Paper:

Montoye, A., Nelson, M., Bock, J., Imboden, M., Kaminsky, L., Mackintosh, K., McNarry, M. & Pfeiffer, K. (2017). Raw and Count Data Comparability of Hip-Worn ActiGraph GT3X+ and Link Accelerometers. *Medicine & Science in Sports & Exercise*, 1

http://dx.doi.org/10.1249/MSS.0000000000001534

Medicine & Science IN Sports & Exercise

The Official Journal of the American College of Sports Medicine
www.acsm-msse.org

. . . Published ahead of Print

Raw and Count Data Comparability of Hip-Worn ActiGraph GT3X+ and Link Accelerometers

Alexander H.K. Montoye[1,2], M. Benjamin Nelson[2], Joshua M. Bock[2], Mary T. Imboden[2], Leonard A. Kaminsky[2], Kelly A. Mackintosh[3], Melitta A. McNarry[3], and Karin A. Pfeiffer[4]

[1]Alma College, Alma, MI; [2]Ball State University, Muncie, IN; [3]Swansea University, Swansea, Wales, United Kingdom; [4]Michigan State University, East Lansing, MI

Accepted for Publication: 17 December 2017

Medicine & Science in Sports & Exercise, Publish Ahead of Print
DOI: 10.1249/MSS.0000000000001534

Raw and Count Data Comparability of Hip-Worn ActiGraph GT3X+ and Link Accelerometers

Alexander H.K. Montoye[1,2], M. Benjamin Nelson[2], Joshua M. Bock[2], Mary T. Imboden[2], Leonard A. Kaminsky[2], Kelly A. Mackintosh[3], Melitta A. McNarry[3], and Karin A. Pfeiffer[4]

[1]Alma College, Alma, MI; [2]Ball State University, Muncie, IN; [3]Swansea University, Swansea, Wales, United Kingdom; [4]Michigan State University, East Lansing, MI

Corresponding author:

Alex Montoye

Research in Applied Physiology Laboratory

Integrative Physiology and Health Science Department

Alma College

614 W. Superior

Alma, MI 48801

Email: montoyeah@alma.edu

This study was supported by a Ball State University ASPiRE grant. The authors declare that they have no conflicts of interest to report. The results of this study are presented clearly, honestly, and without fabrication, falsification, or inappropriate data manipulation. The findings of the present study do not constitute endorsement by the American College of Sports Medicine (ACSM).

ABSTRACT

PURPOSE: To enable inter- and intra-study comparisons it is important to ascertain comparability among accelerometer models. This study compared raw and count data between hip-worn ActiGraph GT3X+ and GT9X Link accelerometers. **METHODS**: Adults (n=26 [n=15 women]; aged 49.1±20.0 years) wore GT3X+ and Link accelerometers over the right hip for an 80-min protocol involving 12-21 sedentary, household, and ambulatory/exercise activities lasting 2-15 min each. For each accelerometer, mean and variance of the raw (60 Hz) data for each axis and vector magnitude (VM) were extracted in 30-s epochs. A machine learning model (Montoye 2015) was used to predict energy expenditure in METs from the raw data. Raw data were also processed into activity counts in 30-s epochs for each axis and VM, with Freedson 1998 and 2011 count-based regression models used to predict METs. Time spent in sedentary, light, moderate, and vigorous intensities were derived from predicted METs from each model. Correlations were calculated to compare raw and count data between accelerometers, and percent (%) agreement was used to compare epoch-by-epoch activity intensity. **RESULTS**: For raw data, correlations for mean acceleration were 0.96±0.05, 0.89±0.16, 0.71±0.33, and 0.80±0.28 and for variance 0.98±0.02, 0.98±0.03, 0.91±0.06, and 1.00±0.00 in the X, Y, and Z axes and VM, respectively. For count data, corresponding correlations were 1.00±0.01, 0.98±0.02, 0.96±0.04, and 1.00±0.00, respectively. Freedson 1998 and 2011 count-based models had significantly higher %agreement for activity intensity (95.1±5.6% and 95.5±4.0%) than the Montoye 2015 raw data model (61.5±27.6%; p<0.001). **CONCLUSIONS**: Count data were more highly comparable than raw data between accelerometers. Data filtering and/or more robust raw data models are needed to improve raw data comparability between ActiGraph GT3X+ and Link accelerometers.

Keywords: reliability, activity monitor, agreement, physical activity, energy expenditure

INTRODUCTION

Over the last few decades, accelerometer-based activity monitors have become a popular field-based physical activity (PA) assessment method. Despite some limitations, there is general agreement that accelerometry can provide an appealing blend of validity and feasibility for measuring free-living PA and sedentary behavior across various settings. ActiGraph (ActiGraph Corp., Pensacola, FL), formerly known as Computer Science and Applications (CSA) and Manufacture Technology Incorporated (MTI), is one of the most commonly used accelerometer brands in intervention and surveillance research and has produced several generations over the past few decades, most recently the GT9X Link in 2014 (1). As technology has evolved, ActiGraph accelerometers have progressed from piezoelectric to capacitive sensors, uniaxial to triaxial data capture, and epoch-based count data to raw data sampled at high frequencies (i.e., 30-100 Hz; GT3X+ and Link). Raw acceleration data may contain valuable movement information such as direction and orientation, which are lost when data are condensed into activity counts; this prospect has led some researchers to use raw data in an attempt to improve the measurement of PA and sedentary behavior (2-6).

Previous studies have generally shown good comparability of vertical axis counts from the 7164 and GT1M accelerometers both in laboratory-based mechanical shaker testing and when worn on the hip (7-10); moreover, the majority of past work supports the comparability of the GT1M and GT3X vertical axis counts when worn by adults in both laboratory and free-living settings (11-13). However, Sasaki et al. (14) demonstrated good comparability in the vertical axis but poor comparability in the anterior-posterior axis between GT1M (anterior-posterior axis of GT1M unlocked by ActiGraph for use several years after initial production) and GT3X, thereby recommending caution when drawing comparisons between more than the vertical axis data from different ActiGraph models. Despite a wealth of studies comparing older ActiGraph models, little work has established comparability of data collected from the GT3X+ or the Link to older models or, indeed, to each other.

The newest ActiGraph model, the GT9X Link, differs from the GT3X+ model in several important ways, many of which affect usability. For example, unlike the GT3X+, the Link has a screen that allows researchers to select whether wearer PA feedback (e.g., steps) is provided and contains other measurement technologies including a triaxial magnetometer and gyroscope. Of greater relevance to the current study, the GT3X+ is larger (4.6×3.3×1.5 cm vs. 3.5×3.5×1.0 cm), heavier (19 g vs. 14 g), and has slightly different methods of attachment to the hip than the Link (slots for elastic band on GT3X+ vs. clip for Link); these attributes could subtly change the device orientation when worn, which may directly impact the raw acceleration output from each device. Although the accelerometer type, sampling frequency, and data filtering methods are reportedly the same between the GT3X+ accelerometer and the Link's primary accelerometer, to our knowledge the comparability of these accelerometer models has not been assessed. It is important to establish if data collected from GT3X+ and Link models are comparable when evaluating studies or interventions using these different accelerometer models. Therefore, the purpose of this study was to assess comparability of the ActiGraph GT3X+ and Link accelerometers for raw and count data and for predictions of energy expenditure (EE) and activity intensity during human testing.

METHODS

Participants

Participants (n=30) were apparently healthy adults aged 18-79 years who were able to safely participate in activities of at least a moderate intensity (i.e., able to complete a brisk walk). The study protocol was approved by Ball State University's Institutional Review Board, where data collection took place, and all participants provided written informed consent prior to beginning the study.

Equipment

Participants wore two ActiGraph accelerometers (one GT3X+, one GT9X Link), which were initialized at a sampling rate of 60 Hz; previous work by Brønd et al. (15) indicate that, for correct conversion of raw data to counts, the sampling frequency must be a multiple of 30 Hz (e.g., 30, 60, or 90 Hz). Accelerometers were placed immediately next to each other, with placement of accelerometers randomized among participants, on the right hip using an elastic belt. The GT3X+ has two slots into which the elastic band feeds; conversely, the Link comes with a separate clip which can be attached to the elastic band. Accelerometers were oriented so that the X axis was primarily the anterior-posterior axis, the Y axis was the vertical axis, and the Z axis was primarily the medial-lateral axis (see Figure, Supplemental Digital Content 1, Link and GT3X monitor orientation, http://links.lww.com/MSS/B166). A total of seven GT3X+ and four Link devices were used in testing. The GT3X+ devices can capture a dynamic range of ±6g, whereas the Link devices have a range of ±8g.

