Montoye, A. H.K., Vondrasek, J. D., Neph, S. E., Basu, N. , Paul, L., Bachmair, E.-M., Stefanov, K. and Gray, S. R. (2022) Comparison of the activPAL CREA and VANE algorithms for characterization of posture and activity in free-living adults. *Journal for the Measurement of Physical Behaviour*, 5(1), pp. 49-57. (doi: 10.1123/jmpb.2021-0053)

© 2022 Human Kinetics

This is the author's version of the work. It is posted here for your personal use. You are advised to consult the published version if you wish to cite from it: DOI: https://doi.org/10.1123/jmpb.2021-0053

http://eprints.gla.ac.uk/263462/

Enlighten – Research publications by members of the University of Glasgow
http://eprints.gla.ac.uk

Comparison of the activPAL CREA and VANE algorithms for characterization of posture and activity in free-living adults

ABSTRACT

Background: The activPAL accelerometer is used widely for assessment of free-living activity and postural data. Two algorithms, VANE (traditional) and CREA (new), are available to analyze activPAL data, but the comparability of metrics derived from these algorithms is unknown. **Purpose:** To determine the comparability of physical activity and sedentary behaviour metrics from activPAL's VANE and CREA algorithms. **Methods:** Individuals enrolled in the LIFT trial (n=354) wore an activPAL accelerometer on the right thigh continuously for seven days on four occasions, resulting in 5,851 valid days of data for analysis. Daily data were downloaded in the PALbatch software using the VANE and CREA algorithms. Correlations, mean absolute percent error, effect sizes, and equivalence (within 3%) were calculated to evaluate comparability of the algorithms. **Results:** Steps, activity score, stepping time, bouts of stepping, and upright time metrics were statistically equivalent, highly correlated ($r \geq 0.98$), and had small mean absolute percent errors (≤2.5%) and trivial effect sizes (ES<0.07) between algorithms. Stepping bouts also had good comparability. Conversely, sedentary-upright and upright-sedentary transitions and bouts of sitting were not equivalent, with large mean absolute percent differences (17.4-141.3%) and small to very large effect sizes (ES=0.45-3.80) between algorithms. **Conclusions:** Stepping and upright metrics are highly comparable between activPAL's VANE and CREA algorithms, but sitting metrics had large differences as the VANE algorithm does not capture non-wear or differentiate between sitting and lying down. Researchers using the activPAL should explicitly describe the analytic algorithms used in their work to facilitate data pooling and comparability across studies.

Keywords: equivalence, physical activity, sedentary behaviour, accelerometer, activity monitor, MVPA

INTRODUCTION

Over the last four decades, accelerometer-based physical activity monitors have evolved from bulky but simple, step- or Calorie-estimating devices to miniaturized, sophisticated devices capable of capturing rich acceleration data from which advanced activity and postural information can be derived (Intille, Lester, Sallis, & Duncan, 2012; Wong, Webster, Montoye, & Washburn, 1981). Today, accelerometers can be placed on a variety of body locations, depending on the type of data to be collected and to balance the desired accuracy and wear compliance (Troiano, McClain, Brychta, & Chen, 2014). The thigh placement is appealing in that it allows for the capture and differentiation of seated and upright postures as well as between stationary and non-stationary activity (Steeves et al., 2015), thereby allowing for differentiation between sedentary, light, and moderate- to vigorous-intensity activities based on a combination of body position and acceleration (Edwardson et al., 2017). The placement on the leg also allows for accurate step counting across a variety of activity types and movement speeds (except for very slow walking) (Grant, Ryan, Tigbe, & Granat, 2006; Stansfield, Hajarnis, & Sudarshan, 2015) and may have good comparability across device brands (Crowley et al., 2019). Additionally, as the device is typically taped in place on the thigh and designed to be worn continuously, wear compliance on the thigh has been shown to be high in large, cohort-based studies (Edwardson et al., 2017; Hamer et al., 2020; Stevens et al., 2020).

