Ward SC, Schnelle JF, & Buchowski MS. (2012). Assessment of wear/nonwear time classification algorithms for triaxial accelerometer. Medicine & Science in Sports & Exercise, 44(10), 2009–2016. 10.1249/MSS.0b013e318258cb36 [PubMed: 22525772]

Dempsey PC, Biddle SJH, Buman MP, Chastin S, Ekelund U, Friedenreich CM, ::: Bull F. (2020). New global guidelines on sedentary behaviour and health for adults: broadening the behavioural targets. International Journal of Behavioral Nutrition and Physical Activity, 17(1), 1–12. 10.1186/S12966-020-01044-0 [PubMed: 31898547]

Diaz KM, Howard VJ, Hutto B, Colabianchi N, Vena JE, Blair SN, & Hooker SP. (2016). Patterns of sedentary behavior in US middle-age and older adults: The REGARDS study. Medicine & Science in Sports & Exercise, 48(3), 430–438. 10.1249/MSS.0000000000000792 [PubMed: 26460633]

Dunstan DW, Kingwell BA, Larsen R, Healy GN, Cerin E, Hamilton MT, ::: Owen N. (2012). Breaking up prolonged sitting reduces postprandial glucose and insulin responses. Diabetes Care, 35(5), 976–983. 10.2337/dc11-1931 [PubMed: 22374636]

Dunstan DW, Zimmet PZ, Welborn TA, Cameron AJ, Shaw J, de Courten M, ::: McCarty DJ. (2002). The Australian diabetes, obesity and lifestyle study (AusDiab)—Methods and response rates. Diabetes Research and Clinical Practice, 57(2), 119–129. 10.1016/S0168-8227(02)00025-6 [PubMed: 12062857]

Greenwood-Hickman MA, Nakandala S, Jankowska MM, Rosenberg D, Tuz-Zahra F, Bellettiere J, ::: Natarajan L. (2021). The CNN hip accelerometer posture (CHAP) method for classifying sitting patterns from hip accelerometers: A validation study. Medicine & Science in Sports & Exercise, 53(11), 2445–2454. 10.1249/mss.0000000000002705 [PubMed: 34033622]

Hartman SJ, Marinac CR, Cadmus-Bertram L, Kerr J, Natarajan L, Godbole S, ::: Sears DD. (2018). Sedentary behaviors and biomarkers among breast cancer survivors. Journal of Physical Activity and Health, 15(1), 1–6. 10.1123/jpah.2017-0045 [PubMed: 28682735]

Healy GN, Matthews CE, Dunstan DW, Winkler EAH, & Owen N. (2011). Sedentary time and cardio-metabolic biomarkers in US adults: NHANES 2003–06. European Heart Journal, 32(5), 590–597. 10.1093/eurheartj/ehq451 [PubMed: 21224291]

Healy GN, Winkler EA, Owen N, Anuradha S, & Dunstan DW. (2015). Replacing sitting time with standing or stepping: Associations with cardio-metabolic risk biomarkers. European Heart Journal, 36(39), 2643–2649. 10.1093/eurheartj/ehv308 [PubMed: 26228867]

Hibbing PR, LaMunion SR, Hilafu H, & Crouter SE. (2020). Evaluating the performance of sensor-based bout detection algorithms: The transition pairing method. Journal for the Measurement of Physical Behaviour, 3(3), 219–227. 10.1123/jmpb.2019-0039 [PubMed: 34258524]

Hovell MF, Wahlgren DR, & Adams MA. (2009). The logical and empirical basis for the behavioral ecological model. In DiClemente R, Crosby R, & Kegler M (Eds.), Emerging theories in health promotion practice and research (2nd ed., pp. 415–450). Jossey-Bass, Inc.

