J Meas Phys Behav. Author manuscript; available in PMC 2021 May 14.

Published in final edited form as:

J Meas Phys Behav. 2020 June ; 3(2): 110–117. doi:10.1123/jmpb.2019-0049.

Alternative Wear-time Estimation Methods Compared to Traditional Diary Logs for Wrist-Worn ActiGraph Accelerometers in Pregnant Women

Samantha F. Ehrlich, PhD MPH
Division of Research, Kaiser Permanente Northern California and The University of Tennessee, Knoxville;

Amanda J. Casteel, MPH, **Scott E. Crouter, PhD**, **Paul R. Hibbing, PhD**
The University of Tennessee, Knoxville;

Monique M. Hedderson, PhD, **Susan D. Brown, PhD**, **Maren Galarce, MPH**
Division of Research, Kaiser Permanente Northern California;

Dawn P. Coe, PhD, **David R. Bassett, PhD**
The University of Tennessee, Knoxville;

Assiamira Ferrara, MD PhD
Division of Research, Kaiser Permanente Northern California;

Abstract

Background: This study sought to compare three sensor-based wear-time estimation methods to conventional diaries for ActiGraph wGT3X-BT accelerometers worn on the non-dominant wrist in early pregnancy.

Methods: Pregnant women (n= 108) wore ActiGraph wGT3X-BT accelerometers for 7 days and recorded their device on and off times in a diary (criterion). Average daily wear-time estimates from the Troiano and Choi algorithms and the wGT3X-BT accelerometer wear sensor were compared against the diary. The Hibbing 2-regression model was used to estimate time spent in activity (during periods of device wear) for each method. Wear-time and time spent in activity were compared with multiple repeated measures ANOVAs. Bland Altman plots assessed agreement between methods.

Results: Compared to the diary [825.5 minutes (795.1, 856.0)], the Choi [843.0 (95% CI 812.6, 873.5)] and Troiano [839.1 (808.7, 869.6)] algorithms slightly overestimated wear-time, whereas the sensor [774.4 (743.9, 804.9)] underestimated it, although only the sensor differed significantly from the diary ($P < .0001$). Upon adjustment for average daily wear-time, there were no statistically significant differences between the wear-time methods in regards to minutes per day of moderate to vigorous physical activity (MVPA), vigorous PA, and moderate PA. Bland Altman plots indicated the Troiano and Choi algorithms were similar to the diary and within ≤ 0.5% of each other for wear-time and MVPA.

sehrlic1@utk.edu.

Conclusions: The Choi or Troiano algorithms offer a valid and efficient alternative to diaries for the estimation daily wear-time in larger-scale studies of MVPA during pregnancy, and reduce burden for study participants and research staff.

Keywords

ActiGraph accelerometers; wrist-worn; wear-time; pregnancy

INTRODUCTION

The Physical Activity Guidelines for Americans and the American College of Obstetricians and Gynecologists recommend that pregnant women participate in at least 150 minutes of moderate-intensity aerobic physical activity (PA) per week (Piercy et al., 2018; The American College of Obstetricians and Gynecologists, 2015). PA during pregnancy is safe (The American College of Obstetricians and Gynecologists, 2015) and beneficial in terms of improving or maintaining physical fitness (Kramer & McDonald, 2006) and pregnancy and postpartum weight management (Ruchat et al., 2018). Epidemiologic studies of PA during pregnancy have largely relied on self-report via questionnaires (Evenson et al., 2012) although they are prone to measurement error (Matthews et al., 2012).

Device-based measures of PA require the estimation of wear-time, or the amount of time the tracking device was worn. Calculation of the amount of time spent in activity of various intensities is then limited to periods of device wear. Historically, diaries have been used to assess wear time; participants record on/off times, and the logs also serve as a reminder to wear the device (Lee & Shiroma, 2014). More recently, large epidemiologic studies favor the use of objective wear-time algorithms, such as the Troiano (Troiano et al., 2008a) and Choi (Choi, Liu, et al., 2011) algorithms, though only the Choi algorithm has been validated in wrist-worn devices (Choi et al., 2012). The ActiGraph wGT3X-BT is a popular research grade accelerometer equipped with capacitive coupling technology that detects when the device is in proximity to the skin, thereby serving as a new option for wear-time estimation.

Despite the proliferation of wrist-worn devices in PA research, there are few published data to guide investigators on selecting a wear-time estimation method for use with wrist-worn devices. To the best of our knowledge, no published study has compared various wear-time estimation methods to traditional diary logs for wrist-worn devices in regard to wear-time and the corresponding amounts of time spent in activity, by intensity level. As such, the aim of the current study was to compare three wear-time estimation methods (the Choi and Troiano algorithms, and the ActiGraph wGT3X-BT sensor) with traditional diary logs for ActiGraph wGT3X-BT accelerometers worn on the non-dominant wrist during early pregnancy.

METHODS

Study Setting and Parent Trial

The study setting was Kaiser Permanente Northern California (KPNC), a large, integrated healthcare delivery system with an expansive electronic health record (EHR) system. The membership of KPNC is broadly representative of the underlying geographic areas served. (Gordon, 2012) Data for the current study came from the GestationaL Weight Gain and Optimal Wellness (GLOW) randomized clinical trial [ClinicalTrials.gov: NCT02130232]. The GLOW trial is described in detail elsewhere (Brown et al., 2019). Briefly, in 2014–2016, pregnant women participating in the GLOW trial were randomized to receive usual care or usual care plus a lifestyle intervention to prevent excess gestational weight gain. Eligibility criteria included ≥18 years of age, pre-pregnancy BMI between 25.2 and 40.0 kg/m^2, singleton pregnancy, and identification for recruitment at <8 weeks' gestational age in the EHR. The primary outcome of the GLOW trial was gestational weight gain; change in diet and PA were secondary outcomes.