One such commonly-used device, the activPAL (PAL Technologies, Ltd., Glasgow, UK), has been developed specifically for wear on the thigh and has been validated for assessment of steps and posture in laboratory and free-living settings (Grant et al., 2006; Kozey-Keadle, Libertine, Lyden, Staudenmayer, & Freedson, 2011; Lyden, Kozey Keadle, Staudenmayer, & Freedson, 2012; Steeves et al., 2015). The traditional algorithm, called "VANE" in the activPAL software, allows for derivation of stepping, limited postures (seated, upright, and stepping), and transitions from sedentary to upright and upright to sedentary postures. In December 2018, a beta version of a new, enhanced "CREA" algorithm was added to the activPAL software, and the full release of the CREA algorithm occurred in June 2019 (version 8.11.1.47; https://kb.palt.com/knowledge-base/palanalysis-version-history/). This new algorithm provided the original metrics included in the VANE algorithm but added determination of non-wear; splitting of stepping into cycling and non-cycling stepping; and parsing of sitting into sitting, seated transport, and both primary (nighttime) and secondary (daytime) lying down. This new algorithm advances the ability to understand the health implications of posture, activity, and sleep. However, prior to widespread use of the CREA algorithm, clarity is needed in several areas. First, metrics which are new to the CREA algorithm should be validated; preliminary work has been done for some of the new activPAL variables including non-wear (Carlson et al., 2021; Winkler et al., 2016) and primary lying down (Courtney et al., 2021; Hettiarachchi et al., 2021; Lyden, John, Dall, & Granat, 2016; van der Berg et al., 2016). Second, the CREA should be compared against the traditional VANE algorithm to determine if and how to compare the results across studies using different algorithms. For this latter question, we are unaware of any past research that has compared outputs from activPAL's VANE and CREA algorithms. While it would seem that the CREA simply decomposes some of VANE's metrics into more granular postures (e.g., CREA stepping + CREA cycling = VANE stepping), preliminary analysis by our group revealed that outputs from CREA could not be used to perfectly reconstruct the VANE variables. Therefore, our study's purpose was to perform a thorough analysis comparing outputs from the VANE and CREA algorithms in free-living adults and determine which variables are equivalent.

METHODS

Participants

Data for this study came from a UK-based clinical trial; individuals with inflammatory rheumatic disease (n=368) were recruited for participation in the LIFT randomized clinical trial (trial registration number: NCT03248518), and the main goal of the trial is to lessen fatigue in this population. Details of the LIFT trial inclusion criteria are described in detail elsewhere (Martin et al., 2019). Individuals who completed one week of accelerometer wear (n=354) were eligible for inclusion in this study. Participants were primarily female (~75%) and had rheumatoid arthritis (~56%), with an average age of 57.9±12.5 years and average duration of living with their rheumatic condition of 11.5±10.0 years. All participants provided informed consent to participate in the study, and study methods were approved by the Research Ethics Committee Wales REC 7 (17/WA/0065).

Protocol

Participants reported to a study site or research facility for four assessment visits over the course of a year, with visits 1, 2 and 3 spaced roughly three months apart and visit 4 roughly six months after visit 3 (one year following visit 1). At each visit, participants were fitted with an activPAL3 micro (PAL Technologies Ltd., Glasgow, UK) accelerometer. The accelerometer was affixed to the midline of the anterior portion of the right leg using Tegaderm (3M, Maplewood, MN, USA) hypoallergic tape. The Tegaderm allowed the accelerometer to be worn continuously. The accelerometer was set to begin recording the day after participants were fitted with the device and recorded triaxial data at 20 Hz for seven days. Participants were instructed to wear the accelerometer at all times until the seven measurement days were completed, at which time participants mailed it back to research staff in a pre-paid envelope.

Upon receipt of the accelerometer, data were downloaded from the "Daily summaries" export selection option using the PALbatch software (version 8.11.5.64; PAL Technologies Ltd., Glasgow, UK). First, summary metrics for each valid day were downloaded using activPAL's VANE algorithm, which the PALbatch software describes as "The standard PAL analysis algorithm." Then, using the same software defaults, summary metrics for each valid day were downloaded using activPAL's CREA algorithm, which the PALbatch software describes as the "Enhanced analysis algorithm." For both algorithms, the minimum sedentary and upright postures for determination of transitions was set at the default of 10 seconds. The VANE algorithm outputs steps, activity score (MET-hours), time spent in sitting (all seated and lying body positions plus non-wear)/upright/stepping postures, and sedentary-upright and upright-sedentary transitions. The CREA algorithm captures the same metrics but parses VANE's sitting time metric into non-wear, sitting, seated transport (which is also counted as part of the sitting time metric), and both primary (overnight) and secondary (during the daytime) lying down. With the CREA algorithm, sitting and secondary lying would be considered "sedentary time", while primary lying would be used to derive sleep estimates. Additionally, stepping time is divided into cycling time and non-cycling stepping time. Postures are determined by a combination of accelerometer orientation and movement data, and METs are determined by a combination of posture and step counts (Edwardson et al., 2017). Transitions (both sedentary-upright and upright-sedentary) are calculated in the software as times where the wearer's body position shifts from "non-upright" (e.g., sitting, lying) to "upright" (e.g., standing, stepping). VANE and CREA data were aligned by participant and day for analysis.

