# Estimating Energy Expenditure with ActiGraph GT9X Inertial Measurement Unit

PAUL R. HIBBING, SAMUEL R. LAMUNION, ANDREW S. KAPLAN, and SCOTT E. CROUTER

*Department of Kinesiology, Recreation, and Sport Studies, The University of Tennessee Knoxville, Knoxville, TN*

**ABSTRACT**

HIBBING, P. R., S. R. LAMUNION, A. S. KAPLAN, and S. E. CROUTER. Estimating Energy Expenditure with ActiGraph GT9X Inertial Measurement Unit. *Med. Sci. Sports Exerc.*, Vol. 50, No. 5, pp. 1093–1102, 2018. **Purpose**: The purpose of this study was to explore whether gyroscope and magnetometer data from the ActiGraph GT9X improved accelerometer-based predictions of energy expenditure (EE). **Methods**: Thirty participants (mean ± SD: age, 23.0 ± 2.3 yr; body mass index, 25.2 ± 3.9 $kg \cdot m^{-2}$) volunteered to complete the study. Participants wore five GT9X monitors (right hip, both wrists, and both ankles) while performing 10 activities ranging from rest to running. A Cosmed $K4b^2$ was worn during the trial, as a criterion measure of EE (30-s averages) expressed in METs. Triaxial accelerometer data (80 Hz) were converted to milli-G using Euclidean norm minus one (ENMO; 1-s epochs). Gyroscope data (100 Hz) were expressed as a vector magnitude (GVM) in degrees per second (1-s epochs) and magnetometer data (100 Hz) were expressed as direction changes per 5 s. Minutes 4–6 of each activity were used for analysis. Three two-regression algorithms were developed for each wear location: 1) ENMO, 2) ENMO and GVM, and 3) ENMO, GVM, and direction changes. Leave-one-participant-out cross-validation was used to evaluate the root mean square error (RMSE) and mean absolute percent error (MAPE) of each algorithm. **Results**: Adding gyroscope to accelerometer-only algorithms resulted in RMSE reductions between 0.0 METs (right wrist) and 0.17 METs (right ankle), and MAPE reductions between 0.1% (right wrist) and 6.0% (hip). When direction changes were added, RMSE changed by ≤0.03 METs and MAPE by ≤0.21%. **Conclusions**: The combined use of gyroscope and accelerometer at the hip and ankles improved individual-level prediction of EE compared with accelerometer only. For the wrists, adding gyroscope produced negligible changes. The magnetometer did not meaningfully improve estimates for any algorithms. **Key Words:** ACCELEROMETER, GYROSCOPE, MAGNETOMETER, TWO-REGRESSION, MULTISENSOR

To identify links between physical activity (PA) and health outcomes, it is essential to have valid tools for measuring PA. Sensors that track movement and physiological responses to PA have gained widespread use, and accelerometers are among the most common (1,2). Accelerometers record linear acceleration, which is one aspect of movement that can be analyzed to predict PA outcomes. Other aspects of movement can be measured with a variety of sensors (3), but these sensors have not been used to the same extent as accelerometers. Two key examples are gyroscopes and magnetometers, which measure rotational velocity and magnetic flux, respectively.

In general, the primary focus of previous studies that have used gyroscopes and magnetometers has been on classification tasks related to PA recognition, with limited attention given to the prediction of energy expenditure (EE) (4,5). Gyroscopes have been used more than magnetometers and have most commonly been applied to gait analysis and recognition of posture and locomotion (6–9). More recently, the gyroscopes, magnetometers, and accelerometers in smartphones have been used in concert to predict specific activities individuals perform (10,11).

Previous research with accelerometers has shown that using a two-step process to predict EE leads to more accurate estimates (12,13). In the first step, activity or activity type is classified, and in the second step, an EE estimate is made using a model developed specifically for that classification. This idea was first introduced in the two-regression algorithms developed by Heil (14) and Crouter et al. (15,16), and has since received increased attention (12,17–19). However, these studies have exclusively used accelerometers. Thus, it is unknown whether gyroscopes and magnetometers may improve these types of algorithms.

Until recently, gyroscopes and magnetometers have not been readily available to researchers, and the increased participant burden from having to wear multiple devices has further contributed to their lack of use. In 2014, ActiGraph (ActiGraph, LLC, Pensacola, FL) released the GT9X, which

Address for correspondence: Paul Hibbing, M.Sc., Department of Kinesiology, Recreation, and Sport Studies, The University of Tennessee Knoxville, 1914 Andy Holt Ave, Knoxville, TN 37996; E-mail: phibbing@vols.utk.edu.
Submitted for publication July 2017.
Accepted for publication December 2017.
Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's Web site (www.acsm-msse.org).