Eligible women attended a study clinic visit at their regular KPNC facility, where informed consent was obtained and baseline questionnaire data collected. Prior to the baseline study clinic visit, ActiGraph wGT3X-BT devices were fully charged and initialized with a 30Hz sampling rate. At the visit, research staff explained the PA assessment protocol and provided a device and wrist strap, a diary log (Lee & Shiroma, 2014) (i.e., small paper booklet) to record their daily on and off times, and a postage paid envelope to return the device and diary log at the end of the assessment period. Participants were instructed to securely strap the device to their non-dominant wrist with the ActiGraph logo facing up and to wear it for 7 days (non-consecutive days allowed), though some participants opted to wear it longer. Since negative feedback was received regarding overnight wear when piloting the devices prior to the initiation of the study, and because sleep was not an outcome of the GLOW trial, participants were instructed to remove the device before going to bed and to put it back on upon waking, though some participants opted to wear the device overnight (i.e., continuously; overnight wear was included as wear-time in all analyses). They were instructed to remove the device when bathing and swimming (i.e., to not submerge the device in water) and to record all breaks in wear in the diary log (Brown et al., 2019). The diary logs were manually examined for errors (e.g., am/pm mix ups, missing on/off times) and corrected by logic, examination of the on/off times reported on neighboring days and/or visual inspection of the device data in ActiLife software; there were errors necessitating visual inspection for <2% of the diary data points. Commonly reported reasons for device removal in the diary logs included sleep and bathing.

Data Processing

For the purpose of this study, the first consecutive 108 GLOW participants with baseline PA data (i.e., accelerometer and diary log data) were used for the analysis. Device data were downloaded using ActiLife software version 16.13.3 (ActiGraph, Pensacola, FL, USA), then converted to raw (30Hz) and count (1-s epoch with the low-frequency extension enabled) .csv files. Estimates of wear-time were then obtained using the Choi and Troiano algorithms and the GT3X wear sensor for comparison to the wear-time obtained from the participant diary log.

The 1-s count files were read into R and reintegrated to 1-min epochs using the 'AGread' R package (Hibbing et al., 2019). The Choi algorithm was then implemented using the 'PhysicalActivity' R package (Choi, Zhouwen, et al., 2011), with vertical axis counts. The algorithm of Choi et al. (Choi, Liu, et al., 2011; Choi et al., 2012) defines non-wear-time as periods of 90 or more consecutive minutes of 0 counts/minute, using three 30-minute sliding windows (Choi, Liu, et al., 2011; Choi et al., 2012). For the Troiano algorithm, the 1-min count data were processed in R using code that was replicated from the original SAS code (Aguilar-Farias et al., 2014; Troiano et al., 2008b). In the Troiano wear-time algorithm, non-wear-time is defined as 60 or more consecutive minutes of 0 counts/minute, with allowance for 1–2 minutes of 0–99 counts/min during this time. For the wear-sensor, G3TX files were converted into .agd files in the ActiLife software and then individually run through the software's wear-time validation function to determine periods of wear and non-wear. The GT3X-BT wear sensor operates via capacitive coupling to detect whether the wrist-worn device is being worn, with data gathered continuously following device initialization.

For estimates of time spent in sedentary behavior, light intensity PA (LPA), moderate intensity PA (MPA), vigorous intensity PA (VPA), and moderate to vigorous intensity PA (MVPA), the raw .csv files were read into R and reduced to 1-s epochs using the 'AGread' package (Hibbing et al., 2019). Afterwards, the two-regression algorithms of Hibbing et al. (Hibbing et al., 2018) were implemented using the 'TwoRegression' R package (Hibbing & van Hees, 2018). The two-regression algorithms (i.e., the Hibbing approach hereafter) were developed using the ActiGraph GT9X, with specific algorithms for right and left wrist data (raw acceleration in 1-s epochs), and validated against indirect calorimetry in healthy adults in a structured laboratory setting. Briefly, the algorithms first classify sedentary behavior and assign an energy expenditure of 1.25 METs. The remaining data are classified as continuous walk/run or intermittent activity based on the signal variability (i.e., coefficient of variation), then fed into corresponding regression equations to predict a MET value each second. Once the second-by-second MET values have been obtained, they are averaged over each minute to obtain 1-min MET values. Estimates of time spent in PA intensity categories are then calculated based on: sedentary behavior (≤ 1.5 METs), LPA (1.6–2.9 METs), MPA (3.0–5.9 METs), VPA (≥ 6 METs), and MVPA (≥ 3 METs).

Data containing the minute-by-minute PA intensity classifications were then then merged with the minute-by-minute wear time estimates (diary log, Choi, Troiano, and wear sensor). For each participant, the first day of data was removed (i.e., the day of the study clinic visit), and the total number of days included was capped at six days (non-consecutive days allowed). For the purpose of this analysis, the diary log is considered the criterion method against which the other methods are compared. As such, days were only included in the calculations of average daily wear-time and time spent per day in activity, by intensity category, if any wear-time was identified on that calendar day by the diary log (Keadle et al., 2014). Discrepancies between the wear-time methods regarding the number of days the device was worn were also examined. Specifically, the per participant average number of valid days, defined as ≥600 minutes of device wear-time, was calculated and compared between the methods. In addition, the percentage of the cohort to be retained if ≥3 valid days or if ≥4 valid days were used as the criterion for inclusion in study analysis was compared between methods, as these are commonly used criterion for accelerometer data in lifestyle intervention studies (Ekelund et al., 2019; Jake-Schoffman et al., 2019; Migueles et al., 2017).