Statistical analysis

All valid days (>20 hours of wear time, no data errors noted by the software, and 1440 minutes of recording time) were analyzed (rather than averages across a participant) across all visits; participants could therefore contribute up to 28 data points (7 days * 4 visits) for analysis. Total daily steps, activity score (MET-hours), sitting time (both unadjusted [as reported by software] and "adjusted" by summing composite sitting + primary lying + secondary lying + non-wear), stepping time, standing time, upright time, the number of sedentary-upright and upright-sedentary were compared between VANE and CREA algorithms using Pearson correlations, mean absolute differences, and mean absolute percent differences, with a mean absolute percent difference of <3.0% indicating excellent agreement (Holbrook, Barreira, & Kang, 2009). Additionally, equivalence testing was conducted. In equivalence testing, an equivalence region must be defined; we used a 3% limit to indicate equivalence between algorithms as this is a strict threshold which would require a high level of agreement between algorithms (Dixon et al., 2018). We implemented equivalence testing using two one-sided tests of equivalence using the TOSTER package (version 0.3.4) in R). Finally, effect sizes were calculated for differences between algorithms and were interpreted as ES<0.20 = trivial, 0.20≤ES<0.50 = small, 0.50≤ES<0.80 = Medium, 0.80≤ES<1.30 = Large, and ES>1.30 = very large (Cohen, 1988).

These same statistical tests were conducted for comparing several secondary metrics reported by the software for sitting bouts (number of bouts >30 minutes, number of bouts >60 minutes, time spent in bouts >30 minutes, time spent in bouts >60 minutes), stepping bouts

(bouts <1 minute, bouts 1-5 minutes, bouts 5-10 minutes, bouts 10-20 minutes, and bouts >20 minutes, along with number of steps taken in each of these bout durations), and stepping cadences (time at a cadence ≥75 steps/minute, time at a cadence ≥75 steps/minute for a duration >1 minute, number of steps taken at a cadence ≥75 steps/minute, number of steps taken at a cadence ≥75 steps/minute for a duration >1 minute, time at a cadence ≥100 steps/minute, time at a cadence ≥100 steps/minute for a duration >1 minute, number of steps taken at a cadence ≥100 steps/minute, number of steps taken at a cadence ≥100 steps/minute for a duration >1 minute).

To evaluate if differences between algorithms was a function of activity level or type of activity participation, Bland-Altman plots (Bland & Altman, 1986) were constructed comparing the average of CREA and VANE algorithms against the difference between the CREA and VANE algorithms for each variable. Finally, because steps, METs, and postural data were not identical between algorithms (see Results), we did some follow-up in subjects whose data were outliers for each variable, defined as difference between algorithms that fell more than 3.0 standard deviations away from the mean. In these files, we looked at daily totals as well as 60-second epochs to determine the cause of disagreement. Additionally, the CREA algorithm reports accelerometer alignment and data errors in separate columns, and these were evaluated when difference between algorithms occurred. All analyses were conducted in Microsoft Excel 2016 (Microsoft Corp., Redmond, WA, USA) and RStudio (version 1.1.423; R Core Development Team, Vienna, Austria).

RESULTS

Across the four possible one-week visits, participants (n=354) provided a total of 907 visits of data and a total of 5,851 valid days, at an average of 6.5 valid days per visit. Data from primary outcome metrics are shown in Table 1. Correlations were high (≥0.97) for all metrics except unadjusted sitting time. Additionally, mean absolute differences in time spent in each posture were <8 minutes (and mean absolute percent differences <4%) for all measures except unadjusted sitting time, with the disagreement between sedentary-upright transitions being ~7-8 per day (~17% difference). Effect sizes were trivial for all except unadjusted sitting time (very large effect size) and both sedentary-upright and upright-sedentary transitions (small effect sizes).

** INSERT TABLE 1 AND TABLE 2 HERE**

Using the 3% equivalence threshold (primary outcomes shown in Figure 1), all primary outcomes were equivalent (signified by having a mean difference and its 95% confidence interval falling entirely within 3% of the mean for the metric) except for unadjusted sitting, standing, and both the sedentary-upright and upright-sedentary transitions. Data from secondary outcome metrics can be found in Table 2, with equivalence graphs shown in Supplemental Data 1. For most metrics, there were significant differences and a lack of equivalence between algorithms. However, effect sizes were trivial, mean absolute differences were <3%, and correlations were high (≥0.96) for all metrics except the sitting bout metrics.

** INSERT FIGURE 1 HERE**

Bland-Altman plots (primary outcomes shown in Figure 2, secondary outcomes shown in Supplemental Data 2) show narrow limits of agreement for total steps, MET-hours, standing time, upright time, and stepping time as well as most of the secondary metrics other than sitting bouts.