0195-9131/18/5005-1093/0
MEDICINE & SCIENCE IN SPORTS & EXERCISE®

DOI: 10.1249/MSS.0000000000001532

houses multiple sensors in a single device. The GT9X has a primary triaxial accelerometer that is the same as its predecessor (wGT3X-BT), whereas the other sensors are contained in an inertial measurement unit (IMU). These include a triaxial gyroscope, triaxial magnetometer, thermometer, and a secondary triaxial accelerometer. The purpose of this study was to investigate whether combining accelerometer, gyroscope, and magnetometer data improved EE predictions, compared with predictions made using only an accelerometer.

## METHODS

### Participants

Thirty participants were recruited from the University of Tennessee, Knoxville, and the Knoxville Community via e-mail, flyers, and word of mouth. The exclusion criteria were as follows: 1) contraindications for exercise, based on an answer of "no" to any question on the Physical Activity Readiness Questionnaire (20); 2) recent (within the last 6 months) orthopedic or musculoskeletal injury; 3) currently pregnant; or 4) body mass index of $\geq 35$ kg·m$^{-2}$. Participants were not screened for pretrial food intake or PA. The study was approved by The University of Tennessee, Knoxville Institutional Review Board, and participants gave written informed consent before participation in the study.

### Experimental Protocol

Participants reported to the Applied Physiology Laboratory and had their height and weight measured, in light clothing without shoes, using a wall-mounted stadiometer and a calibrated physician's scale, respectively. They were then fitted with five GT9X monitors and a portable metabolic system (Cosmed K4b$^2$). The primary accelerometer for the GT9X was initialized to sample at 80 Hz and the IMU was activated. The GT9X monitors were positioned on the right hip (along the anterior axillary line, at the level of the iliac crest), both wrists (between the ulnar and styloid processes), and both ankles (immediately superior to the lateral malleoli). The hip and ankle devices were attached to elastic and Velcro bands, respectively, using custom clips from the manufacturer, and the wrist devices were worn in dedicated wristbands provided by the manufacturer. The start time for the K4b$^2$ was manually recorded and synchronized with the system clock for merging with the GT9X data.

Each participant performed 10 activities in the following order: supine rest, computer work, table cleaning, sweeping, overground walking, ascending and descending stairs, basketball, tennis, overground slow run, and overground fast run. The first four activities were performed in the Applied Physiology Laboratory, and the others were performed nearby in a gym, stairwell, or outdoor recreation area, as appropriate. All activities were performed for 7 min and were self-paced where applicable. The start and stop time of each activity was recorded manually and used as a reference to annotate the data from the GT9X and the portable metabolic system.

### Instruments

**ActiGraph GT9X link.** The GT9X is a small (3.5 × 3.5 × 1.0 cm) and lightweight (14 g) triaxial accelerometer–based PA monitor that contains a primary accelerometer and an IMU that contains four additional sensors. The primary accelerometer measures acceleration in the dynamic range of ±8*g*. Acceleration is digitized using a 12-bit analog to digital converter and sampled at user-selected rate between 30 and 100 Hz in increments of 10 Hz. These data are low-pass filtered at a proprietary cutoff and stored and downloaded as raw acceleration.

The four sensors in the IMU are a thermometer and three triaxial sensors (secondary accelerometer, gyroscope, and magnetometer). Each sensor samples at 100 Hz, independently of the selected sampling rate for the primary accelerometer, and the IMU must be specifically activated when the device is initialized. The dynamic ranges of the triaxial IMU sensors are ±16*g* (secondary accelerometer), ±2000°·s$^{-1}$ (gyroscope), and ±4800 $\mu$T (magnetometer). The thermometer in the IMU is used for the temperature of the sensor, not the skin. No low-pass filter is applied to the raw data obtained from the IMU sensors.

**Cosmed K4b$^2$.** A Cosmed K4b$^2$ (Cosmed, Rome, Italy) portable metabolic system was used for the measurement of oxygen consumption ($\dot{V}O_2$) and carbon dioxide production. The K4b$^2$ consists of a molded-rubber facemask and a chest-mounted harness which holds a gas analyzer unit and a battery unit. The device is lightweight (approximately 800 g) and has been validated for use in measuring $\dot{V}O_2$ during a range of intensities on a cycle ergometer (21). A four-step calibration was performed before each test according to the manufacturer's specifications. First, a room air calibration was performed using the ambient temperature and relative humidity of the room. Second, a reference gas calibration was performed using a gas mixture of 15.980% $O_2$ and 4.008% carbon dioxide. Third, a flow meter calibration was performed using a 3-L Hans Rudolf syringe. Finally, a delay calibration was performed to account for lag between sampled air at the flow meter in the face mask and the gas sensors in the metabolic unit.