Protocol

Once fitted with the accelerometers, participants completed an 80-min, semi-structured protocol in which they performed 12-21 activities from a pre-defined list of 21 (see Table 1 for activities performed and number of total epochs for which activities were performed). Participants were given considerable freedom in the protocol; duration, intensity, order, and method of performing each activity (e.g., how to hold a broom, walking speeds, etc.) were left up to the participants. Activities were broadly grouped into three categories (sedentary, household, and ambulatory/exercise), and participants were required to perform at least four activities in each category. Chosen activities were performed for 2-15 min each; additionally, participants were asked to spend ≥40 min (≥50%) of their time performing sedentary behaviors to better replicate adults' free-living behaviors (16). During the protocol, a trained research assistant recorded exact start and end times for each activity on a tablet computer which was time-synchronized with the accelerometers; these times were used to conduct activity-specific accelerometer comparisons.

Data processing and analysis

Data were downloaded and converted to counts in ActiLife version 6.9 (ActiGraph Corp., Pensacola, FL); the raw data were then exported to Microsoft Excel 2010 (Microsoft Corp., Redmond, WA) for feature extraction and conversion to 30-s epochs.

In terms of accelerometer orientation (see Figure, Supplemental Digital Content 1, Link and GT3X monitor orientation, http://links.lww.com/MSS/B166), the Y axis is oriented in the same orientation for both accelerometers, but the X and Z axes are reversed in sign; therefore, the signs (+ or -) for the X and Z axes were reversed for the Link accelerometer when making raw data comparisons to the GT3X+ accelerometer and before extracting features for the machine learning model (as described below).

For all analyses (except Bland-Altman plots) data were calculated for each individual participant and averaged across the sample. Several statistical tests were completed to analyze comparability of data from the Link and GT3X+ accelerometers. For the raw data, the mean and variance of the raw signal in each accelerometer axis and vector magnitude (VM; $VM = \sqrt{X^2 + Y^2 + Z^2}$) were computed in 30-s epochs) and were compared on an epoch-by-epoch basis using Pearson correlations; a similar analysis was conducted for counts accrued in 30-s epochs in each axis and VM. Correlations between devices were computed for each individual participant, and data were normalized using a Fisher Z transform to normalize data prior to statistical analyses.

The 30-s epoch length was chosen for consistency with previously-developed machine learning predictions of activity intensity (5, 17). Repeated measures analysis of variance tests were used to compare correlations for each axis and VM among the raw and count data; in the case of a significant test statistic, a least significant difference correction was used for *post hoc* pairwise comparisons.

Three previously-developed EE prediction equations/cut-points were used to predict the activity intensity (sedentary, light, moderate, and vigorous) of each 30-s epoch, and the total time spent in each intensity was summed for the entire visit. For the raw data, mean and variance of the raw signal in each axis ([2 features·axis^{-1}]·3 axes = 6 features) were input into a previously developed, machine learning EE (in METs) prediction equation; this equation was developed for the ActiGraph GT3X+ accelerometer by Montoye et al. (3) in 2015 (subsequently referred to as Montoye 2015 model). This was done separately for the raw data from each accelerometer. Standard absolute MET thresholds (sedentary: ≤1.5; light: >1.5 and <3.0; moderate: ≥3.0 and <6.0; vigorous: ≥6.0) were then used to determine activity intensity from the EE predictions.

For the count data, 30-s data were input into the Freedson 1998 and 2011 count-based linear regression equations (14, 18), which were used to determine epoch-by-epoch and total time spent in each activity intensity. It should be noted that neither of these cut-points originally differentiated between sedentary behavior and light-intensity PA, but we included cut-points of 100 counts·min^{-1} and 200 counts·min^{-1} for the Freedson 1998 and Freedson 2011 cut-points, respectively, as these are commonly used for determining sedentary behavior (19, 20). Specifically, the Freedson 1998 equation only uses the Y (vertical) axis (sedentary: 0-100; light: 101-1951; moderate: 1952-5724; vigorous: >5724 in counts·min^{-1}), whereas the Freedson 2011 equation uses VM of the accelerometer axes (sedentary: 0-200; light: 201-2690; moderate: 2691-6166; vigorous: >6166 counts·min^{-1}).

Independently for the raw and count data, paired t-tests were used to determine whether statistically significant differences in time spent in each activity intensity existed between GT3X+ and the Link accelerometers. Additionally, both percent (%) agreement and kappa statistics were computed and compared between raw and count data using paired t-tests.

Activity-specific EE values, mean and variance of the raw data, and count data were also compared between accelerometers using paired t-tests. Statistical significance was determined at a p-value of $p<0.05$ for the repeated measures analysis of variance tests and paired t-tests.

Finally, Bland-Altman plots were developed using pooled data from all participants to assess agreement between accelerometers for the raw and count data. The Montoye 2015 model for EE prediction was employed using R statistical software (R-project, Vienna, Austria), and the Freedson 1998 and 2011 models were employed using ActiLife version 6.9. Statistical testing was conducted using SPSS version 24.0 (IBM Corp., Armonk, NY).

RESULTS

Four participants had accelerometer initialization or malfunction issues (one Link malfunction, one instance where accelerometers were initialized to the incorrect time, and two instances where Link was placed in the incorrect orientation). Their data were excluded from analysis, leaving 26 (n=15 female) with data available for analysis. Mean ± standard deviation was 49.1±20.0 years for age (males: 50.7±21.6 years; females: 47.9±19.5 years), 173.5±9.2 cm for height (males: 181.8±5.9 cm; females: 167.5±5.6 cm), 78.3±15.8 kg for mass (males: 88.8±11.8 kg; females: 70.6±14.1 kg), and 25.9±4.3 kg·m^{-2} for body mass index (males: 26.8±2.8 kg·m^{-2}; females: 25.2±5.1 kg·m^{-2}).

Since raw acceleration data may not be perfectly calibrated to gravity, we attempted to use the g.calibrate function of the GGIR package in R to calibrate the data. However, of the 52 data files (26 participants × 2 devices), only 22 files from 11 participants had enough data to assess calibration; for those, GGIR reported calibration errors ranging from 0.01081321-0.04571884 (average: 0.022458135) for the GT3X+ and 0.01878736-0.04958797 (average: 0.028124825) for the Link. None of the devices had enough data to recalibrate the devices; therefore, all data were left uncalibrated for analyses.

Correlations between devices

Figure 1 displays Pearson correlations for both the raw and count data. Both the count data (r=0.96-1.00) and variance of the raw data (r=0.91-1.00) had significantly higher correlations between accelerometers for each axis and VM compared to the mean of the raw data (r=0.71-0.96). Moreover, the count data had significantly higher correlations for the X and Z axes than the variance of the raw signal. With the exception of the mean raw data for the Z axis (r=0.71), all correlations found in this study would be considered "strong" (≥0.80) according to Cohen's correlation strength thresholds (21).

Activity-specific, raw and count data comparisons

Activity-specific comparisons of raw and count data are shown in Tables 1 and 2. Small but statistically significant differences between accelerometers were observed for mean raw data (Table 1) for Y and Z axes as well as the VM with all 6 sedentary behaviors for the Y axis and VM and all but one sedentary behavior (lying down) for the Z axis. Small but statistically significant differences were also shown for the Y and Z axes and VM for most household activities (5, 5, and 7 of the 8 household activities, respectively) and X axis, Y axis and VM for most ambulatory/exercise activities (4, 5, and all 7 of the 7 ambulatory/exercise activities, respectively). For the Y axis, the GT3X+ accelerometer had a greater positive signal magnitude in 13 of the 16 different activities deemed statistically significant, and for the Z axis, the GT3X+ accelerometer had a greater positive magnitude in all 11 of the activities deemed statistically significant; no such trend was present for the X axis. The X axis had significant between-accelerometer differences for 7 activities (reading, picking up items, vacuuming, walking briskly, over-ground and treadmill jogging, and stairs).