** INSERT FIGURE 2 HERE**

Outliers were rare for most variables. Total daily steps were <200 steps different between algorithms for all data. For MET-hours, upright time, standing time, stepping time, and adjusted sitting time, two observations were made. First, the outliers seemed to cluster in files with poorer alignment of the accelerometer on the thigh. For example, the average thigh alignment reported by the software for the outlier files was lower (93.9%) than non-outlier files (96.7%). Second, the outliers in standing and adjusted sitting time occurred more frequently when there were outliers in sedentary-upright and upright-sedentary transitions. For the sedentary-upright and upright-sedentary transitions, the VANE algorithm always had at least as many as the CREA algorithm, with the difference between algorithms increasing as the total number of transitions increased. These occurred most frequently when the VANE algorithm noted a change from sedentary to upright, non-stepping posture (rather than transitions from sedentary to stepping behaviors), whereas the CREA did not note such change in postures. The same was true for upright-sedentary transitions. In the transition outlier files, the disagreement in transitions was most pronounced in the waking hours, with 79% of disagreements happening between the hours of 9am-8pm.

Finally, for unadjusted sitting and all secondary sitting bout metrics, the largest differences between metrics occurred in files with the most lying down (primary or secondary) detected by the CREA algorithm.

DISCUSSION

This study's purpose was to evaluate the comparability of metrics derived from the traditional (VANE) and new (CREA) algorithms available within activPAL's data analysis software. None of the metrics were identical between algorithms (i.e., having point estimates that matched perfectly), signaling that there is a difference in how the VANE and CREA algorithms calculate steps and postural metrics, and therefore also the transitions between sedentary and upright postures. However, for many of the primary metrics (total steps, MET-hours, upright and stepping time, adjusted sitting time), the data were statistically equivalent to within 3% and had mean absolute percent differences <3%, both of which are rigorous criteria demonstrating excellent comparability between algorithms. The only posture not statistically equivalent, standing time, narrowly missed equivalence at 3% (upper bound 7.4 minutes, where 7.1 minutes would have been equivalent) but would have been equivalent at a 4% threshold (where an upper bound under 9.4 minutes is considered equivalent)). Additionally, the high correlations, small absolute and absolute percent differences, and trivial effect sizes further demonstrate excellent comparability between algorithms for most primary metrics. Fewer of the secondary metrics, which focus mostly on detecting bouts of sitting and stepping, were statistically equivalent due to occasional outliers resulting in wide standard deviations in differences between algorithms. Nevertheless, for all but the bouts of sitting metrics, mean absolute percent differences were <3% and therefore indicate good agreement.

However, metrics that included sitting behaviours demonstrated poor comparability. This is not surprising, as the VANE algorithm groups sitting, lying down, and non-wear into a single "sitting" metric, while the CREA assesses them as separate constructs. Encouragingly, a metric equivalent to VANE's sitting can be computed by summing the CREA's sitting, non-wear, and primary and secondary lying down metrics, and we called this "adjusted sitting" in our analysis. Seated transport was not included into adjusted sitting as transport is already included in CREA's sitting time.

The other primary metrics which showed poor comparability were transitions between sedentary and upright postures, and vice versa. The VANE algorithm had more sedentary-upright transitions than the CREA algorithm in 99.9% (5,844 out of 5,851) of files and more upright-sedentary transitions in 99.6% (5,826 out of 5,851) of files, and in no files did the CREA algorithm have more transitions than the VANE. In examining second-, minute-, and hour-level data from outlier files, disagreements in postural transitions occurred almost exclusively where the CREA algorithm classified the time as mostly/all sitting (not lying down, non-wear, or any other posture) but the VANE algorithm recorded a mix of sitting and upright postures (resulting in postural transitions), but no stepping. ~~As noted in the Results, these disagreements occurred primarily during waking hours.~~ In the file with the greatest disagreement in total transitions, 1.2-1.3% of 1-minute epochs had different numbers of sedentary-upright or upright-sedentary transitions. As noted in the Results, much of the discrepancy in transitions between algorithms occurred during waking hours, and very few disagreements occurred when the CREA algorithm indicated primary lying activity. Therefore, it is unlikely that awakenings during the night to use the toilet, tend to a child, or do other nighttime activities would be classified differently or lead to major dissimilarities between algorithms. It is unclear why the algorithms disagree on such transitions, as sitting and upright postures are distinctly recognizable through changes in thigh angle, as has been demonstrated in past work (Steeves et al., 2015). However,~~, but~~ ~~it~~ our results do provide evidence~~does show~~ that the algorithms have a slight difference in the way they detect sitting vs. standing, resulting in~~ under certain circumstances the algorithms~~ some disagreement when differentiating between~~identify~~ seated and standing postures under certain circumstances~~postures differently~~. ~~As noted in the Results, these disagreements occurred primarily during waking hours.~~

Notably, the disagreement in sitting time as well as transitions to and from sedentary postures led to large differences in some of the secondary metrics, mainly the number of prolonged sedentary bouts and the time spent in prolonged sitting (>30 minutes and >60 minutes). And unlike our adjusted sitting metric, there is not an easy way to recalculate these metrics to make them comparable between the VANE and CREA algorithms. Prolonged sedentary behaviour appears to have negative health effects, with the breaking up of sedentary bouts being important for reducing some of their negative health implications (Dunstan et al., 2012; Loh, Stamatakis, Folkerts, Allgrove, & Moir, 2020). Therefore, accurate measures of transitions from sedentary to non-sedentary behaviours are key to better understanding health implications of sedentary behaviour patterns. It is not possible from this study to determine which algorithm is more accurate for the transition metrics, so future work should examine which one more accurately assesses breaks in sedentary behaviours.