Katzmarzyk PT, Powell KE, Jakicic JM, Troiano RP, Piercy K, & Tennant B. (2019). Sedentary behavior and health: Update from the 2018 physical activity guidelines advisory committee. Medicine & Science in Sports & Exercise, 51(6), 1227–1241. 10.1249/MSS.0000000000001935 [PubMed: 31095080]

Koster A, Shiroma EJ, Caserotti P, Matthews CE, Chen KY, Glynn NW, & Harris TB. (2016). Comparison of sedentary estimates between activPAL and hip- and wrist-worn ActiGraph. Medicine & Science in Sports & Exercise, 48(8), 1514–1522. 10.1249/MSS.0000000000000924 [PubMed: 27031744]

Kuster RP, Grooten WJA, Baumgartner D, Blom V, Hagströmer M, & Ekblom Ö. (2020). Detecting prolonged sitting bouts with the ActiGraph GT3X. Scandinavian Journal of Medicine and Science in Sports, 30(3), 572–582. 10.1111/sms.13601 [PubMed: 31743494]

Lin LI-K. (1989). A concordance correlation coefficient to evaluate reproducibility. Biometrics, 45(1), 255. 10.2307/2532051 [PubMed: 2720055]

Lyden K, Keadle SK, Staudenmayer J, & Freedson PS. (2017). The activPAL accurately classifies activity intensity categories in healthy adults. Medicine & Science in Sports & Exercise, 49(5), 1022–1028. 10.1249/MSS.0000000000001177 [PubMed: 28410327]

Lyden K, Kozey-Keadle SL, Staudenmayer JW, & Freedson PS. (2012). Validity of two wearable monitors to estimate breaks from sedentary time. Medicine & Science in Sports & Exercise, 44(11), 2243–2252. 10.1249/MSS.0b013e318260c477 [PubMed: 22648343]

Migueles JH, Cadenas-Sanchez C, Ekelund U, Nyström CD, MoraGonzalez J, Löf M, … Ortega FB. (2017). Accelerometer data collection and processing criteria to assess physical activity and other outcomes: A systematic review and practical considerations. Sports Medicine, 47(9), 1821–1845. 10.1007/s40279-017-0716-0 [PubMed: 28303543]

Nakandala S, Jankowska MM, Tuz-Zahra F, Bellettiere J, Carlson JA, LaCroix AZ, … Natarajan L. (2021). Application of convolutional neural network algorithms for advancing sedentary and activity bout classification. Journal for the Measurement of Physical Behaviour, 4(2), 102–110. 10.1123/jmpb.2020-0016 [PubMed: 34458688]

Owen N, Healy GN, Dempsey PC, Salmon J, Timperio A, Clark BK, … Dunstan DW. (2020). Sedentary behavior and public health: Integrating the evidence and identifying potential solutions. Annual Review of Public Health, 41(1), 265–287. 10.1146/annurev-publhealth-040119-094201

Rosenberg D, Walker R, Greenwood-Hickman MA, Bellettiere J, Xiang Y, Richmire KR, … Lacroix AZ. (2020). Device-assessed physical activity and sedentary behavior in a community-based cohort of older adults. BMC Public Health, 20(1), 1–13. 10.1186/s12889-020-09330-z [PubMed: 31898494]

Saunders TJ, Atkinson HF, Burr J, MacEwen B, Skeaff CM, & Peddie MC. (2018). The acute metabolic and vascular impact of interrupting prolonged sitting: A systematic review and meta-analysis. Sports Medicine, 48, 2347–2366. 10.1007/s40279-018-0963-8 [PubMed: 30078066]

Saunders TJ, McIsaac T, Douillette K, Gaulton N, Hunter S, Rhodes RE, … Healy GN. (2020). Sedentary behaviour and health in adults: An overview of systematic reviews. Applied Physiology, Nutrition, and Metabolism, 45(10, Suppl. 2), S197–S217. 10.1139/APNM-2020-0272

Stamatakis E, Koster A, Hamer M,Rangul V,Lee IM, Bauman AE, … Holtermann A. (2020). Emerging collaborative research platforms for the next generation of physical activity, sleep and exercise medicine guidelines: The prospective physical activity, sitting, and sleep consortium (ProPASS). British Journal of Sports Medicine, 54, 435–437. 10.1136/bjsports-2019-100786 [PubMed: 31076396]