For variance of the raw data (Table 1), the X axis, Y axis, and VM were significantly different between accelerometers for most activities (14, 14, and 18 of the 21 activities, respectively). Additionally, variance in the Z axis was significantly different between accelerometers for 4 of the 7 ambulatory/exercise activities but only 1 household activities and 2 sedentary behaviors. In most cases, the detected accelerations for variance were <0.01g different between accelerometers. For count data (Table 2), the X axis data were significantly different between accelerometers for the majority of household (5 out of 8) and ambulatory/exercise activities (4 out of 7) as well as computer use, whereas the Y axis and VM data were significantly different between accelerometers in only a few activities interspersed among categories (5 and 7 of the 21 activities, respectively). The between-accelerometer Z axis data were significantly different for most ambulatory/exercise activities (4 out of 7) and a few sporadic sedentary behaviors (3 out of 6) and household (2 out of 8) activities. The largest differences in count data occurred during the ambulatory/exercise activities, with a mean difference as large as 472.6 Z axis counts during the treadmill jog activity.

Activity-specific EE predictions are shown in Table 3. For the sedentary behaviors, there were few inter-accelerometer differences in predicted EE between the GT3X+ and Link accelerometers. Statistically significant differences were present for several of the household activities using the Freedson 1998 and 2011 equations (sweeping and vacuuming for both, picking up items for Freedson 1998), but the absolute magnitude of the difference was small

(0.1-0.2 METs). There were significant differences between the GT3X+ and Link in predicted METs (ranging 0.1-0.9 METs) for the Montoye 2015 model for most ambulatory/exercise activities (6 out of 7), but few differences existed between accelerometers for the Freedson 1998 (2 out of 7) or 2011 (1 out of 7) equations.

Comparison of predicted physical activity intensities

Predictions of time spent in each activity intensity as well as % agreement and kappa coefficients for activity intensity classification are shown in Table 4. The Montoye 2015 model predicted significantly less time spent in light-intensity PA but significantly more time spent in moderate-, vigorous-, and moderate- or vigorous-intensity PA with the Link accelerometer compared to the GT3X+ accelerometer. Conversely, for the count data, there were no differences in the predicted times spent in any intensity for the Freedson 1998 or 2011 equations. For % agreement and kappa coefficients, the Montoye 2015 model had significantly lower inter-accelerometer agreement than the Freedson equations. Additionally, using kappa strength thresholds developed by Altman (22), the Montoye 2015 model fell in the low end of "moderate" agreement (0.41-0.60), whereas the Freedson 1998 and 2011 equations had "very good" agreement (>0.80).

Agreement between devices

In order to further examine agreement between the GT3X+ and Link data, Bland-Altman plots were constructed (Figure 2). For both raw and count data, mean differences between accelerometers clustered close to 0, indicating good overall agreement between devices. For raw data, 95% limits of agreement (in g) for the mean raw data for the X axis, Y axis, Z axis, and VM, respectively were [-0.29, 0.29], [-0.17, 0.21], [-0.34, 0.48], and [-0.08, 0.04] and for variance of the raw data were [-0.07, 0.06], [-0.04, 0.05], [-0.05, 0.05], and [-0.08, 0.14]. For variance of the raw data, differences in variance between devices appear to become more pronounced with increasing signal magnitude. For count data, 95% limits of agreement (in counts) were [-207.84, 207.27], [-290.27, 290.74], [-426.67, 411.12], and [-174.93, 170.60] for the X axis, Y axis, Z axis, and VM, respectively. Translated into percentages, these limits of agreement (for X, Y, Z, and VM, respectively) were [-138.57%, 138.57%], [-19.69%, 24.32%], [-261.86%, 369.69%], and [-7.73%, 3.87%] for means of the raw data; [-259.19%, 222.17%], [-77.15%, 96.44%], [-297.38%, 297.38%], and [-159.73%, 279.52%] for variances of the raw data; and [-44.14%, 44.02%], [-86.49%, 86.63%], [-140.44%, 135.33%], and [-24.85%, 24.24%] for count data.

DISCUSSION

The purposes of the present study were to ascertain the comparability of raw and count data collected from GT3X+ and GT9X Link accelerometers worn on the hip and, secondly, to assess the comparability of activity intensity predictions from these accelerometers when assessed using previously validated EE/activity intensity prediction equations/models. Results indicated that the raw data were less comparable than the count data between accelerometers. Additionally, when features of the raw data were used as inputs into the Montoye 2015 prediction model (3), the two accelerometers produced significantly different estimates of time spent in all activity intensity categories except sedentary behavior. Conversely, count data were more comparable than raw data between accelerometers and yielded similar estimates of activity intensity when the count data were used as inputs into two popularly used prediction equations, the Freedson 1998 and 2011 equations (14, 18).

Comparability of count data and raw data

Given that most studies seeking to evaluate PA using ActiGraph accelerometers still rely on count data (1), the high comparability between the GT3X+ and Link accelerometers for count data is encouraging. Importantly, the comparability of activity intensity estimates derived from both the count Freedson 1998 and 2011 equations (14, 18) provides evidence to support the use of either vertical-axis counts or VM to predict EE and activity intensity interchangeably between the GT3X+ and Link accelerometers when worn on the hip. This is a reassuring finding for researchers using more than one version of ActiGraph accelerometer along with count-based regression or machine learning data models for EE and/or activity type prediction, such as those developed by Freedson et al. (14, 18), Lyden et al. (23), Trost et al. (24), and Mackintosh et al. (25). However, if using count data it is important to note that the sampling frequency of raw ActiGraph data affects the conversion to counts (15) and, thus, the sampling rate chosen for a study would need to match the sampling frequency used to validate the models if optimal accuracy is to be obtained. An additional limitation of count data is that counts are brand-specific and often proprietary, precluding inter-brand comparability or usability of predictive models (26).

Despite the limitations of activity counts, the high comparability of activity count data between devices in this study may provide insight into potential ways to process, filter, and/or summarize raw data to improve raw data comparability within and among accelerometer brands. Determination of transparent, replicable steps to improve raw data comparability across different accelerometers is important as it would offer the best potential for pooling raw data from multiple studies, comparing across studies, and developing highly accurate prediction models across different populations and a variety of accelerometer brands. It is not yet clear if this is possible, however, with limited available evidence providing mixed support of the comparability of raw data across brands. For example, John et al. (27) compared raw data collected from GENEActiv and ActiGraph GT3X+ data and found significantly higher VM for the GENEActiv compared to the GT3X+ during mechanical shaker testing; these differences persisted in human testing, where activity type recognition accuracy was significantly decreased when developing a predictive model for one device and applying it to raw data from the other device. In two subsequent studies comparing the GENEActiv to the ActiGraph GT3X+, Rowlands et al. (28) and Hildebrand et al. (29) found significant differences in raw data between GENEActiv and ActiGraph GT3X+ when worn on either the wrist and hip. However, these studies have predominately evaluated raw data and/or time-domain features (similar to those used in this study) when evaluating comparability. Importantly, John et al. (27) found better comparability of frequency- than time-domain features between GENEActiv and ActiGraph accelerometer brands. Rowlands et al. (30) also found that some frequency-domain features were translatable across GENEActiv and ActiGraph brands. However, whether frequency-domain features can solely provide sufficient information for accurate activity prediction remains unknown. These and other aspects, such as accelerometer angle, should be explored in future work given their potential to be used for improving translatability within and across accelerometer brands and in the accurate prediction of physical activity.

Accelerometer orientation

The presence of subtle differences between accelerometer brands and within models of the same brand, and the subsequent influence of these differences on activity prediction, has important implications for inter-study comparisons and interpretations. The sedentary behaviors in this study elicited a predicted EE of 1.1-1.7 METs for both accelerometers. Given the limited hip motion for these activities (as evidenced by variances close to 0), this indicates that mean of the raw acceleration signal, which is orientation-dependent (see Figure, Supplemental Digital

Content 1, Link and GT3X monitor orientation, http://links.lww.com/MSS/B166), affected EE estimates for these activities. Therefore, the machine learning model used in this study is also orientation-dependent. While standardization of accelerometer placement should be routine practice in studies utilizing accelerometers, the present study indicates that even small variations in accelerometer placement and/or orientation, which inevitably occur when wearing an accelerometer over a sustained period and which may be inherent in different methods of attachment used (e.g., tape vs. strap vs. clip), may affect accuracy of prediction models which use raw data.