In our literature search, few papers using the activPAL mentioned which algorithm they used for analyzing their data. For studies that analyzed activPAL data prior to June 2019, there was less need to report such information since there were no algorithm choices to be made. However, now that two algorithms exist for analyzing the data and produce different metrics (some of which we found to be non-equivalent), it is important that studies report which algorithm they choose when analyzing data. In some cases, it may be possible to infer which algorithm was used based on the variables reported (e.g., if a study reports time lying down from the activPAL, it is likely the CREA algorithm). However, some researchers employ methods not available in the activPAL analysis software for deriving similar metrics. For example, past research has developed several methods available for detecting activPAL non-wear and wake/sleep time (Bassett et al., 2014; Hettiarachchi et al., 2021; van der Berg et al., 2016; Winkler et al., 2016) without using activPAL's CREA algorithm. Given the push for pooling of thigh-worn data by groups such as the ProPASS consortium (Crowley et al., 2019; Stevens et al., 2020), it is critical for researchers using the activPAL to be explicit in the protocols they employ when analyzing activPAL data. Especially as ProPASS identified that many researchers want to maintain control of their own data analyses (rather than centralized processing) (Stevens et al., 2020), data download and analysis methods must be explicit for replication and comparability purposes.

Study strengths and limitations

Our study had several notable strengths. Our sample size was large and diverse in age and physical ability. The wide range in daily step counts in this study (range 82-34,490; mean 7,172; median 6,308) is comparable to other adult populations (Bassett, Wyatt, Thompson, Peters, & Hill, 2010; Tudor-Locke et al., 2011), providing support of the generalizability of the data. Further, downloading data from the same device rather than collecting data in monitors worn side-by-side eliminated any possible issues with device placement or inter-device reliability.

Our study also had notable limitations. The population, while large and diverse, comprised adults with some form of diagnosed rheumatic condition (ranging from minimal to severe symptoms). Thus, these results should be confirmed in apparently healthy adults as well as in other populations such as children, who have very different movement and behavioural patterns. Additionally, due to the free-living nature of this study, we did not have a criterion measure for determination of which algorithm was most accurate for capturing the metrics deemed not equivalent between algorithms. Finally, agreement between algorithms would likely have been higher (more favorable) if primary lying would have been removed from the dataset (i.e., prior to processing the VANE variables), as is often done using self-reported bed times in past work (Lynch et al., 2019). However, as our results show, the improvements in agreement would likely have been modest as most of the disagreement between algorithms occurred during the daytime hours.

Conclusions

Our study found equivalence to a strict 3% threshold for steps, stepping time, upright time, MET-hours, and adjusted sitting time as well as many bouted stepping metrics when comparing outputs from activPAL's VANE and CREA data analysis algorithms. However, transitions between sedentary and upright postures as well as bouted and unadjusted sitting were not equivalent, and had large differences, between algorithms. Given our findings, it is possible to compare some of the activPAL's metrics across studies using different analysis algorithms. However, when using the activPAL, researchers should be explicit about what algorithms/methods they use to analyze data. Only with clear reporting of analysis choices will pooling and comparability of activPAL data from studies across different research groups and over time be fully possible.

REFERENCES

Bassett, D. R., Jr., John, D., Conger, S. A., Rider, B. C., Passmore, R. M., & Clark, J. M. (2014). Detection of lying down, sitting, standing, and stepping using two activPAL monitors. *Med Sci Sports Exerc, 46*(10), 2025-2029. doi:10.1249/MSS.0000000000000326

Bassett, D. R., Jr., Wyatt, H. R., Thompson, H., Peters, J. C., & Hill, J. O. (2010). Pedometer-measured physical activity and health behaviors in U.S. adults. *Med Sci Sports Exerc, 42*(10), 1819-1825. doi:10.1249/MSS.0b013e3181dc2e54

Bland, J. M., & Altman, D. G. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet, 1*(8476), 307-310. doi:10.1016/S0140-6736(86)90837-8

Carlson, J. A., Tuz-Zahra, F., Bellettiere, J., Ridgers, N. D., Steel, C., Bejarano, C., . . . Natarajan, L. (2021). Validity of Two Awake Wear-Time Classification Algorithms for activPAL in Youth, Adults, and Older Adults. *J Meas Phys Behav, 4*(2), 151-162. doi:10.1123/jmpb.2020-0045

Cohen, S. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2 ed.). Hillsdale (NJ): Lawrence Erlbaum Associates.