Steel C, Bejarano C, & Carlson JA. (2019). Time drift considerations when using GPS and accelerometers. Journal for the Measurement of Physical Behaviour, 2(3), 203–207. 10.1123/jmpb.2019-0004

Tanamas SK, Magliano DJ, Lynch B, Sethi P, Willenberg L, Polkinghorne KR, Chadban S, Dunstan D, & Shaw JE. (2013). AUSDIAB 2012: The Australian diabetes, obesity and lifestyle study. https://apo.org.au/node/35278

Tremblay MS, Aubert S, Barnes JD, Saunders TJ, Carson V, Latimer-Cheung AE, … SBRN Terminology Consensus Project Participants. (2017). Sedentary behavior research network (SBRN)— Terminology consensus project process and outcome. International Journal of Behavioral Nutrition and Physical Activity, 14(75), 1–17. 10.1186/s12966-017-0525-8 [PubMed: 28057008]

Wijndaele K, Westgate K, Stephens SK, Blair SN, Bull FC, Chastin SFM, … Healy GN. (2015). Utilization and harmonization of adult accelerometry data: Review and expert consensus. Medicine & Science in Sports & Exercise, 47(10), 2129. 10.1249/MSS.0000000000000661 [PubMed: 25785929]

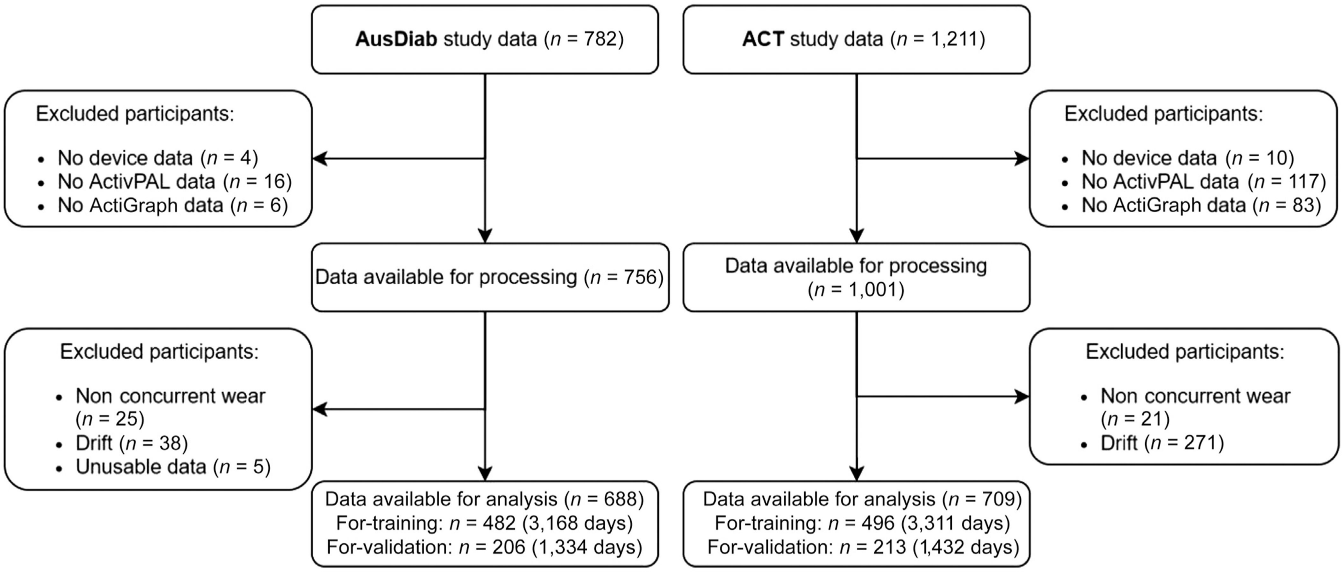

Figure 1 —.