### Data Reduction

The final calibration data set consisted of one entry per participant for each activity completed. All variables (metabolic values and GT9X features) were averages from steady-state activity, based on 30-s epochs of metabolic data and 1-s epochs of GT9X data, as described hereinafter. Steady-state activity was considered minutes 4–6 of each activity except supine rest, where only minute 6 was used, to allow more time for metabolism to adjust to a near-resting value. The same calibration data set was used to develop and cross-validate the algorithms.

The K4b$^2$ data were saved as breath-by-breath measurements and converted to 30-s averages. For the calibration data set, the mean 30-s $\dot{V}O_2$ (mL·min$^{-1}$) for each participant and activity was calculated, then converted to relative $\dot{V}O_2$

($mL \cdot kg^{-1} \cdot min^{-1}$) by dividing by the participant's body mass. Relative $\dot{V}O_2$ was then converted to METs ($3.5\ mL \cdot kg^{-1} \cdot min^{-1}$ = 1 MET), which served as the criterion measure of EE in the calibration data set. For weight-bearing activities, 2 kg was added to the participant's measured body mass to account for the weight of the devices.

The GT9X data came from the primary accelerometer, gyroscope, and magnetometer. For the primary accelerometer, the data were first collapsed using a function written by the developer of the R package GGIR (22). The function calculated the Euclidean norm minus one (ENMO; in milli-G) (23) with a 1-s output window, which involved combining all three axes of 80-Hz acceleration data into 1-s epochs of vector magnitude with the acceleration due to gravity subtracted. For the calibration data set, the mean 1-s ENMO value for each participant and activity was used.

For the gyroscope, the 100-Hz data for each axis were low-pass filtered at 35 Hz (24) using a second-order Butterworth filter. The data were then expressed as a gyroscope vector magnitude (GVM; $GVM = \sqrt{X^2 + Y^2 + Z^2}$) and averaged into 1-s epochs. The mean 1-s GVM value for each participant and activity was used in the calibration data set.

The magnetometer data can be recoded into a categorical indicator of cardinal direction using one of two orientation-specific conversion schemes provided by the manufacturer (25). A total of 16 directions are possible, comprising the four cardinal directions with three intermediate directions (e.g., NNE, NE, ENE) between each. For the analyses in the present study, raw data from each axis were averaged into 1-s epochs, and then recoded to reflect cardinal direction for that second (using the conversion scheme for vertical orientation; that is, when worn on the hip, the monitor's *y*-axis points vertically in the frontal plane in the anatomical position). Each second of cardinal direction was then compared with the two previous and two following seconds, and the number of direction changes during that 5-s window was determined by comparing consecutive epochs. If two consecutive epochs were the same, no direction change was registered, but if they differed, a direction change was registered and added to the total for that window. Thus, the range of possible values for each window was 0–4 direction changes, and there was a unique 5-s window for each 1-s epoch. Consequently, the variable functioned as a second-by-second variable (i.e., having 1-s epoch length), although the units reflected direction change per 5 s. The first and last two windows had insufficient epochs and thus were assigned missing values. However, this did not affect the calibration data set because these 4 s of data occurred before the start or after the finish of the trial (i.e., just after initialization and just before download). For the calibration data set, the mean direction changes per 5 s for each participant and activity were used.

## Modeling

Previous research (14–16) has demonstrated that a two-regression approach to predicting EE is more effective than other regression-based methods. Thus, two-regression algorithms were developed for the present study using similar procedures to that used by Crouter et al. (15,16). For each accelerometer placement site, three algorithms were developed using the following predictors: 1) ENMO only, 2) ENMO and GVM, and 3) ENMO, GVM, and direction changes. Hereafter, these are referred to as algorithms 1–3, respectively.

Each two-regression algorithm was developed in the same general way. First, a classifier for distinguishing between sedentary behavior (SB) and non-SB was developed using receiver operating characteristic (ROC) analysis. The data for this analysis were taken from two SBs (supine rest and computer work) and two light-intensity activities (table cleaning and sweeping). Classifications were made based on ENMO for algorithm 1 and GVM for algorithms 2 and 3.