Courtney, J. B., Nuss, K., Lyden, K., Harrall, K. K., Glueck, D. H., Villalobos, A., . . . Litt, J. S. (2021). Comparing the activPAL software's Primary Time in Bed Algorithm against Self-Report and van der Berg's Algorithm. *Meas Phys Educ Exerc Sci, 25*(3), 212-226. doi:10.1080/1091367x.2020.1867146

Crowley, P., Skotte, J., Stamatakis, E., Hamer, M., Aadahl, M., Stevens, M. L., . . . Holtermann, A. (2019). Comparison of physical behavior estimates from three different thigh-worn accelerometers brands: a proof-of-concept for the Prospective Physical Activity, Sitting, and Sleep consortium (ProPASS). *Int J Behav Nutr Phys Act, 16*(1), 65. doi:10.1186/s12966-019-0835-0

Dixon, P. M., Saint-Maurice, P. F., Kim, Y., Hibbing, P., Bai, Y., & Welk, G. J. (2018). A Primer on the use of equivalence testing for evaluating measurement agreement. *Med Sci Sports Exerc, 50*(4), 837-845. doi:10.1249/MSS.0000000000001481

Dunstan, D. W., Kingwell, B. A., Larsen, R., Healy, G. N., Cerin, E., Hamilton, M. T., . . . Owen, N. (2012). Breaking up prolonged sitting reduces postprandial glucose and insulin responses. *Diabetes Care, 35*(5), 976-983. doi:10.2337/dc11-1931

Edwardson, C. L., Winkler, E. A. H., Bodicoat, D. H., Yates, T., Davies, M. J., Dunstan, D. W., & Healy, G. N. (2017). Considerations when using the activPAL monitor in field-based research with adult populations. *J Sport Health Sci*, *6*(2), 162-178. doi:10.1016/j.jshs.2016.02.002

Grant, P. M., Ryan, C. G., Tigbe, W. W., & Granat, M. H. (2006). The validation of a novel activity monitor in the measurement of posture and motion during everyday activities. *British journal of sports medicine, 40*(12), 992-997. doi:10.1136/bjsm.2006.030262

Hamer, M., Stamatakis, E., Chastin, S., Pearson, N., Brown, M., Gilbert, E., & Sullivan, A. (2020). Feasibility of Measuring Sedentary Time Using Data From a Thigh-Worn Accelerometer. *Am J Epidemiol, 189*(9), 963-971. doi:10.1093/aje/kwaa047

Hettiarachchi, P., Aili, K., Holtermann, A., Stamatakis, E., Svartengren, M., & Palm, P. (2021). Validity of a Non-Proprietary Algorithm for Identifying Lying Down Using Raw Data from Thigh-Worn Triaxial Accelerometers. *Sensors (Basel)*, *21*(3). doi:10.3390/s21030904

Holbrook, E. A., Barreira, T. V., & Kang, M. (2009). Validity and reliability of Omron pedometers for prescribed and self-paced walking. *Med Sci Sports Exerc*, *41*(3), 670-674. doi:10.1249/MSS.0b013e3181886095

Intille, S. S., Lester, J., Sallis, J. F., & Duncan, G. (2012). New horizons in sensor development. *Medicine and science in sports and exercise, 44*(1 Suppl 1), S24-31. doi:10.1249/MSS.0b013e3182399c7d

Kozey-Keadle, S., Libertine, A., Lyden, K., Staudenmayer, J., & Freedson, P. S. (2011). Validation of wearable monitors for assessing sedentary behavior. *Medicine and science in sports and exercise, 43*(8), 1561-1567. doi:10.1249/MSS.0b013e31820ce174

Loh, R., Stamatakis, E., Folkerts, D., Allgrove, J. E., & Moir, H. J. (2020). Effects of Interrupting Prolonged Sitting with Physical Activity Breaks on Blood Glucose, Insulin and Triacylglycerol Measures: A Systematic Review and Meta-analysis. *Sports Med, 50*(2), 295-330. doi:10.1007/s40279-019-01183-w

Lyden, K., John, D., Dall, P., & Granat, M. H. (2016). Differentiating Sitting and Lying Using a Thigh-Worn Accelerometer. *Med Sci Sports Exerc, 48*(4), 742-747. doi:10.1249/MSS.0000000000000804

Lyden, K., Kozey Keadle, S. L., Staudenmayer, J. W., & Freedson, P. S. (2012). Validity of two wearable monitors to estimate breaks from sedentary time. *Medicine and science in sports and exercise, 44*(11), 2243-2252. doi:10.1249/MSS.0b013e318260c477