STROBE diagram for data used to train and validate the CHAP-Adult algorithm. ACT = Adult Changes in Thought study; AusDiab = Australian Diabetes, Obesity, and Lifestyle study; CHAP = **C**onvolutional Neural Network **H**ip **A**ccelerometer **P**osture.

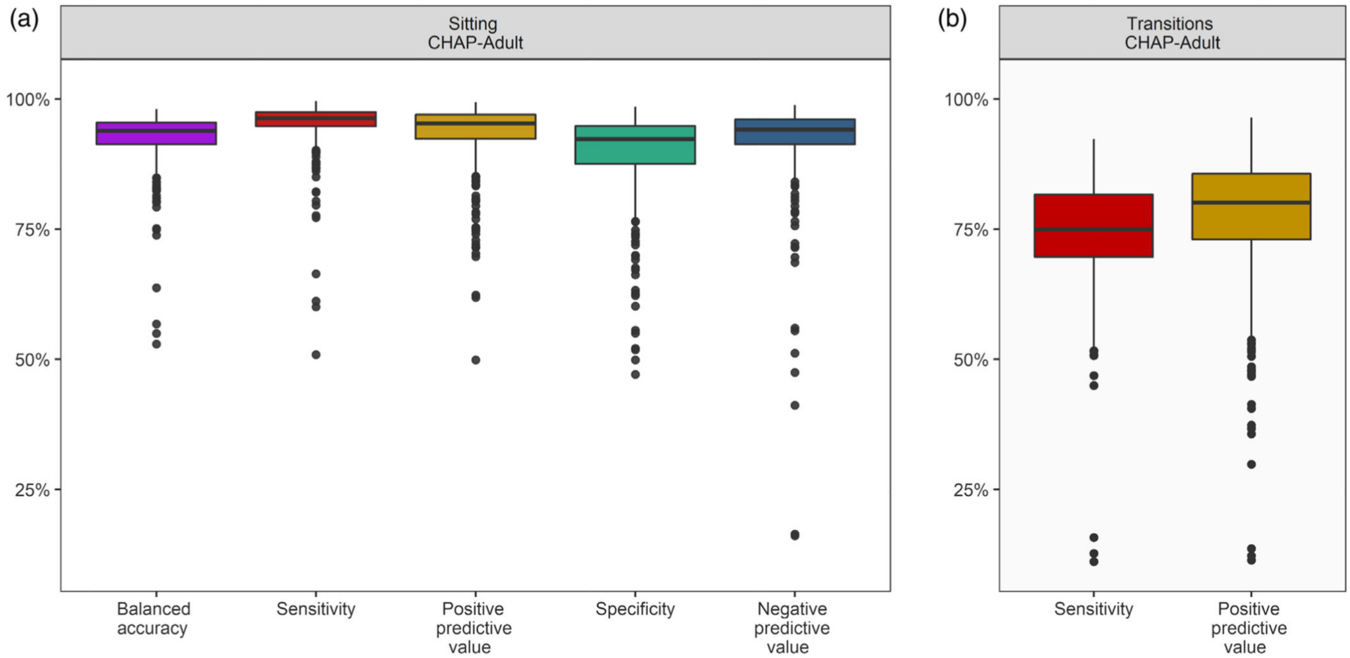

Figure 2 —.

Accuracy metrics for sitting (a) and postural transitions (b) for CHAP-Adult using the ACT and AusDiab combined for-validation sample (*n* = 419). ACT = Adult Changes in Thought study; AusDiab = Australian Diabetes, Obesity, and Lifestyle study; CHAP = **C**onvolutional Neural Network **H**ip **A**ccelerometer **P**osture.