Statistical Analyses

Multiple repeated measures ANOVAs were used to investigate the differences in average daily wear-time (per participant) and the corresponding estimates of time spent in activity, by intensity category. Total days of device wear were used to calculate means (i.e., calculations of means ± SD were not limited to valid days). Analyses of time spent in activity by intensity category were additionally adjusted for average daily wear-time. The tests were completed using the Proc Mixed procedure in SAS, with compound symmetry as the covariance structure. To account for multiple comparisons, the threshold for significance was set conservatively at $p < .01$. Bland-Altman plots were constructed in Excel to examine the individual level of agreement between the methods.

This study was approved by the Institutional Review Boards of the University of Tennessee Knoxville and Kaiser Permanente Northern California.

RESULTS

Participant characteristics (n= 108) are presented in Table 1. Over half of the women were from minority racial-ethnic groups, and 55% were pregnant with their first child. The mean age was 32.7 (SD= 4.8) years and mean gestational age at the PA assessment was 11.8 (SD= 1.2) weeks. There were seven women (6.5%) who opted to wear the device overnight for at least some part of the assessment period. On average, there were 4.9 (SD= 0.7) days of device wear per PA assessment period, as identified by the diary. The other wear-time methods identified wear-time on days with no wear-time recorded in the diary: there were 5 such days (in 5 separate women) identified by the Choi algorithm, 8 days (in 8 women) by the Troiano algorithms, and 4 days (in 4 women) by the sensor.

Table 2 presents the average daily wear-time estimate for each method. No differences in wear-time were detected for estimates obtained by the Choi and Troiano algorithms compared to the diary.. Daily wear-time estimated by the sensor differed significantly from that by the diary; daily wear-time was, on average, 51.1 minutes lower for the sensor compared to the diary (95% CI: 34.5, 67.7).

Minutes spent in each intensity category by the Hibbing approach, both unadjusted and adjusted for average daily wear-time, are also presented in Table 2. In the unadjusted analyses, with the Choi and Troiano algorithms, estimates of minutes per total day spent in MVPA, MPA and LPA were similar to one another other and significantly different from the diary log (all $p < .002$). For MVPA/MPA, the differences between estimates obtained using both algorithms and the diary log were under 2 minutes. For LPA, use of the Choi and Troiano algorithms resulted in an estimated additional 8.6 minutes (95% CI 2.7, 14.5) and 8.4 minutes (95% CI 2.5, 14.4) per day, respectively, compared to use of the diary log for wear time. Estimates of time spent in activity based upon the sensor were significantly different from those obtained using the diary log for wear-time across all intensity categories (all $p < .01$). For MVPA/MPA, the difference between estimates obtained using the sensor and the diary log was also under 2 minutes.

In analyses adjusted for average daily wear-time, no differences were observed between the three methods in regard to estimated minutes spent in MVPA, VPA, and MPA per day by the Hibbing approach (Table 2). For LPA and sedentary behavior, the adjusted estimates from the Troiano and sensor wear-time methods were statistically significantly different from those obtained using the diary log for wear time. For LPA, use of the Troiano algorithm an additional 4.4 minutes (95% CI 1.2, 7.7) per day and the sensor an additional 5.3 minutes (95% CI 1.9, 8.7) per day. For sedentary time, use of the Troiano algorithm 5.2 fewer minutes (95% CI 1.4, 9.0) per day and the sensor 5.7 fewer minutes (95% CI 1.7, 9.6) per day than use of the diary log for wear time.

Bland-Altman plots comparing the diary log to the three other methods' estimates of wear-time and corresponding MVPA, LPA and sedentary time by the Hibbing approach are presented in Figure 1. Variability was consistent across all plots. As compared to the diary, there was little bias in the Choi and Troiano algorithms' estimates of device wear-time (−17.5 minutes and −13.6 minutes per day, respectively) and corresponding estimated time spent in MVPA (both −1.3 minutes per day) and LPA (−8.6 and −8.4 minutes per day, respectively), with narrow limits of agreement. All plots comparing the sensor to the diary had wide limits of agreement (i.e., −168.8 to 273.8 minutes per day for wear-time, −6.2 to 3.6 minutes per day for MVPA, −76.6 to 96.1 minutes per day for LPA, and −124.3 to 203.8 minutes per day for sedentary time). Figure 2 displays the Bland-Altman plots comparing the Choi and Troiano methods to one another. The Choi and Troiano methods' estimates were within 0.5% of each other for wear-time and 0.09% for MVPA, 0.06% for LPA, and 0.7% for sedentary time by the Hibbing approach.

The difference in the number of valid days (i.e., ≥600 minutes of device wear-time) identified by the four wear-time methods attained statistical significance ($P < .0001$), although the differences were not practically meaningful. Specifically, the diary log identified 4.6 (SD= 0.9) valid days, compared to 4.5 (SD= 1.1) valid days for the Choi algorithm, 4.5 (SD= 1.0) for the Troiano algorithm, and 4.2 (SD= 1.2) valid days for the sensor. If 3+ valid days had been the criteria for inclusion, use of the diary log to estimate wear-time would result in 97.2% (n= 105) of the cohort being retained for analyses, compared to 93.5% (n= 101) for the Choi algorithm, 92.6% (n= 100) for the Troiano algorithm, and 83.3% (n= 90) for the sensor. If 4+ valid days had been the criteria for inclusion, use of the diary log to estimate wear-time would result in 88.0% (n= 95) of the cohort being retained for analyses, compared to 86.1% (n= 93) for the Choi algorithm, 85.2% (n= 92) for the Troiano algorithm, and 77.8% (n= 84) for the sensor.