Lynch, B. M., Nguyen, N. H., Moore, M. M., Reeves, M. M., Rosenberg, D. E., Boyle, T., . . . English, D. R. (2019). A randomized controlled trial of a wearable technology-based intervention for increasing moderate to vigorous physical activity and reducing sedentary behavior in breast cancer survivors: The ACTIVATE Trial. *Cancer, 125*(16), 2846-2855. doi:10.1002/cncr.32143

Martin, K. R., Bachmair, E. M., Aucott, L., Dures, E., Emsley, R., Gray, S. R., . . . Basu, N. (2019). Protocol for a multicentre randomised controlled parallel-group trial to compare the effectiveness of remotely delivered cognitive-behavioural and graded exercise interventions with usual care alone to lessen the impact of fatigue in inflammatory rheumatic diseases (LIFT). *BMJ Open, 9*(1), e026793. doi:10.1136/bmjopen-2018-026793

Stansfield, B., Hajarnis, M., & Sudarshan, R. (2015). Characteristics of very slow stepping in healthy adults and validity of the activPAL3 activity monitor in detecting these steps. *Med Eng Phys, 37*(1), 42-47. doi:10.1016/j.medengphy.2014.10.003

Steeves, J. A., Bowles, H. R., McClain, J. J., Dodd, K. W., Brychta, R. J., Wang, J., & Chen, K. Y. (2015). Ability of thigh-worn ActiGraph and activPAL monitors to classify posture and motion. *Med Sci Sports Exerc, 47*(5), 952-959. doi:10.1249/MSS.0000000000000497

Stevens, M. L., Gupta, N., Inan Eroglu, E., Crowley, P. J., Eroglu, B., Bauman, A., . . . Stamatakis, E. (2020). Thigh-worn accelerometry for measuring movement and posture across the 24-hour cycle: a scoping review and expert statement. *BMJ Open Sport Exerc Med, 6*(1), e000874. doi:10.1136/bmjsem-2020-000874

Troiano, R. P., McClain, J. J., Brychta, R. J., & Chen, K. Y. (2014). Evolution of accelerometer methods for physical activity research. *British journal of sports medicine, 48*(13), 1019-1023. doi:10.1136/bjsports-2014-093546

Tudor-Locke, C., Craig, C. L., Aoyagi, Y., Bell, R. C., Croteau, K. A., De Bourdeaudhuij, I., . . . Blair, S. N. (2011). How many steps/day are enough? For older adults and special populations. *Int J Behav Nutr Phys Act, 8*, 80. doi:10.1186/1479-5868-8-80

van der Berg, J. D., Willems, P. J., van der Velde, J. H., Savelberg, H. H., Schaper, N. C., Schram, M. T., . . . Koster, A. (2016). Identifying waking time in 24-h accelerometry data in adults using an automated algorithm. *J Sports Sci, 34*(19), 1867-1873. doi:10.1080/02640414.2016.1140908

Winkler, E. A., Bodicoat, D. H., Healy, G. N., Bakrania, K., Yates, T., Owen, N., . . . Edwardson, C. L. (2016). Identifying adults' valid waking wear time by automated estimation in activPAL data collected with a 24 h wear protocol. *Physiol Meas, 37*(10), 1653-1668. doi:10.1088/0967-3334/37/10/1653

Wong, T. C., Webster, J. G., Montoye, H. J., & Washburn, R. (1981). Portable accelerometer device for measuring human energy expenditure. *IEEE transactions on bio-medical engineering*, *28*(6), 467-471. doi:10.1109/TBME.1981.324820

TABLES AND FIGURES

Table 1. Comparison of outputs from the VANE and CREA algorithms in PALbatch software for primary outcome metrics.

	VANE algorithm	CREA algorithm	Correlation	Mean absolute difference	Mean absolute percent difference	Effect size
Steps	7171.6 (4126.4)	7171.5 (4126.3)	1.00	2.4 (4.3)	0.0 (0.1)	0.00
MET-hours	33.4 (1.7)	33.4 (1.8)	0.98	0.1 (0.4)	0.2 (1.3)	0.04
Sitting time unadjusted (min)	1100.4 (139.0)	546.5 (152.6)*	0.55	554.0 (138.5)	69.3 (21.1)	3.80
Sitting time adjusted (Sitting + non-wear + lying down (min)	1100.4 (139.0)	1107.7 (137.4)	0.99	7.5 (19.4)	0.7 (1.8)	0.00
Standing) time (min)	245.4 (110.4)	237.9 (107.9)*	0.98	7.6 (20.8)	3.4 (9.5)	0.07
Upright time (min)	339.6 (139.0)	332.2 (137.4)	0.99	7.6 (20.8)	2.5 (7.6)	0.05
Stepping time (min)	94.2 (47.7)	94.2 (47.7)	1.00	0.0 (0.1)	0.1 (0.1)	0.00
Sedentary-upright transitions	50.0 (17.2)	42.5 (16.0)*	0.97	7.5 (4.5)	17.6 (10.7)	0.45
Upright-sedentary transitions	50.0 (17.2)	42.5 (16.0)*	0.97	7.5 (4.5)	17.4 (10.6)	0.45

*Indicates that VANE and CREA algorithms are not equivalent at a 3% threshold.