Table 1

Participant Characteristics of the AusDiab and ACT Samples, and the Combined AusDiab and ACT Samples Used to Train and Test (For-Training) and to Validate (For-Validation) CHAP-Adult

	ACT	AusDiab[a]	For-training	For-validation
n	709	693	981	421
Age, mean (SD)	76.70 (6.52)	58.31 (10.43)	67.70 (12.67)	67.40 (12.60)
Male, *n* (%)	294 (41.5)	303 (43.7)	416 (42.4)	181 (43.0)
BMI ≥30 kg/m^2, *n* (%)	157 (22.6)	172 (24.8)	241 (24.9)	88 (21.1)
Education—Secondary school[b] or less, *n* (%)	62 (8.7)	204 (29.6)	191 (19.5)	75 (17.9)

Note. ACT = Adult Changes in Thought study; AusDiab = Australian Diabetes, Obesity, and Lifestyle study; BMI = body mass index; CHAP = **C**onvolutional Neural Network **H**ip **A**ccelerometer **P**osture.

[a]Data from five participants who were not included in the training process due to unusable accelerometer data.

[b]Secondary school includes high school.

Table 2

Validity and Reliability of Sitting and Sitting Pattern Metrics Derived From Data Processed Using CHAP-Adult Among 419 Adults From the Combined ACT and AusDiab For-Validation Sample

	Total sitting time (min/day)	Mean bout duration (min)	Number of breaks in sitting time	Usual bout duration (min)	Alpha	Time spent in bouts ≥30 min (min/day)
AP, mean (*SD*)	532.7 (118.7)	12.7 (5.9)	46.7 (14)	31.8 (15.5)	1.32 (0.04)	278.9 (127.8)
CHAP, mean (*SD*)	543.2 (116.4)	13.3 (5.9)	44.9 (12.7)	33.2 (17.1)	1.32 (0.04)	294 (133.9)
Person-level agreement						
Mean error[a] (PE)	−10.5 (−2.0%)	−0.6 (−4.7%)	1.9 (4.1%)	−1.4 (−4.4%)	0.0001 (0.0%)	−15.1 (−5.4%)
95% LoA	−63, 42	−4, 2.7	−9.2, 12.9	−8.3, 5.4	−0.04, 0.04	−84.3, 54.1
Mean absolute error (MAPE)	21.1 (4.0%)	1.5 (12.2%)	5.4 (11.6%)	3.0 (9.6%)	0.02 (1.4%)	26.7 (9.6%)
Pearson correlation	.96 (.95, .97)	.92 (.91, .94)	.86 (.83, .88)	.92 (.91, .94)	.78 (.74, .81)	.96 (.96, .97)
Concordance correlation	.96 (.95, .96)	.92 (.9, .93)	.85 (.82, .87)	.91 (.9, .93)	.78 (.74, .82)	.96 (.95, .96)

Note. ACT = Adult Changes in Thought study; AP = activPAL; AusDiab = Australian Diabetes, Obesity, and Lifestyle study; CHAP = **C**onvolutional Neural Network **H**ip **A**ccelerometer **P**osture; LoA = limit of agreement; MAPE = mean absolute percent error; PE = percent error which was computed as the mean error/AP.

[a] Mean error = AP – CHAP-Adult.

Table 3

Agreement Metrics for Total Sedentary Time, Breaks in Sitting, and Mean Bout Duration Derived From Data Processed Using CHAP-Adult Among 419 Adults From the Combined ACT and AusDiab For-Validation Sample by Age, Gender, and BMI