Next, another classifier was developed using ROC analysis, to distinguish between continuous walking or running (CWR) and intermittent activity. This analysis included only data that the first classifier labeled non-SB when it was applied to the full data set. Thus, any of the 10 activities could have been used for analysis, depending on the effectiveness of the first classifier. Of the data passed to the second classifier for analysis, observations from overground walking and overground slow and fast running were considered CWR, and all other activities were considered intermittent activity. Classifications were made on the basis of the coefficient of variation (CV) per 10 s of ENMO for algorithm 1 and the CV per 10 s of GVM for algorithms 2 and 3.

To calculate CV per 10 s, an extension of the method in the refined two-regression algorithm by Crouter et al. (13) was used. In this method, multiple CVs are calculated for each epoch, which compare it with (A) preceding epochs, (B) succeeding epochs, and (C) intermediate combinations of preceding and succeeding epochs. The minimum of these CVs is then assigned to the original epoch and used to classify CWR or intermittent activity. In the original study (13), each window for calculating CV consisted of six 10-s epochs, and there were six CVs considered for each epoch, corresponding to windows where it occupied the first through sixth position. For the present study, each window consisted of 10 1-s epochs, and there were 10 CVs considered for each epoch, corresponding to windows where it occupied the first through 10th position. The purpose of this technique is to capture transitions into and out of CWR in free-living. More detail is available in Ref. (13). The mean CV per 10 s for each participant and activity was used for analysis.

After the classifiers were developed, the two-regression algorithms were constructed using the following steps: 1) apply the SB classifier, and assign anything at or below that threshold an estimate of 1.25 METs; 2) apply the CWR classifier to the remaining data, dividing it into two sets (i.e., a CWR set and an intermittent activity set); and 3) develop separate regression equations for the data sets from step 2.

The value of 1.25 METs for SB was a departure compared with most other SB classifiers, which have used a value of

1.0 MET. The decision to use 1.25 METs was supported by a variety of evidence suggesting that SB in free-living frequently has elevated EE (>1.0 MET). Specifically, the 2011 compendium of physical activities (26) includes 35 activities that explicitly meet the Sedentary Behavior Research Network's recently published consensus definition of SB (27); that is, nonsleeping behaviors in seated, reclining, or lying posture, with an EE of ≤1.5 METs. Of those 35 activities, only 3 ("lying quietly and watching television," "laughing, sitting," and "sitting, playing traditional video game, computer game") have a value of 1.0 MET, whereas the other 32 activities are ≥1.3 METs, including "sitting quietly and watching television" and various computer tasks (typing, working, etc.). Thus, even under the historical assumption (28) that television and computer time are representative of total SB, EE is still greater than 1.0 MET in most of those cases. Furthermore, time use surveys in both the United States and Australia have suggested that television and computer activities may only constitute half or less of total SB (28,29). Thus, it seems even more likely that a considerable proportion of free-living SB occurs with an EE of greater than 1.0 MET, and that 1.25 METs is a more reasonable value to use in SB classifiers.

### Statistical Analysis

Descriptive statistics for the participants and device data were calculated. The thresholds for the SB and CWR classifiers were selected using the "closest.topleft" method in the pROC package for R (30), which maximizes sensitivity and specificity based on a Pythagorean distance from the ROC space's top-left corner (31). These thresholds were evaluated in terms of area under the curve (AUC), sensitivity, and specificity.

In the final algorithms, the classifier thresholds and regression parameters were obtained using the whole calibration data set (one data point per participant and activity). The same calibration data set was used to cross-validate the algorithms using leave-one-participant-out cross-validation (LOOCV). This process involved repeating the entire modeling process (i.e., development of the classifiers and the regression equations) separately for all possible subsamples of 29 participants, with the 30th participant held out for cross-validation. Thus, there were 30 iterations of the modeling process, and each participant's data were used one time for cross-validation. The 30 sets of cross-validation estimates were then appended into one set and evaluated using root mean square error (RMSE), mean absolute percent error (MAPE), and equivalence testing. Equivalence testing was performed with a ±10% zone of equivalence. For each algorithm, this consisted of two one-sided *t*-tests (each at $\alpha = 0.05$) to test whether the estimate was greater than 90% and less than 110% of the criterion, respectively. The *P* value for each equivalence test was the larger of the *P* values from the two one-sided tests. Data processing and analysis were carried out in R.

## RESULTS

Two participants did not complete the fast run activity and one participant had abnormal K4b$^2$ values for tennis and slow run, resulting in exclusion of these cases. There were two cases with partial steady-state data (1 min of fast running and 2.5 min of basketball) that were included in analysis. Participant characteristics are summarized in Table 1, and descriptive statistics from the calibration data set are shown in Table 2.