Table 2. Comparison of outputs from the VANE and CREA algorithms in PALbatch software for secondary outcome metrics.

	VANE algorithm	CREA algorithm	Correlation	Mean absolute difference	Mean absolute percent difference	Effect size
Number of sitting bouts >30 min	8.2 (2.6)	4.7 (2.3)*	0.76	3.4 (1.7)	60.4 (37.0)	1.45
Number of sitting bouts >60 min	4.2 (1.7)	1.5 (1.3)*	0.66	2.7 (1.3)	113.1 (60.2)	1.41
Time (min) spent in sitting bouts >30 min	843.4 (175.6)	280.0 (162.9)*	0.63	563.5 (146.0)	106.9 (35.5)	3.03
Time (min) spent in sitting bouts >60 min	674.6 (175.4)	143.2 (146.5)*	0.62	531.5 (142.8)	141.3 (45.3)	2.81
Stepping time (min) with duration ≤1 min	67.9 (30.2)	67.9 (30.2)	1.00	0.1 (0.1)	0.1 (0.2)	0.00
Stepping time (min) with duration >1 to 5 min	17.4 (17.0)	17.0 (16.5)*	0.98	0.4 (3.2)	2.0 (12.3)	0.04
Stepping time (min) with duration >5 to 10 min	4.8 (9.0)	4.7 (8.9)	0.99	0.1 (1.5)	1.8 (18.0)	0.02
Stepping time (min) with duration >10 to 20 min	2.6 (7.8)	2.5 (7.5)*	0.97	0.1 (1.9)	1.4 (16.0)	0.04
Stepping time (min) with duration >20 min	1.5 (8.5)	1.3 (7.9)*	0.96	0.2 (2.4)	1.3 (13.4)	0.04
Number of steps with duration ≤1 min	4592.2 (2101.9)	4591.9 (2101.5)	1.00	3.1 (8.6)	0.1 (0.2)	0.00
Number of steps with duration >1 to 5 min	1607.6 (1616.4)	1566.2 (1573.4)*	0.98	42.7 (328.3)	2.0 (12.9)	0.03
Number of steps with duration >5 to 10 min	512.6 (970.3)	497.2 (953.8)*	0.98	15.5 (175.0)	1.8 (18.0)	0.02
Number of steps with duration >10 to 20 min	287.3 (865.9)	269.3 (826.9)*	0.96	18.0 (231.2)	1.4 (16.1)	0.04
Number of steps with duration >20 min	171.9 (976.3)	151.7 (914.5)*	0.96	20.3 (271.5)	1.0 (13.5)	0.04
Stepping time (min) at cadence ≥75 steps/min	43.7 (33.2)	42.9 (32.5)	0.99	0.8 (5.7)	1.5 (9.5)	0.04
Stepping time (min) at cadence ≥75 steps/min, duration >1 min	23.1 (26.8)	22.3 (26.1)*	0.98	0.8 (5.7)	2.6 (17.0)	0.05
Number of steps at cadence ≥75 steps/min	4217.2 (3432.4)	4128.5 (3346.9)	0.98	91.2 (654.0)	1.6 (10.6)	0.04
Number of steps at cadence ≥75 steps/min, duration >1 min	2367.9 (2871.0)	2279.3 (2777.9)*	0.97	89.1 (654.3)	2.7 (17.6)	0.06

Stepping time (min) at cadence ≥100 steps/min	17.1 (22.6)	16.5 (22.0)*	0.98	0.6 (4.6)	2.3 (16.2)	0.05
Stepping time (min) at cadence ≥100 steps/min, duration >1 min	13.3 (20.8)	12.8 (20.2)*	0.98	0.6 (4.6)	2.9 (20.2)	0.05
Number of steps at cadence ≥100 steps/min	1903.4 (2602.9)	1832.5 (2520.8)*	0.98	71.8 (561.9)	2.4 (16.7)	0.05
Number of steps at cadence ≥100 steps/min, duration >1 min	1499.4 (2401.2)	1428.5 (2321.1)*	0.97	71.3 (562.0)	2.9 (20.6)	0.05

*Indicates that VANE and CREA algorithms are not equivalent at a 3% threshold.

Figure 1. Equivalence testing for VANE and CREA algorithms for primary outcome metrics.

Data are considered equivalent if the mean difference (square marker on each plot) and its 95% confidence interval lies entirely within the 3% equivalence zone (represented as black, dotted vertical lines on each plot).

Figure 2. Bland-Altman plots comparing VANE and CREA algorithms for primary outcome metrics.