	AP mean (*SD*)	Mean error[a] (95% LoA)	Percent error[b]	Mean absolute error (MAPE)	Pearson correlation	Concordance correlation
Total sitting time (min/day)						
Age						
35–49	451.4 (113.6)	4.2 (−44.4, 52.8)	0.9%	18.7 (3.5%)	.98 (.95, .99)	.98 (.95, .99)
50–64	491.6 (109.2)	−3.2 (−63.2, 56.8)	−0.6%	20.6 (3.9%)	.96 (.94, .97)	.96 (.94, .97)
65+	562.0 (113.1)	−15.8 (−68.3, 36.8)	−2.7%	21.7 (4.1%)	.95 (.94, .96)	.94 (.93, .95)
Gender						
Women	518.6 (121.4)	−13.2 (−68.7, 42.3)	−2.5%	22.9 (4.3%)	.96 (.95, .97)	.95 (.94, .96)
Men	551.4 (112.7)	−6.9 (−54.7, 40.8)	−1.2%	18.7 (3.5%)	.96 (.95, .97)	.96 (.94, .97)
BMI (kg/m^2)						
<30	521.8 (113.5)	−9.4 (−58.7, 39.8)	−1.8%	19.7 (3.7%)	.96 (.95, .97)	.96 (.95, .97)
≥30	568.4 (128.4)	−15.5 (−91.5, 60.5)	−2.7%	26.6 (5.0%)	.95 (.92, .97)	.94 (.91, .96)
Breaks in sitting (no. per day)						
Age35–49	51.2 (10.4)	1.8 (−10.2, 13.8)	3.7%	5.0 (10.6%)	.83 (.69, .91)	.82 (.67, .90)
50–64	52.7 (14.1)	2.5 (−13.2, 18.2)	4.9%	6.5 (13.8%)	.82 (.75, .87)	.81 (.74, .86)
65+	43.5 (13.3)	1.6 (−9.5, 12.8)	3.9%	5.0 (10.7%)	.87 (.83, .89)	.85 (.81, .88)
Gender						
Women	46.9 (12.8)	1.6 (−10.0, 13.1)	3.4%	5.4 (11.6%)	.84 (.79, .87)	.83 (.78, .86)
Men	46.6 (15.5)	2.4 (−9.3, 14.0)	5.3%	5.4 (11.5%)	.89 (.86, .92)	.87 (.83, .90)
BMI (kg/m^2)						
<30	48.3 (14.1)	2.6 (−7.8, 13.0)	5.7%	5.3 (11.4%)	.88 (.85, .90)	.86 (.83, .88)
≥30	41.2 (12.2)	−0.6 (−15.1, 14.0)	−1.4%	5.5 (11.8%)	.79 (.70, .86)	.79 (.70, .86)
Mean bout duration (min)						
Age						
35–49	9.1 (2.9)	−0.3 (−2.8, 2.1)	−3.6%	1.0 (7.5%)	.92 (.84, .96)	.91 (.83, .95)
50–64	9.9 (3.3)	−0.6 (−3.7, 2.4)	−5.9%	1.2 (9.7%)	.90 (.86, .93)	.88 (.83, .92)
65+	14.4 (6.4)	−0.7 (−4.4, 3.0)	−4.6%	1.8 (13.9%)	.91 (.89, .93)	.90 (.88, .92)
Gender						

	AP mean (*SD*)	Mean error[a] (95% LoA)	Percent error[b]	Mean absolute error (MAPE)	Pearson correlation	Concordance correlation
Women	11.9 (4.8)	−0.7 (−4.0, 2.6)	−5.2%	1.5 (11.7%)	.89 (.87, .92)	.88 (.85, .91)
Men	13.6 (7.0)	−0.6 (−4.2, 3.0)	−4.3%	1.6 (12.8%)	.94 (.92, .95)	.93 (.91, .95)
BMI (kg/m^2)						
<30	11.9 (5.3)	−0.8 (−3.7, 2.0)	−6.6%	1.4 (10.9%)	.95 (.94, .96)	.93 (.92, .95)
≥30	15.3 (7.2)	0.1 (−5.1, 5.2)	0.3%	2.1 (16.7%)	.88 (.82, .92)	.86 (.80, .91)

Note. Sample size per category—Age: 35–49 = 36, 50–64 = 118, and ≥65 = 265; gender: women = 239 and men = 180; BMI: <30 kg/m^2 = 328 and ≥30 kg/m^2 = 88. ACT = Adult Changes in Thought study; AP = activPAL; AusDiab = Australian Diabetes, Obesity, and Lifestyle study; BMI = body mass index; CHAP = **C**onvolutional Neural Network **H**ip **A**ccelerometer **P**osture; LOA = limit of agreement; MAPE = mean absolute percent error.

[a]Mean error = AP – CHAP-Adult.

[b]Percent error was computed as the mean error divided by the AP mean. This is included to contextualize the magnitude of the observed error.