DISCUSSION

This study compared wear-time estimates obtained from three sensor-based methods (the Choi and Troiano algorithms, and the ActiGraph wGT3X-BT wear sensor) to traditional diary logs (i.e., self-report) and additionally examined the impact of wear-time method on corresponding estimates of time spent in activity, by intensity level. The study is unique compared to previous studies in that the data were obtained from wrist-worn devices, instead of hip-worn devices, and device on and off times were recorded in traditional diary logs. In this study, with adjustment for device wear-time, use of the Choi and Troiano algorithms and the sensor for wear time all provided comparable estimates of time spent in MVPA, VPA and MPA by the Hibbing approach as compared to use of the diary log. However, in Bland-Altman plots comparing the sensor to the diary log, the limits of agreement were wide for all metrics examined, indicating that the sensor should not be used in place of the diary log for wear-time estimation. Our results suggest that burden may be reduced by automating the identification of non-wear with the Choi or Troiano algorithms—particularly for studies of MVPA, but perhaps not for studies of sedentary behavior and LPA. Discrepancies between the wear time methods' corresponding estimates of time spent in activity would be expected for LPA and/or sedentary behavior, since low signal in those categories could potentially promote misclassification as non-wear.

There is no true 'gold standard' method for wear time estimation. Although diary logs are commonly used (particularly in smaller studies), they are burdensome for participants, which often leads to partial or total incompletion. Furthermore, recording device on and off times is a behavior susceptible to recall and social desirability biases (Peeters et al., 2013; Pulakka et al., 2018; Rillamas-Sun et al., 2015). Diaries create a similar burden for research staff, who must enter, clean, and process the diary data. Despite their limitations, diaries are useful for reminding participants to wear the devices, and may be used to identify the start (calendar date) of the assessment period (Keadle et al., 2014). Diaries also allow participants to report what they were doing during periods of non-wear, which can provide contextual information, which may potentially allow for imputation (i.e., of MET values for activities they reportedly engaged in during periods of non-wear, such as swimming).

The Choi and Troiano wear time algorithms are often used instead of diaries, to automate the process of addressing non-wear time. The algorithms provide a quicker, objective, and more cost-effective solution for accounting for non-wear time when processing accelerometer data (Rillamas-Sun et al., 2015). In the current study, estimates of the amount of time spent in MVPA obtained using the Choi and Troiano wear-time methods and the Hibbing approach were nearly identical to each other and highly compatible with the corresponding estimate of MVPA obtained using the diary log to identify wear time.

This finding is consistent with the conclusions of analysis of over 500 older adults in the Finnish Retirement and Aging Study (FIREA) who wore ActiGraph wActiSleep-BT accelerometers on their non-dominant wrist for a 7-day period (Pulakka et al., 2018). FIREA study participants were instructed to wear the device at all times and record their sleep and wake times in a diary log, whereas the current study asked participants to record device on and off times specifically. The FIREA study reported no substantial difference in the wake wear-times estimated by the sleep log and the Choi algorithm, and little variability in wake wear-time derived from the sleep log, the Choi algorithm, and the sensor (i.e., only up to 24 minutes per day). Visual inspection of the FIREA data suggested that the sensor had indicated non-wear during periods of apparent device wear, leading the authors to question the accuracy of the sensor and only include data from functioning sensors in the analyses (i.e., 71% of the cohort) (Pulakka et al., 2018).

Also compatible with the current study are the results of a follow-up to the Women's Health Study (WHS), conducted in over 8,000 participants who wore ActiGraph GT3X+ devices on the hip for 7 consecutive days during waking hours (Keadle et al., 2014). Similar to the current study, a wear-time diary log was used to identify the calendar dates of wear and several wear-time estimation methods subsequently compared. Wear-time estimates obtained from the diary log were found to be very similar to the wear-time estimates obtained from the Choi algorithm [i.e., median daily wear times, using vector magnitude data, of 898 minutes (IQR 851, 937) and 896 minutes (IQR 848, 940), respectively]. Wear time estimates obtained from the Troiano algorithm were slightly lower than those from the Choi algorithm in the WHS follow up study, but the difference did not attain statistical significance [i.e., median daily wear times of 867 minutes (IQR 813, 912) and 896 minutes (IQR 848, 940), respectively] (Keadle et al., 2014). As in the current study, wear-time estimation method did not meaningfully impact estimates of time spent in MVPA (Keadle et al., 2014).

The current study has several strengths and limitations worth noting. The racial-ethnic diversity of the GLOW participants is a clear strength. Placing the accelerometers on the wrist, although preferable for pregnant populations, may overestimate PA due to the detection of upper body movement in pregnant and non-pregnant adults alike. As previously mentioned, the diary log is subject to bias and though traditionally used to estimate device wear-time, it serves as an imperfect wear-time criterion. As such, it is difficult to pinpoint why the alternative wear-time methods identified wear-time on days with no wear-time recorded in the diary (e.g., erroneous diary log entries, devices being carried in bags, and/or other sources of error). In the current study, estimates of time spent in activity, by intensity category, are based on an algorithm (i.e., the Hibbing approach) and thus potentially subject to error. The Hibbing approach was developed using data from the GT9X primary accelerometer, the same used in the wGT3X-BT devices used in this study. Study results should therefore not be affected by differences between the GT9X and the wGT3X-BT. Fortunately, any error due to device differences would be consistent across the wear time methods compared in the current study. Finally, the current study examined count-based non-wear methods, although methods have been developed for raw acceleration data. It tends to be easier to implement count-based methods than raw-based methods, thus count-based methods are more commonly used. Future work should examine the effectiveness of non-wear methods for raw acceleration data and aim to make those methods more user friendly.