Use of VM or ENMO would avoid orientation issues since they summarize triaxial data into a single, non-negative value. However, similar to count data, it is thought that collapsing raw data into activity counts or metrics, such as VM or ENMO, loses some of the information contained in the raw signal and, further, that this extra information may help improve the accuracy of predictive models. This especially impacts accelerometers that are placed further from the center of mass, have higher movement variability (such as the wrist), and/or are used to measure EE in children, whose movements are more irregular, shorter in duration, and of higher intensity than adults (2, 6, 31). A previous study by Montoye et al. (32) found that machine learning models using triaxial data had higher accuracy in a simulated free-living setting than machine learning models using VM for wrist-worn accelerometers. Therefore, a current challenge is determining how to effectively utilize the extra information provided by using triaxial, raw data to maximize EE prediction accuracy while minimizing the effect of orientation differences and/or potential overfitting of predictive models. One potential approach would be to use filtering methods to remove noise from the raw accelerometer data; such filtering is routinely conducted when translating raw ActiGraph data to activity counts (15). Several different kinds of filtering techniques are available (e.g., Butterworth, band-pass), and each has parameters that must be predetermined. Studies suggest that filtering can be utilized/included as part of using raw data to assess physical activity (33, 34); however, currently there are not sufficient data to recommend a certain kind of filtering technique or parameters for a given technique to allow standardization.

Strengths and limitations

This study has several important strengths. The simulated free-living setting allowed for a thorough examination of the accelerometers' raw and count data comparability across a wide variety of activities and activity types (i.e., sedentary, lifestyle, ambulatory) while utilizing high-quality criterion measures to capture the ground truth/criterion activity types and timings. Therefore, our findings likely generalize well due to similarity in our protocol to typical accelerometer use practices. Additionally, although relatively small in number, our sample encompassed a wide age range and had considerable variability in fitness level and body composition, increasing the likelihood of these results being generalizable to other adult populations. Finally, the use of a machine learning model for raw data and two commonly used, count-based models for EE prediction allowed for determination of potential practical implications of differences in the accelerometer data on physical activity outcomes.

This study also had several noteworthy limitations. First, this study utilized only one machine learning model available for assessment of EE from raw accelerometer data, and other models may be affected differently by the small differences in the raw acceleration signal. Another limitation is that this study did not have access to a mechanical shaker to assess comparability of the raw data between accelerometers. Given the differences found in the raw signal when the devices were worn side-by-side, a mechanical shaker study may help elicit further differences in raw data between devices (i.e., frequencies, magnitudes, orientations). On a related note, while the randomization of the accelerometers in this study should account for any potential influence of accelerometer placement on the data, inter- or intra-instrument variability of the devices may, at least in part, contribute to the slight variation seen in the raw data collected from the GT3X+ and Link accelerometers. We are not aware of studies assessing reliability of raw data collected by ActiGraph accelerometers, although it seems likely that the raw data are highly reliable given that numerous studies have shown high intra- and inter-instrument reliability of various models of the ActiGraph accelerometer for count data (especially at lower intensities) (35-39). An additional limitation is that the short length of data collection (~2 hours) did not allow for use of functions to calibrate sensor data for gravity, as is possible in the GGIR package in R (40); slight calibration errors presented in the results may have contributed to the differences seen between devices.

Conclusions

In conclusion, our study found good comparability between hip-worn ActiGraph GT3X+ and GT9X Link accelerometers for activity counts (all three axes and VM) and count-based activity intensity prediction. Therefore, it appears that hip-worn GT3X+ and Link accelerometers can be used interchangeably for collection of count data. Conversely, small but significant differences were uncovered in the raw signal between accelerometers, and these translated to significantly different predictions of time spent in different activity intensities. Studies utilizing raw data should be cognizant of potential differences between the GT3X+ and Link and consider data filtering or alternate modeling techniques that may improve comparability of raw data between accelerometers.

ACKNOWLEDGMENTS

This study was supported by a Ball State University ASPiRE grant.

CONFLICT OF INTEREST

The authors declare that they have no conflicts of interest to report. The results of this study are presented clearly, honestly, and without fabrication, falsification, or inappropriate data manipulation. The findings of the present study do not constitute endorsement by the American College of Sports Medicine (ACSM).

REFERENCES

1. Montoye AHK, Moore RW, Bowles HR, Korycinski R, Pfeiffer KA. Reporting accelerometer methods in physical activity intervention studies: A systematic review and recommendations for authors. Br J Sports Med. 2016. doi: 10.1136/bjsports-2015-095947.

2. Staudenmayer J, He S, Hickey A, Sasaki J, Freedson P. Methods to estimate aspects of physical activity and sedentary behavior from high-frequency wrist accelerometer measurements. J Appl Physiol. 2015; 119(4): 396-403. doi: 10.1152/japplphysiol.00026.2015.

3. Montoye AHK, Mudd LM, Biswas S, Pfeiffer KA. Energy expenditure prediction using raw accelerometer data in simulated free living. Med Sci Sports Exerc. 2015; 47(8): 1735-46. doi: 10.1249/MSS.0000000000000597.

4. Montoye AHK, Pivarnik JM, Mudd LM, Biswas S, Pfeiffer KA. Comparison of activity type classification accuracy from accelerometers worn on the wrists, hip, and thigh. Meas Phys Educ Exerc Sci. 2016; 20(3): 173-83. doi: 10.1080/1091367X.2016.1192038.

5. Montoye AHK, Pivarnik JM, Mudd LM, Biswas S, Pfeiffer KA. Validation and comparison of accelerometers worn on the hip, thigh, and wrists for measuring physical activity and sedentary behavior. AIMS Public Health. 2016; 3(2): 298-312. doi: 10.3934/publichealth.2016.2.298.

6. Ellis K, Kerr J, Godbole S, Staudenmayer J, Lanckriet G. Hip and wrist accelerometer algorithms for free-living behavior classification. Med Sci Sports Exerc. 2016; 48(5): 933-40. doi: 10.1249/MSS.0000000000000840.

7. Kozey SL, Staudenmayer JW, Troiano RP, Freedson PS. Comparison of the ActiGraph 7164 and the ActiGraph GT1M during self-paced locomotion. Med Sci Sports Exerc. 2010; 42(5): 971-6. doi: 10.1249/MSS.0b013e3181c29e90.

8. Corder K, Brage S, Ramachandran A, Snehalatha C, Wareham N, Ekelund U. Comparison of two actigraph models for assessing free-living physical activity in Indian adolescents. J Sports Sci. 2007; 25(14): 1607-11. doi: 10.1080/02640410701283841.

9. John D, Tyo B, Bassett DR. Comparison of four ActiGraph accelerometers during walking and running. Med Sci Sports Exerc. 2010; 42(2): 368-74. doi: 10.1249/MSS.0b013e3181b3af49.

10. Rothney MP, Apker GA, Song Y, Chen KY. Comparing the performance of three generations of ActiGraph accelerometers. J Appl Physiol. 2008; 105(4): 1091-7. doi: 10.1152/japplphysiol.90641.2008.

11. Vanhelst J, Mikulovic J, Bui-Xuan G, et al. Comparison of two ActiGraph accelerometer generations in the assessment of physical activity in free living conditions. BMC Res Notes. 2012; 5: 187. doi: 10.1186/1756-0500-5-187.

12. Santos-Lozano A, Santin-Medeiros F, Cristi-Montero C, Jaen-Jimenez R, Casajus JA, Garatachea N. GT1M, GT3X, and Actitrainer counts comparison during standardized activities in young, adults and older adults. Nutr Hosp. 2016; 33(3): 692-7.

13. Kaminsky LA, Ozemek C. A comparison of the ActiGraph GT1M and GT3X accelerometers under standardized and free-living conditions. Physiol Meas. 2012; 33(11): 1869-76. doi: 10.1088/0967-3334/33/11/1869.

14. Sasaki JE, John D, Freedson PS. Validation and comparison of ActiGraph activity monitors. J Sci Med Sport. 2011; 14(5): 411-6. doi: 10.1016/j.jsams.2011.04.003.

15. Brond JC, Arvidsson D. Sampling frequency affects the processing of ActiGraph raw acceleration data to activity counts. J Appl Physiol. 2016; 120(3): 362-9. doi: 10.1152/japplphysiol.00628.2015.

16. Donaldson SC, Montoye AH, Tuttle MS, Kaminsky LA. Variability of objectively measured sedentary behavior. Med Sci Sports Exerc. 2016; 48(4): 755-61. doi: 10.1249/MSS.0000000000000828.

17. Montoye AHK, Conger SA, Connolly CP, Imboden MT, Nelson MB, Bock JM, Kaminsky LA. Validation of accelerometer-based energy expenditure prediction models in structured and simulated free-living settings. Meas Phys Educ Exerc Sci. 2017; 21(4): 223-234. doi: 10.1080/1091367X.2017.1337638.

18. Freedson PS, Melanson E, Sirard J. Calibration of the Computer Science and Applications, Inc. accelerometer. Med Sci Sports Exerc. 1998; 30(5): 777-81.

19. Treuth MS, Schmitz K, Catellier DJ, et al. Defining accelerometer thresholds for activity intensities in adolescent girls. Med Sci Sports Exerc. 2004; 36(7): 1259-66. doi: 00005768-200407000-00026 [pii].

20. Aguilar-Farias N, Brown WJ, Peeters GM. ActiGraph GT3X+ cut-points for identifying sedentary behaviour in older adults in free-living environments. J Sci Med Sport. 2014; 17(3): 293-9. doi: 10.1016/j.jsams.2013.07.002.

21. Cohen J. Statistical Power Analysis for the Behavioral Sciences. New York, Academic Press; 1977. 109-43 p.

22. Altman DG. Practical Statistics for Medical Research. London, UK., Chapman and Hall; 1991. 403-8 p.

23. Lyden K, Keadle SK, Staudenmayer J, Freedson PS. A method to estimate free-living active and sedentary behavior from an accelerometer. Med Sci Sports Exerc. 2014; 46(2): 386-97. doi: 10.1249/MSS.0b013e3182a42a2d.

24. Trost SG, Wong WK, Pfeiffer KA, Zheng Y. Artificial neural networks to predict activity type and energy expenditure in youth. Med Sci Sports Exerc. 2012; 44(9): 1801-9. doi: 10.1249/MSS.0b013e318258ac11.

25. Mackintosh KA, Montoye AHK, Pfeiffer KA, McNarry M. Investigating optimal accelerometer placement for energy expenditure prediction in children using a machine learning approach. Physiol Meas. 2016; 37(10): 1728-40. doi: 10.1088/0967-3334/37/10/1728.

26. Paul DR, Kramer M, Moshfegh AJ, Baer DJ, Rumpler WV. Comparison of two different physical activity monitors. BMC Med Res Methodol. 2007; 7: 26. doi: 10.1186/1471-2288-7-26.

27. John D, Sasaki J, Staudenmayer J, Mavilia M, Freedson PS. Comparison of raw acceleration from the GENEA and ActiGraph GT3X+ activity monitors. Sensors. 2013; 13(11): 14754-63. doi: 10.3390/s131114754.

28. Rowlands AV, Yates T, Davies M, Khunti K, Edwardson CL. Raw accelerometer data analysis with GGIR R-package: Does accelerometer brand matter? Med Sci Sports Exerc. 2016; 48(10): 1935-41. doi: 10.1249/MSS.0000000000000978.

29. Hildebrand M, Van Hees VT, Hansen BH, Ekelund U. Age group comparability of raw accelerometer output from wrist- and hip-worn monitors. Med Sci Sports Exerc. 2014; 46(9): 1816-24. doi: 10.1249/MSS.0000000000000289.

30. Rowlands AV, Fraysse F, Catt M, Stiles VH, Stanley RM, Eston RG, Olds TS. Comparability of measured acceleration from accelerometry-based activity monitors. Med Sci Sports Exerc. 2015; 47(1): 201-10. doi: 10.1249/MSS.0000000000000394.

31. Bailey RC, Olson J, Pepper SL, Porszasz J, Barstow TJ, Cooper DM. The level and tempo of children's physical activities: An observational study. Med Sci Sports Exerc. 1995; 27(7): 1033-41.

32. Montoye AHK, Pivarnik JM, Mudd LM, Biswas S, Pfeiffer KA. Wrist-independent energy expenditure prediction models from raw accelerometer data. Physiol Meas. 2016; 37(10): 1770-84. doi: 10.1088/0967-3334/37/10/1770.

33. Oshima Y, Kawaguchi K, Tanaka S, et al. Classifying household and locomotive activities using a triaxial accelerometer. Gait Posture. 2010; 31(3): 370-4. doi: 10.1016/j.gaitpost.2010.01.005.

34. Charry E, Lai DT, Begg RK, Palaniswami M. A study on band-pass filtering for calculating foot displacements from accelerometer and gyroscope sensors. Conf Proc IEEE Eng Med Biol Soc. 2009; 2009: 4824, 6-7. doi: 10.1109/IEMBS.2009.5332673.

35. Aadland E, Ylvisaker E. Reliability of the ActiGraph GT3X+ accelerometer in adults under free-living conditions. PLoS One. 2015; 10(8): e0134606. doi: 10.1371/journal.pone.0134606.

36. Jarrett H, Fitzgerald L, Routen AC. Interinstrument reliability of the ActiGraph GT3X+ ambulatory activity monitor during free-living conditions in adults. J Phys Act Health. 2015; 12(3): 382-7. doi: 10.1123/jpah.2013-0070.

37. Vanhelst J, Baquet G, Gottrand F, Beghin L. Comparative interinstrument reliability of uniaxial and triaxial accelerometers in free-living conditions. Percept Mot Skills. 2012; 114(2): 584-94. doi: 10.2466/03.26.PMS.114.2.584-594

38. Sirard JR, Forsyth A, Oakes JM, Schmitz KH. Accelerometer test-retest reliability by data processing algorithms: Results from the Twin Cities Walking Study. J Phys Act Health. 2011; 8(5): 668-74.

39. Ozemek C, Kirschner MM, Wilkerson BS, Byun W, Kaminsky LA. Intermonitor reliability of the GT3X+ accelerometer at hip, wrist and ankle sites during activities of daily living. Physiol Meas. 2014; 35(2): 129-38. doi: 10.1088/0967-3334/35/2/129.

40. van Hees VT, Fang Z, Langford J, Assah F, Mohammad A, da Silva IC, Trenell MI, White T, Wareham NJ, Brage S. Autocalibration of accelerometer data for free-living physical activity assessment using local gravity and temperature: An evaluation on four continents. J Appl Physiol. 2014; 117(7): 738-44. doi: 10.1152/japplphysiol.00421.2014.

FIGURE LEGENDS

Figure 1. Pearson correlations for raw and count data collected from GT3X+ and Link accelerometers.

Error bars represent standard deviation.

$^{\$}$Indicates significant differences between raw – mean and raw – variance data.

*Indicates significant difference between raw – mean and count data.

^Indicates significant difference between raw – variance and count data.

VM: vector magnitude.

Figure 2. Bland-Altman plots for raw and count data in each axis and for vector magnitude.

Black dotted lines represent average difference, and gray dotted lines represent 95% limits of agreement of mean difference.

Axis units for plots a-h are in gravitational (g) units.

Axis units for plots i-l are in counts.

VM: vector magnitude.

a. Plot for X axis raw data – mean.
b. Plot for Y axis raw data – mean.
c. Plot for Z axis raw data – mean.
d. Plot for VM raw data – mean.
e. Plot for X axis raw data – variance.
f. Plot for Y axis raw data – variance.
g. Plot for Z axis raw data – variance.
h. Plot for VM raw data – variance.

i. Plot for X axis count data.

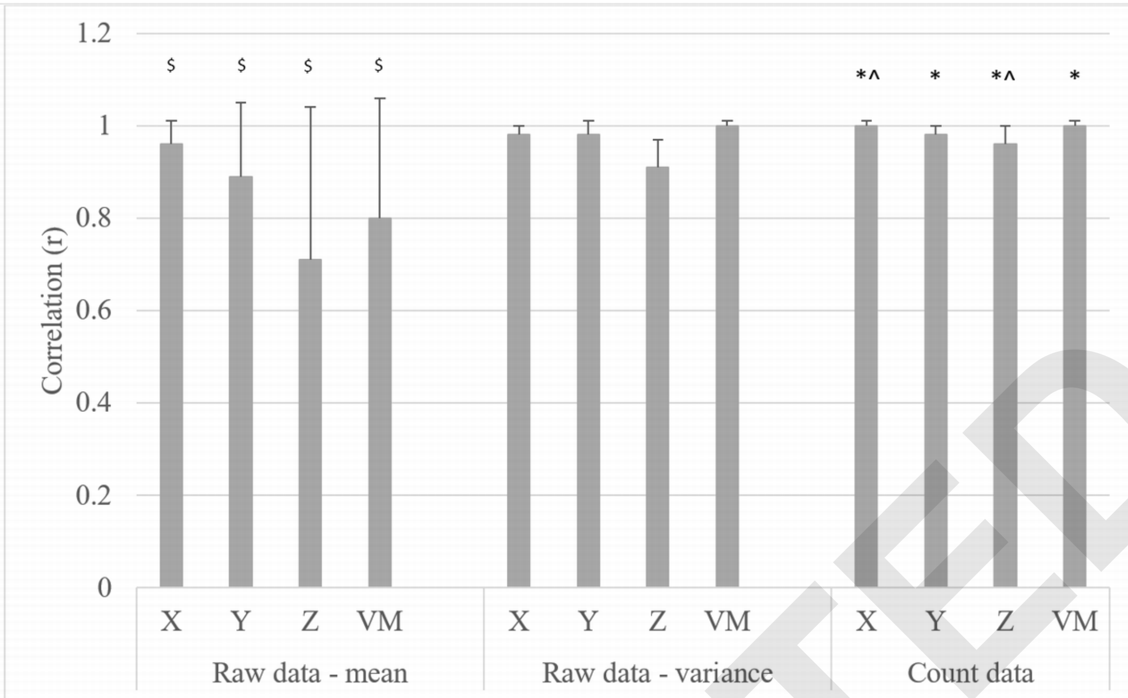
j. Plot for Y axis count data.