In conclusion, for ActiGraph wGT3X-BT devices worn on the non-dominant wrist of pregnant women, the Choi and Troiano wear-time algorithms were equally suitable alternate methods to traditional diary logs for estimating device wear-time and resulted in comparable estimates of minutes per day spent in MVPA by the Hibbing approach. Use of the Choi or Troiano algorithms to estimate device wear-time may substantially reduce the burden of study participants and research staff, thereby increasing data completeness and the efficiency of future epidemiological studies objectively assessing time spent in MVPA.

ACKNOWLEDGEMENTS

The GLOW trial was funded by the National Institutes of Child Health and Human Development, R01 HD073572 to Dr. Ferrara. Additional support for this study was received from the National Institute of Diabetes and Digestive and Kidney Diseases: P30 DK092924 to Dr. Ferrara, K01 DK099404 to Dr. Brown, and K01 DK105106 to Dr.

Ehrlich. We gratefully acknowledge the GLOW trial participants and staff, as well as Pragya Poudel, Jonathan Thomas, Claire Holiday, and Sarah Walsh for processing the accelerometer data.

References

Aguilar-Farias N, Brown WJ, & Peeters GM (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. doi:10.1016/j.jsams.2013.07.002 [PubMed: 23932934]

Brown SD, Hedderson MM, Ehrlich SF, Galarce MN, Tsai AL, Quesenberry CP, & Ferrara A (2019). Gestational weight gain and optimal wellness (GLOW): rationale and methods for a randomized controlled trial of a lifestyle intervention among pregnant women with overweight or obesity. BMC Pregnancy and Childbirth, 19(1), 145. doi:10.1186/s12884-019-2293-8 [PubMed: 31039753]

Choi L, Liu Z, Matthews CE, & Buchowski MS (2011). Validation of Accelerometer Wear and Nonwear Time Classification Algorithm. Medicine and Science in Sports and Exercise, 43(2), 357–364. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3184184/. doi:10.1249/MSS.0b013e3181ed61a3 [PubMed: 20581716]

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

Choi L, Zhouwen L, Matthews CE, & Buchowski MS (2011). PhysicalActivity: Process physical activity accelerometer data. Retrieved from https://CRAN.R-project.org/package=PhysicalActivity

Ekelund U, Tarp J, Steene-Johannessen J, Hansen BH, Jefferis B, Fagerland MW, … Lee IM (2019). Dose-response associations between accelerometry measured physical activity and sedentary time and all cause mortality: systematic review and harmonised meta-analysis. BMJ, 366, l4570. doi:10.1136/bmj.l4570 [PubMed: 31434697]

Evenson KR, Chasan-Taber L, Symons Downs D, & Pearce EE (2012). Review of self-reported physical activity assessments for pregnancy: summary of the evidence for validity and reliability. Paediatric and Perinatal Epidemiology, 26(5), 479–494. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/22882792. doi:10.1111/j.1365-3016.2012.01311.x [PubMed: 22882792]

Gordon NP (2012). Similarity of the Adult Kaiser Permanente Membership in Northern California to the Insured and General Population in Northern California: Statistics from the 2009 California Health Interview Survey. Retrieved from Oakland CA: file:///C:/Users/sehrlic1/Downloads/chis_non_kp_2009%20(1).pdf

Hibbing PR, Lamunion SR, Kaplan AS, & Crouter SE (2018). Estimating Energy Expenditure with ActiGraph GT9X Inertial Measurement Unit. Medicine and Science in Sports and Exercise, 50(5), 1093–1102. doi:10.1249/mss.0000000000001532 [PubMed: 29271847]

Hibbing PR, & van Hees VT (2018). TwoRegression: Process Data from Wearable Research Devices Using Two-Regression Algorithms Retrieved from https://cran.r-project.org/web/packages/TwoRegression/index.html

Hibbing PR, van Hees VT, LaMunion SR, Judge D, & Maygarden J (2019). AGread: Read Data Files from ActiGraph Monitors. Retrieved from https://cran.r-project.org/web/packages/AGread/index.html

Jake-Schoffman DE, Silfee VJ, Sreedhara M, Rosal M, May CN, Lopez-Cepero A, … Haughton CF (2019). Reporting of Physical Activity Device Measurement and Analysis Protocols in Lifestyle Interventions American Journal of Lifestyle Medicine, 8. doi:10.1177/1559827619862179

Keadle SK, Shiroma EJ, Freedson PS, & Lee IM (2014). Impact of accelerometer data processing decisions on the sample size, wear time and physical activity level of a large cohort study. BMC Public Health, 14(1), 1210. Retrieved from 10.1186/1471-2458-14-1210. doi:10.1186/1471-2458-14-1210 [PubMed: 25421941]

Kramer MS, & McDonald SW (2006). Aerobic exercise for women during pregnancy. Cochrane Database Syst Rev(3), Cd000180. doi:10.1002/14651858.CD000180.pub2 [PubMed: 16855953]

Lee IM, & Shiroma EJ (2014). Using accelerometers to measure physical activity in large-scale epidemiological studies: issues and challenges. British Journal of Sports Medicine, 48(3), 197. Retrieved from http://bjsm.bmj.com/content/48/3/197.abstract. [PubMed: 24297837]

Matthews CE, Moore SC, George SM, Sampson J, & Bowles HR (2012). Improving Self-reports of Active and Sedentary Behaviors in Large Epidemiologic Studies. Exercise and Sport Sciences Reviews, 40(3), 118–126. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3388604/. doi:10.1097/JES.0b013e31825b34a0 [PubMed: 22653275]