k. Plot for Z axis count data.

l. Plot for VM count data.

SUPPLEMENTAL DIGITAL CONTENT

SDC 1. Link and GT3X monitor orientation. with title.tiff


1.2
1
0.8
0.6
0.4
0.2
0
Correlation (r)
$ $ $ $
*^ * *^ *
X Y Z VM
X Y Z VM
X Y Z VM
Raw data - mean
Raw data - variance
Count data

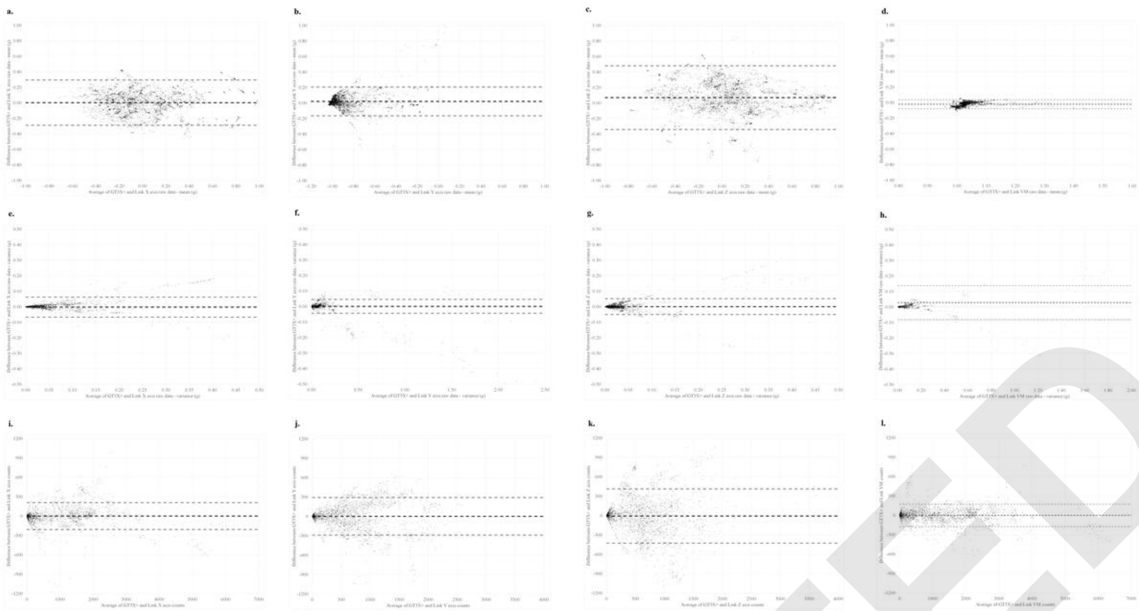


Table 1. Activity-specific, raw data for ActiGraph GT3X+ and Link.

	GT3X+ X	Link X^	GT3X+ Y	Link Y	GT3X+ Z	Link Z^	GT3X+ VM	Link VM
RAW DATA - MEAN								
Sedentary activities								
Lying down (*n*=278)	0.73 (0.25)	0.72 (0.24)	-0.33 (0.23)*	-0.35 (0.20)	-0.40 (0.28)	-0.44 (0.29)	1.00 (0.03)*	1.01 (0.02)
Computer (*n*=303)	0.05 (0.17)	0.07 (0.23)	-0.94 (0.06)*	-0.95 (0.09)	0.02 (0.26)*	-0.11 (0.27)	1.00 (0.03)*	1.03 (0.01)
Watching TV (*n*=440)	0.26 (0.22)	0.26 (0.26)	-0.84 (0.17)*	-0.88 (0.15)	0.09 (0.40)*	0.03 (0.38)	1.02 (0.06)*	1.05 (0.06)
Writing (*n*=100)	0.04 (0.15)	0.06 (0.17)	-0.91 (0.09)*	-0.93 (0.07)	0.09 (0.39)*	-0.06 (0.39)	1.01 (0.03)*	1.03 (0.01)
Reading (*n*=519)	0.21 (0.19)*	0.19 (0.21)	-0.88 (0.14)*	-0.91 (0.11)	0.02 (0.38)*	-0.05 (0.38)	1.01 (0.03)*	1.03 (0.01)
Playing cards (*n*=348)	0.03 (0.16)	0.02 (0.18)	-0.85 (0.27)*	-0.90 (0.17)	0.18 (0.38)*	0.07 (0.42)	1.00 (0.03)*	1.03 (0.01)
Household activities								
Standing (*n*=88)	-0.11 (0.12)	-0.09 (0.18)	-0.92 (0.08)*	-0.95 (0.10)	0.18 (0.27)*	0.09 (0.30)	0.99 (0.03)*	1.03 (0.01)
Dusting (*n*=66)	-0.26 (0.15)	-0.28 (0.18)	-0.84 (0.28)	-0.85 (0.27)	0.17 (0.25)*	0.09 (0.29)	1.02 (0.03)*	1.03 (0.01)
Gardening (*n*=39)	-0.12 (0.29)	-0.09 (0.23)	-0.85 (0.20)	-0.88 (0.20)	0.33 (0.25)	0.20 (0.38)	1.03 (0.03)*	1.05 (0.04)
Laundry (*n*=114)	-0.19 (0.17)	-0.17 (0.19)	-0.93 (0.06)*	-0.94 (0.08)	0.12 (0.19)*	0.07 (0.25)	1.00 (0.03)*	1.03 (0.01)
Making bed (*n*=74)	-0.19 (0.18)	-0.20 (0.17)	-0.88 (0.16)*	-0.89 (0.17)	0.23 (0.27)*	0.18 (0.29)	1.03 (0.03)*	1.04 (0.01)
Picking up items (*n*=75)	-0.29 (0.25)*	-0.23 (0.25)	-0.76 (0.16)*	-0.81 (0.13)	0.27 (0.23)	0.25 (0.27)	1.03 (0.03)*	1.05 (0.01)
Sweeping (*n*=112)	-0.29 (0.15)	-0.27 (0.18)	-0.90 (0.07)	-0.91 (0.09)	0.14 (0.19)*	0.02 (0.28)	1.01 (0.03)*	1.03 (0.01)
Vacuuming (*n*=62)	-0.29 (0.14)*	-0.36 (0.18)	-0.85 (0.16)*	-0.81 (0.18)	0.27 (0.29)	0.21 (0.34)	1.03 (0.01)	1.03 (0.01)
Ambulatory/exercise activities								
Walk – brisk (*n*=165)	-0.15 (0.15)*	-0.18 (0.16)	-0.96 (0.07)*	-0.95 (0.08)	0.14 (0.21)	0.12 (0.28)	1.08 (0.05)*	1.09 (0.04)
Walk – leisure (*n*=151)	-0.18 (0.12)	-0.19 (0.17)	-0.95 (0.07)	-0.95 (0.08)	0.14 (0.25)	0.14 (0.24)	1.05 (0.03)*	1.06 (0.02)
Walk – treadmill (*n*=279)	-0.20 (0.10)	-0.20 (0.14)	-0.96 (0.04)	-0.96 (0.07)	0.15 (0.19)	0.12 (0.24)	1.07 (0.07)*	1.09 (0.06)
Jog – overground (*n*=38)	-0.13 (0.15)*	-0.07 (0.14)	-0.92 (0.11)*	-0.97 (0.11)	0.39 (0.24)	0.36 (0.21)	1.21 (0.12)*	1.23 (0.11)
Jog – treadmill (*n*=58)	-0.22 (0.08)*	-0.13 (0.14)	-0.96 (0.02)*	-0.99 (0.02)	0.12 (0.12)	0.14 (0.17)	1.16 (0.11)*	1.19 (0.13)
Cycling (*n*=188)	-0.08 (020)	-0.07 (0.19)	-0.91 (0.11)*	-0.94 (0.10)	0.14 (0.28)*	-0.01 (0.33)	1.00 (0.03)*	1.03 (0.02)

Stairs (n=148)	-0.22 (0.13)*	-0.26 (0.14)	-0.93 (0.07)*	-0.91 (0.11)	0.19 (0.19)	0.16 (0.29)	1.04 (0.03)*	1.05 (0.02)
RAW DATA - VARIANCE								
Sedentary activities								
Lying down	0.00 (0.00)*	0.00 (0.00)	0.00 (0.00)*	0.00 (0.00)	0.00 (0.01)	0.00 (0.01)	0.00 (0.00)*	0.00 (0.00)
Computer	0.00 (0.01)*	0.00 (0.01)	0.00 (0.00)*	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)*	0.00 (0.00)
Watching TV	0.01 (0.04)*	0.01 (0.06)	0.04 (0.25)*	0.04 (0.30)	0.01 (0.06)*	0.01 (0.04)	0.03 (0.21)*	0.03 (0.23)
Writing	0.00 (0.00)*	0.00 (0.00)	0.00 (0.00)*	0.00 (0.00)	0.00 (0.00)*	0.00 (0.00)	0.00 (0.00)*	0.00 (0.00)
Reading	0.00 (0.01)	0.00 (0.01)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)*	0.00 (0.00)
Playing cards	0.00 (0.01)	0.00 (0.01)	0.00 (0.01)	0.00 (0.00)	0.00 (0.01)	0.00 (0.01)	0.00 (0.00)	0.00 (0.00)
Household activities								
Standing	0.00 (0.00)*	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)*	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
Dusting	0.02 (0.02)	0.02 (0.02)	0.03 (0.05)	0.03 (0.05)	0.02 (0.01)	0.02 (0.01)	0.01 (0.00)*	0.01 (0.00)
Gardening	0.04 (0.06)*	0.04 (0.07)	0.05 (0.09)*	0.06 (0.11)	0.01 (0.02)	0.01 (0.02)	0.01 (0.12)*	0.01 (0.13)
Laundry	0.01 (0.01)	0.01 (0.01)	0.00 (0.00)*	0.00 (0.00)	0.01 (0.01)	0.01 (0.01)	0.00 (0.00)*	0.00 (0.00)
Making bed	0.03 (0.03)	0.03 (0.02)	0.02 (0.01)	0.02 (0.02)	0.02 (0.02)	0.02 (0.01)	0.00 (0.00)*	0.00 (0.00)
Picking up items	0.07 (0.04)	0.08 (0.05)	0.09 (0.06)*	0.08 (0.05)	0.05 (0.03)	0.05 (0.03)	0.03 (0.02)*	0.03 (0.02)
Sweeping	0.02 (0.01)*	0.02 (0.01)	0.01 (0.01)*	0.01 (0.01)	0.02 (0.01)	0.02 (0.01)	0.00 (0.00)*	0.00 (0.00)
Vacuuming	0.02 (0.01)	0.02 (0.01)	0.02 (0.02)*	0.02 (0.02)	0.02 (0.01)	0.02 (0.01)	0.01 (0.00)*	0.01 (0.00)
Ambulatory/exercise activities								
Walk – brisk over-ground	0.11 (0.09)*	0.10 (0.06)	0.14 (0.08)	0.14 (0.10)	0.05 (0.03)*	0.04 (0.03)	0.17 (0.10)*	0.16 (0.11)
Walk – leisure over-ground	0.05 (0.03)*	0.05 (0.02)	0.06 (0.02)*	0.05 (0.02)	0.03 (0.01)	0.03 (0.01)	0.07 (0.03)*	0.07 (0.03)
Walk – treadmill	0.12 (0.13)*	0.10 (0.09)	0.12 (0.09)*	0.13 (0.14)	0.05 (0.03)	0.05 (0.05)	0.15 (0.12)*	0.16 (0.14)
Jog – over-ground	0.12 (0.07)*	0.09 (0.04)	1.06 (0.72)*	1.08 (0.72)	0.29 (0.19)*	0.22 (0.13)	1.07 (0.70)*	1.00 (0.62)
Jog – treadmill	0.18 (0.08)*	0.16 (0.10)	0.61 (0.48)*	0.67 (0.61)	0.07 (0.04)*	0.08 (0.03)	0.50 (0.33)	0.53 (0.42)
Cycling	0.01 (0.01)*	0.01 (0.01)	0.01 (0.01)	0.01 (0.01)	0.01 (0.01)*	0.00 (0.00)	0.01 (0.01)*	0.01 (0.01)
Stairs	0.05 (0.03)*	0.05 (0.02)	0.08 (0.04)*	0.07 (0.03)	0.03 (0.01)	0.03 (0.01)	0.09 (0.04)*	0.09 (0.03)

^Indicates that the signs (+ or -) of data for this axis were reversed.
*Indicates significant difference from Link accelerometer.
n signifies the number of 30-s epochs for which each activity was performed.
Data are shown as mean (standard deviation) in gravitational units.

Table 2. Activity-specific, count data for ActiGraph GT3X+ and Link.

	GT3X+ X	Link X	GT3X+ Y	Link Y	GT3X+ Z	Link Z	GT3X+ VM	Link VM
Sedentary activities								
Lying down	2.59 (22.59)	2.90 (23.71)	3.78 (46.25)	4.30 (52.51)	4.01 (35.60)	3.76 (30.31)	6.49 (62.55)	6.99 (65.04)
Computer	0.57 (4.85)*	1.32 (10.66)	6.50 (34.05)	6.97 (35.14)	7.35 (30.04)*	5.21 (20.10)	11.04 (45.38)	10.68 (41.42)
Watching TV	1.30 (9.06)	1.53 (9.97)	7.63 (35.82)	7.84 (37.31)	9.13 (43.75)*	7.87 (43.36)	14.09 (56.77)	12.96 (57.70)
Writing	0.00 (0.00)	0.22 (1.17)	1.89 (7.01)	3.38 (10.81)	7.82 (25.78)*	4.24 (14.02)	8.44 (26.59)	6.18 (17.49)
Reading	0.88 (10.69)	0.97 (11.36)	4.01 (19.31)	3.76 (17.55)	6.06 (23.39)	5.05 (21.50)	8.49 (31.87)	7.46 (29.74)
Playing cards	5.25 (41.17)	6.20 (40.26)	16.35 (56.66)	18.59 (73.37)	19.87 (54.90)	19.49 (49.39)	30.91 (87.48)*	33.81 (95.20)
Household activities								
Standing	3.19 (14.92)	2.50 (12.84)	6.06 (18.45)*	9.02 (20.87)	21.63 (68.69)*	13.18 (51.36)	24.61 (72.04)*	19.32 (55.91)
Dusting	331.44 (511.48)*	351.45 (502.63)	425.26 (188.34)	426.30 (172.18)	581.52 (281.60)*	539.41 (238.68)	881.74 (474.87)	862.03 (436.02)
Gardening	13.49 (43.45)	17.21 (45.35)	57.97 (120.54)	49.92 (88.88)	52.31 (106.83)	67.13 (154.37)	89.27 (161.54)	98.12 (177.19)
Laundry	28.77 (45.74)*	40.50 (57.22)	252.90 (163.60)	248.81 (171.13)	250.94 (215.05)	250.99 (218.07)	381.56 (239.00)	380.76 (248.03)
Making bed	322.73 (276.28)	324.88 (308.44)	630.07 (289.00)	610.86 (283.21)	625.86 (288.49)	632.80 (304.83)	988.83 (396.22)	983.20 (423.70)
Picking up items	1666.09 (687.12)*	1569.05 (640.53)	1598.81 (548.64)	1597.16 (624.83)	1466.48 (380.03)	1486.16 (365.25)	2801.57 (739.24)	2750.73 (764.51)
Sweeping	147.30 (131.73)*	218.25 (201.24)	607.10 (311.09)*	577.75 (261.47)	758.31 (282.03)	816.07 (350.72)	1025.35 (326.88)*	1069.91 (371.50)
Vacuuming	279.39 (226.56)*	419.00 (385.13)	607.19 (283.68)	676.16 (365.46)	893.34 (468.26)	900.61 (480.36)	1190.88 (418.12)*	1290.85 (535.30)
Ambulatory/exercise activities								
Walk – brisk over-ground	1956.54 (690.12)*	1880.04 (704.11)	1059.19 (460.70)*	1114.27 (452.52)	767.95 (370.32)	786.28 (360.04)	2411.57 (740.24)*	2382.89 (737.99)
Walk – leisure over-ground	1209.27 (401.46)	1196.66 (394.87)	865.21 (334.67)	846.21 (321.99)	683.88 (300.17)	733.37 (281.06)	1679.16 (470.50)	1672.59 (475.57)
Walk – treadmill	1552.78 (580.93)*	1514.39 (573.75)	1089.66 (527.81)	1066.90 (533.98)	616.80 (298.59)*	715.86 (496.73)	2045.37 (707.01)*	2070.61 (718.69)
Jog – over-ground	4357.39 (1024.37)*	4707.79 (1185.50)	1677.82 (433.50)*	1551.00 (516.32)	2068.71 (801.47)*	1596.16 (775.77)	5165.24 (1124.84)*	5288.03 (1185.03)
Jog – treadmill	2767.31 (1692.09)	2771.53 (1673.33)	1113.79 (322.76)	1070.98 (561.51)	623.79 (206.07)*	876.60 (234.08)	3121.33 (1489.41)	3163.98 (1658.00)
Cycling	59.45 (99.64)	60.16 (112.70)	81.51 (128.02)*	120.06 (163.08)	161.08 (226.09)*	101.94 (164.71)	241.90 (234.36)	218.72 (216.60)
Stairs	1551.20 (343.85)*	1521.46 (362.21)	836.70 (267.46)	839.42 (211.40)	1070.59 (299.29)	1102.20 (295.65)	2091.60 (395.22)	2089.24 (363.01)