Migueles JH, Cadenas-Sanchez C, Ekelund U, Delisle Nystrom C, Mora-Gonzalez J, Lof 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. doi:10.1007/s40279-017-0716-0 [PubMed: 28303543]

Peeters G, van Gellecum Y, Ryde G, Aguilar-Farias N, & Brown W (2013). Is the pain of activity log-books worth the gain in precision when distinguishing wear and non-wear time for tri-axial accelerometers. Journal of Science and Medicine in Sport, 16, 515–519. [PubMed: 23294696]

Piercy K, Troiano R, Ballard R, Carlson S, Fulton J, Galuska D, … Olson R (2018). The Physical Activity Guidelines for Americans. JAMA, 320(19), 2020–2028. doi:10.1001/jama.2018.14854 [PubMed: 30418471]

Pulakka A, Shiroma E, Harris T, Pentti J, Vahtera J, & Stenholm S (2018). Classification and processing of 24-hour wrist accelerometer data. Journal for the Measurement of Physical Behavior, 1, 51–59.

Rillamas-Sun E, Buchner DM, Di C, Evenson KR, & LaCroix AZ (2015). Development and application of an automated algorithm to identify a window of consecutive days of accelerometer wear for large-scale studies. BMC Research Notes, 8(1). doi:10.1186/s13104-015-1229-2

Ruchat SM, Mottola MF, Skow RJ, Nagpal TS, Meah VL, James M, … Davenport MH (2018). Effectiveness of exercise interventions in the prevention of excessive gestational weight gain and postpartum weight retention: a systematic review and meta-analysis. British Journal of Sports Medicine, 52, 1347–1356. [PubMed: 30337461]

The American College of Obstetricians and Gynecologists. (2015). ACOG Committee Opinion No. 650: Physical Activity and Exercise During Pregnancy and the Postpartum Period. Obstetrics and Gynecology, 126(6), e135–142. doi:10.1097/aog.0000000000001214 [PubMed: 26595585]

Troiano R, Berrigan D, Dodd K, Masse L, Tilert T, & McDowell M (2008a). Physical activity in the United States measured by accelerometer. Medicine and Science in Sports and Exercise, 40, 181–188. [PubMed: 18091006]

Troiano R, Berrigan D, Dodd K, Masse L, Tilert T, & McDowell M (2008b). SAS Programs for Analyzing NHANES 2003–2004 Accelerometer Data. In: NIH National Cancer Institute