*Indicates significant difference from Link accelerometer.
Data shown as mean (standard deviation) and expressed as counts per 30 s.

Table 3. Activity-specific predicted EE (METs) from each EE prediction equation.

	GT3X+ 2015	**Link 2015**	**GT3X+ 1998**	**Link 1998**	**GT3X+ 2011**	**Link 2011**
Sedentary activities						
Lying down	1.7 (0.8)	1.7 (0.8)	1.4 (0.0)	1.4 (0.0)	0.7 (0.1)	0.7 (0.1)
Computer	1.1 (0.4)	1.1 (0.5)	1.4 (0.0)	1.4 (0.0)	0.7 (0.1)	0.7 (0.1)
Watching TV	1.5 (1.7)	1.5 (1.7)	1.4 (0.0)	1.4 (0.0)	0.7 (0.1)	0.7 (0.1)
Writing	1.3 (0.5) ^	1.1 (0.2)	1.4 (0.0)	1.4 (0.0)	0.7 (0.0)	0.7 (0.0)
Reading	1.2 (0.7)	1.2 (0.7)	1.4 (0.0)	1.4 (0.0)	0.7 (0.1)	0.7 (0.1)
Playing cards	1.5 (1.1)	1.5 (1.0)	1.4 (0.1)	1.4 (0.1)	0.7 (0.2)	0.7 (0.2)
Household activities						
Standing	1.3 (0.7)	1.2 (0.5)	1.4 (0.0)	1.4 (0.0)	0.7 (0.1)	0.7 (0.1)
Dusting	2.5 (1.4)	2.5 (1.5)	2.0 (0.8)	2.0 (0.8)	2.2 (0.8)	2.2 (0.8)
Gardening	3.0 (2.6)	3.2 (2.7)	1.5 (0.1)	1.5 (0.1)	0.8 (0.3)	0.8 (0.3)
Laundry	1.5 (0.5)	1.6 (0.6)	1.5 (0.1)	1.5 (0.1)	1.3 (0.4)	1.3 (0.4)
Making bed	3.0 (1.3)	2.8 (1.4)	2.0 (0.5)	2.0 (0.5)	2.4 (0.7)	2.4 (0.7)
Picking up items	5.3 (1.7)	5.4 (1.8)	4.1 (1.1)^	3.9 (1.0)	5.5 (1.3)	5.4 (1.3)
Sweeping	2.2 (0.6)	2.1 (0.8)	1.7 (0.2)^	1.8 (0.3)	2.4 (0.6)^	2.5 (0.6)
Vacuuming	3.1 (1.3)	3.1 (1.2)	1.9 (0.4)^	2.1 (0.6)	2.7 (0.7)^	2.9 (0.9)
Ambulatory/exercise activities						
Walk – brisk over-ground	6.2 (1.6)^	5.8 (1.5)	4.5 (1.1)	4.4 (1.1)	4.8 (1.3)	4.8 (1.3)
Walk – leisure over-ground	4.7 (0.9)^	4.5 (1.0)	3.4 (0.6)	3.3 (0.6)	3.6 (0.8)	3.6 (0.8)
Walk – treadmill	5.6 (1.5)^	5.4 (1.5)	3.9 (0.9)^	3.8 (0.9)	4.2 (1.2)	4.2 (1.2)
Jog – over-ground	9.2 (4.7)	9.2 (3.7)	8.4 (1.6)^	8.9 (1.9)	9.6 (1.9)^	9.8 (2.0)
Jog – treadmill	6.5 (1.9)^	7.4 (1.5)	5.8 (2.5)	5.8 (2.7)	6.1 (2.6)	6.1 (2.9)
Cycling	1.7 (1.2)^	1.6 (1.0)	1.5 (0.2)	1.5 (0.2)	1.1 (0.4)	1.1 (0.4)
Stairs	5.5 (1.2)^	5.1 (1.2)	3.9 (0.5)	3.9 (0.6)	4.3 (0.7)	4.3 (0.6)

EE: energy expenditure in METs.
GT3X+ 2015: METs predicted from Montoye 2015 model for raw GT3X+ data.
Link 2015: METs predicted from Montoye 2015 model for raw Link data.
GT3X+ 1998: METs predicted from Freedson 1998 equation for count raw GT3X+ data.
Link 1998: METs predicted from Freedson 1998 equation for count Link data.
GT3X+ 2011: METs predicted from Freedson 2011 equation for count raw GT3X+ data.
Link 2011: METs predicted from Freedson 2011 equation for count Link data.
^Indicates significant difference from Link accelerometer.

Table 4. Time spent in each intensity, % agreement, and kappa for raw and count data equations.

Intensity	Raw data – Montoye 2015 model		Count data – Freedson 1998 equation		Count data – Freedson 2011 equation	
	GT3X+	**Link**	**GT3X+**	**Link**	**GT3X+**	**Link**
Sedentary	45.2 (16.8)	41.4 (14.6)	49.0 (6.5)	48.3 (6.4)	42.1 (5.5)	42.3 (5.4)
Light	19.0 (13.0)*	12.9 (5.6)	13.9 (6.1)	15.1 (8.3)	19.7 (7.1)	19.2 (6.8)
Moderate	10.6 (6.8)*	18.2 (11.3)	15.7 (6.3)	15.2 (7.2)	15.6 (7.5)	16.0 (7.4)
Vigorous	5.4 (6.3)*	7.6 (6.7)	1.6 (2.9)	1.6 (2.9)	2.7 (4.0)	2.6 (4.0)
MVPA	16.0 (0.5)*	25.8 (13.8)	17.3 (6.4)	16.7 (7.2)	18.3 (7.5)	18.6 (7.3)
Kappa	0.45 (0.32)[1,2]		0.92 (0.09)		0.93 (0.07)	
Percent agreement	61.5 (27.6)[1,2]		95.1 (5.6)		95.5 (4.0)	

MVPA: moderate- or vigorous-intensity physical activity (moderate + vigorous)
*Indicates significant differences from predicted time spent in given intensity by Link accelerometer.
[1]Indicates significant difference from count data – Freedson 1998.
[2]Indicates significant difference from count data – Freedson 2011.

Supplemental Digital Content 1. ActiGraph GT3X+ and Link axis orientation.

a. ActiGraph GT3X+ **b.** ActiGraph GT9X Link