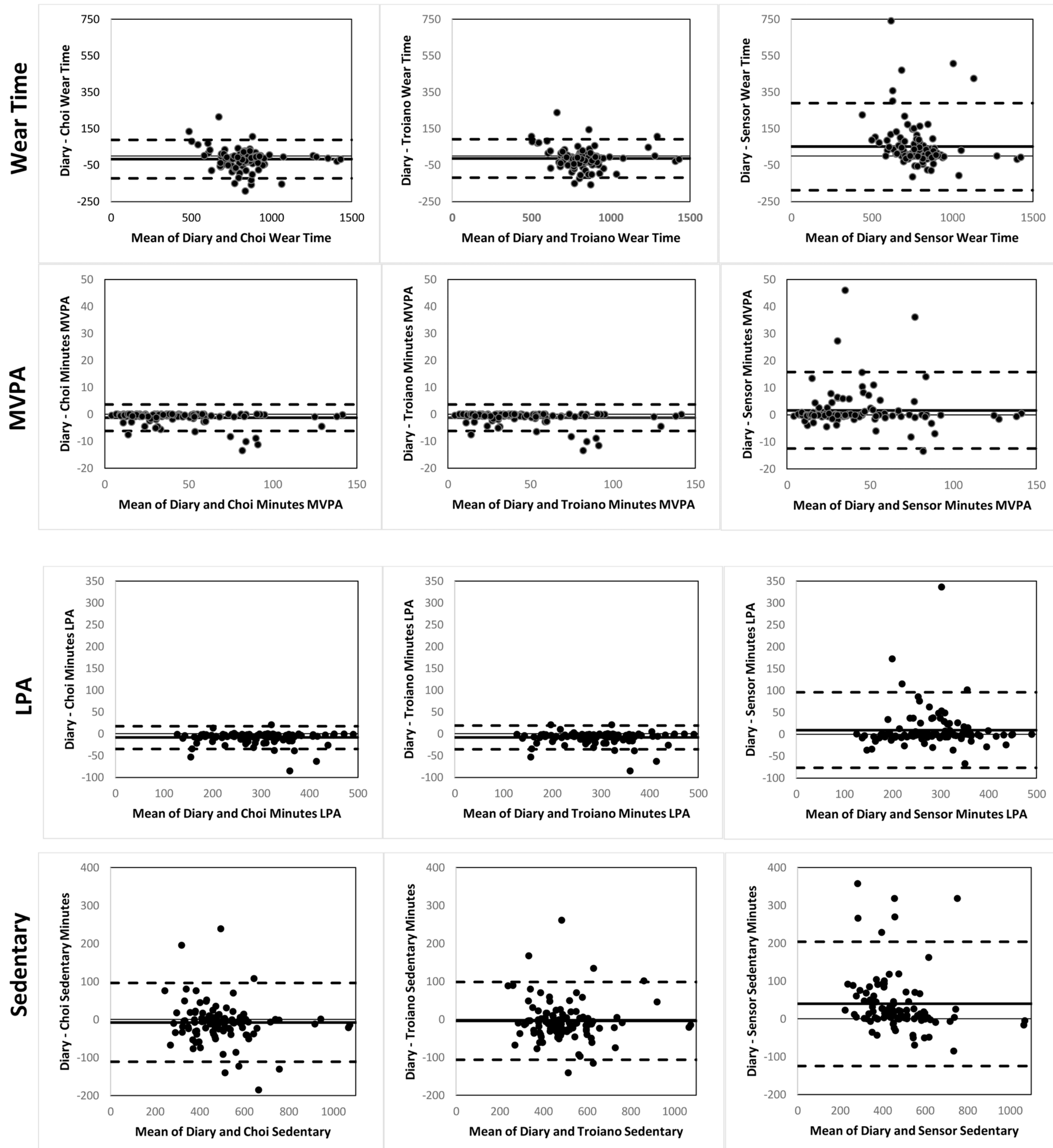
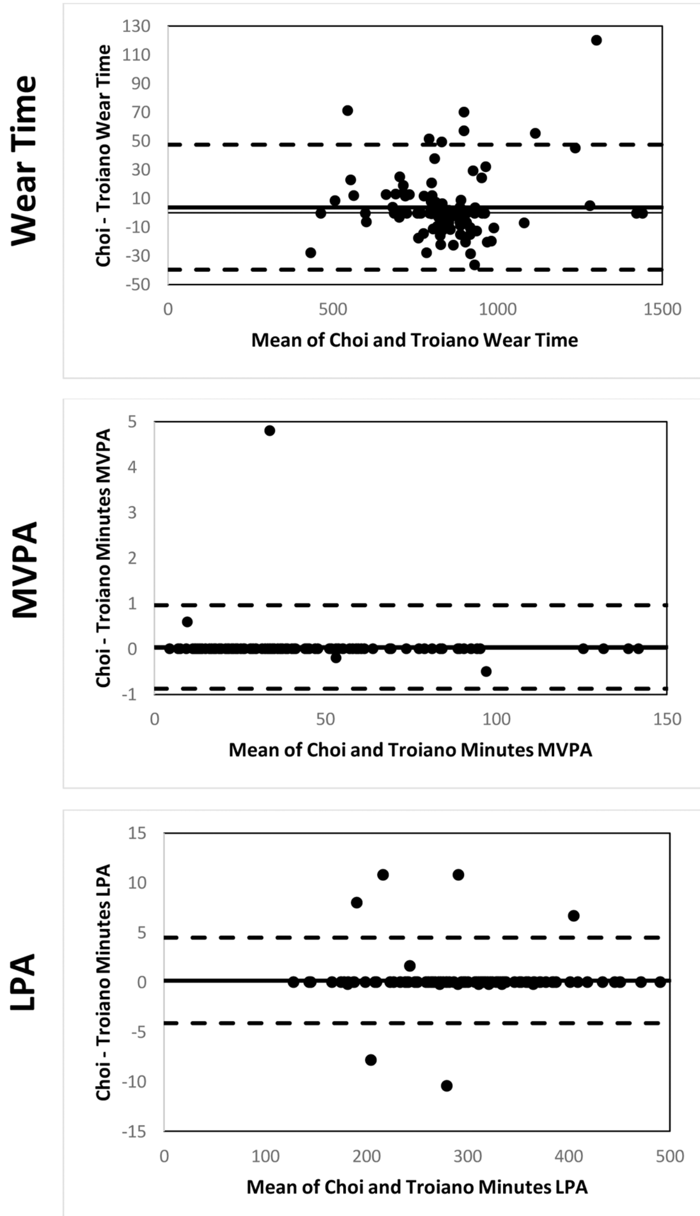


Figure 1.

Bland-Altman plots comparing the diary log to the other methods' estimates of wear-time, MVPA, LPA and sedentary time, the GLOW study (n=108), Kaiser Permanente Northern California, 2014–2015.


Wear Time
Choi - Troiano Wear Time
Mean of Choi and Troiano Wear Time
MVPA
Choi - Troiano Minutes MVPA
Mean of Choi and Troiano Minutes MVPA
LPA
Choi - Troiano Minutes LPA
Mean of Choi and Troiano Minutes LPA

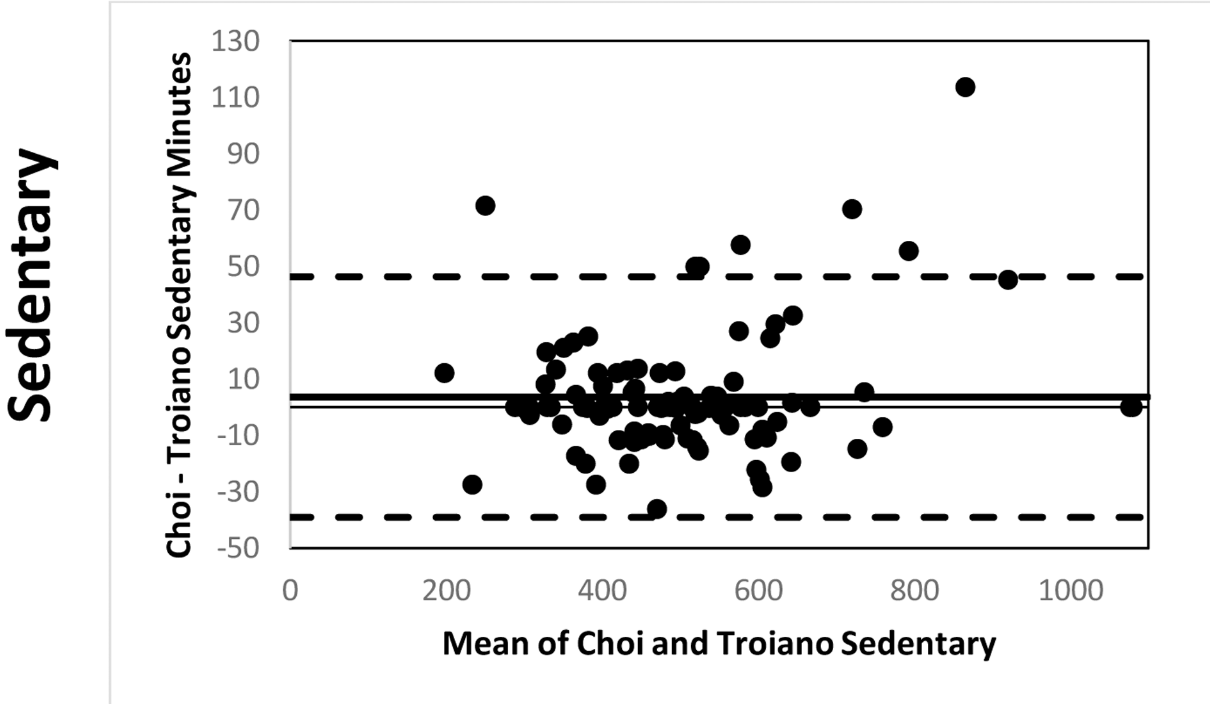


Figure 2.

Bland-Altman plots comparing the Choi and Troiano algorithms' estimates of wear-time, MVPA, LPA and sedentary time, the GLOW study (n=108), Kaiser Permanente Northern California, 2014–2015.

Table 1.

Cohort characteristics, the GLOW study (n=108), Kaiser Permanente Northern California, 2014–2015.

	N(%)
Age (years)	
18–24	6 (5.6%)
25–29	23 (21.3%)
30–34	45 (41.7%)
35+	34 (31.5%)
Gestational age at the physical activity assessment (weeks)	
8–10	28 (25.9%)
11–12	60 (55.6%)
13–15	20 (18.5%)
Pre-pregnancy BMI (kg/m^2)	
Overweight (≥25.2)	68 (63.0%)
Obese (≥30)	40 (37.0%)
Race-ethnicity	
White	41 (38.3%)
Latina (Hispanic/Latin American)	20 (18.7%)
Asian (or Pacific Islander)	18 (16.8%)
Multiracial	18 (16.8%)
Black/African American	10 (9.4%)
Education	
High school or less	10 (9.3%)
Some college	25 (23.2%)
4-year college graduate	30 (27.8%)
Postgraduate	43 (39.8%)
Parity	
0	59 (54.6%)
1	39 (36.1%)
2+	10 (9.3%)
GLOW treatment group	
Usual care	52 (48.1%)
Intervention	56 (51.9%)

Table 2.

Estimates of average daily wear-time and time spent in physical activity* by wear time-method, the GLOW study (n=108), Kaiser Permanente Northern California, 2014–2015.

	Average Daily Wear-Time Mean mins/day (95% CI)	MVPA Mean mins/day (95% CI)		VPA Mean mins/day (95% CI)		MPA Mean mins/day (95% CI)		LPA Mean mins/day (95% CI)		Sedentary Behavior Mean mins/day (95% CI)	
Method		**Unadjusted**	**Adjusted[b]**	**Unadjusted**	**Adjusted[b]**	**Unadjusted**	**Adjusted[b]**	**Unadjusted**	**Adjusted[b]**	**Unadjusted**	**Adjusted[b]**
Diary	825.5 (795.1, 856.0)	42.9 (37.5, 48.2)	42.7 (37.3, 48.0)	1.79 (1.13, 2.45)	1.78 (1.12, 2.44)	41.1 (35.9, 46.2)	40.9 (35.8, 46.0)	290.1 (275.5, 304.7)	288.6 (274.9, 302.4)	492.6 (464.1, 521.1)	489.2 (473.2, 505.3)
Choi	843.0 (812.6, 873.5)	44.2[c] (38.8, 49.6)	43.3 (38.0, 48.6)	1.82 (1.16, 2.47)	1.76 (1.10, 2.42)	42.4[c] (37.2, 47.6)	41.5 (36.4, 46.7)	298.7[c] (284.1, 313.3)	292.1 (278.3, 305.9)	500.1 (471.6, 528.6)	485.1 (469.1, 501.2)
Troiano	839.1 (808.7, 869.6)	44.1[c] (38.8, 49.5)	43.4 (38.1, 48.7)	1.82 (1.16, 2.48)	1.78 (1.12, 2.44)	42.3[c] (37.1, 47.5)	41.6 (36.5, 46.8)	298.5[c] (283.9, 313.1)	293.1[c] (279.3, 306.9)	496.4 (468.0, 524.9)	484.0[c] (468.0, 500.1)
Sensor	774.4[c,d,e] (743.9, 804.9)	41.2[c,d,e] (35.8, 46.6)	43.0 (37.7, 48.4)	1.69[c,d,e] (1.03, 2.35)	1.80 (1.14, 2.46)	39.5[c,d,e] (34.3, 44.7)	41.2 (36.1, 46.4)	280.4[c,d,e] (265.8, 295.0)	293.9[c] (280.1, 307.8)	452.8[c,d,e] (424.3, 481.3)	483.6[c] (467.5, 499.7)
Overall *p*-value[a]	<.0001	<.0001	.19	.002	.76	<.0001	.15	<.0001	.01	<.0001	.02

* Estimates of total wear-time and time spent in PA, by intensity category, are for total days (i.e., not limited to valid days)

[a] Comparisons from multiple repeated measures ANOVA using the Proc Mixed procedure in SAS, with compound symmetry as the covariance structure

[b] Adjusted for average daily wear-time

[c] Significantly different from Diary, $P < .01$

[d] Significantly different from Choi, $P < .01$

[e] Significantly different from Troiano, $P < .01$