The performance of the SB and CWR classifiers is summarized in the supplemental content (see Table, Supplemental Digital Content 1, Results of receiver operating characteristic analysis for developing the thresholds used in the two-regression algorithms, http://links.lww.com/MSS/B164). For the SB classifiers using ENMO, all wear locations except the hip (AUC = 63.7%, sensitivity = 76.7%, specificity = 58.3%) had an AUC of ≥96.9% and sensitivities and specificities of ≥93.3%. When the GVM was used, all classifications for all wear locations had 100% accuracy. Figure 1 depicts the decision boundaries for a subset of the SB classifiers.

For the CWR classifiers using the ENMO CV per 10 s, all wear locations except the hip (AUC = 83.6%, sensitivity = 81.3%, specificity = 93.1%) had an AUC of ≥97.0% and sensitivities and specificities of ≥92.0%. When the CV per 10 s of GVM was used instead, AUC for the hip increased to 100.0%, with 100% accurate classification of all observations. For the wrists and ankles, AUC was ≥98.8%, and sensitivities and specificities were ≥93.1%.

Table 3 shows the thresholds and regression equations for each algorithm at each wear location. For all regression equations, linear regression was initially tested. After examining the fit, additional equations were tested as needed. For all intermittent activity equations, ENMO was fit with third-order polynomial terms. For the CWR equations at the wrists, ENMO was log transformed, and linear regression was applied.

The table in Supplemental Digital Content 2, Activity-specific performance of each two-regression algorithm, http://links.lww.com/MSS/B165, shows activity-specific measured and predicted MET values, as well as RMSE and MAPE, for all three algorithms and all five attachment sites. In terms of overall performance, each algorithm was significantly equivalent ($P < 0.008$) to the K4b$^2$ (mean measured EE = 5.3 METs; zone of equivalence = 4.8–5.8 METs). Table 4 shows the overall error (RMSE and MAPE) from LOOCV for each algorithm. Compared with algorithm 1, the RMSE of algorithm 2 for the hip, left ankle, and right ankle was lower by 0.14 METs (12.3%), 0.15 METs (12.9%), and 0.17 METs (14.4%), respectively. For MAPE, values were

TABLE 1. Physical characteristics of participants.

| | Male ($n$ = 20) | Female ($n$ = 10) | Total ($N$ = 30) |
|---|---|---|---|
| Age, yr | 23.1 ± 2.5 | 22.9 ± 1.9 | 23.0 ± 2.3 |
| Height, cm | 179.0 ± 7.6 | 161.8 ± 6.4 | 173.2 ± 10.9 |
| Weight, kg | 85.1 ± 15.0 | 59.4 ± 12.6 | 76.5 ± 18.7 |
| BMI, kg·m$^{-2}$ | 26.4 ± 3.3 | 22.5 ± 3.5 | 25.1 ± 3.8 |

Values are mean ± SD.
BMI, body mass index.

TABLE 2. Descriptive statistics of the analytical data set.

| | K4b$^2$, METs | ENMO, m*g* | GVM, dps | Magnetometer, dcp5s |
|---|---|---|---|---|
| Hip | | | | |
| Supine rest ($n$ = 30) | 1.4 ± 0.3 | 27.5 ± 11.6 | 2.2 ± 0.3 | 0.1 ± 0.3 |
| Computer work ($n$ = 30) | 1.5 ± 0.3 | 26.5 ± 12.3 | 2.9 ± 0.7 | 0.2 ± 0.3 |
| Table cleaning ($n$ = 30) | 2.8 ± 0.6 | 28.3 ± 8.2 | 30.7 ± 5.8 | 1.2 ± 0.6 |
| Sweeping ($n$ = 30) | 3.5 ± 1.0 | 36.0 ± 8.4 | 33.7 ± 6.2 | 1.3 ± 0.4 |
| Walking ($n$ = 30; 5 km·h$^{-1}$) | 3.7 ± 0.8 | 149.0 ± 42.1 | 48.4 ± 12.7 | 0.7 ± 0.2 |
| Stair walking ($n$ = 30) | 6.7 ± 1.1 | 142.8 ± 30.4 | 52.2 ± 8.8 | 1.4 ± 0.3 |
| Basketball ($n$ = 30) | 7.9 ± 1.7 | 223.5 ± 53.8 | 75.3 ± 11.2 | 1.5 ± 0.6 |
| Tennis ($n$ = 29) | 7.5 ± 1.6 | 193.4 ± 52.7 | 77.0 ± 9.3 | 1.8 ± 1.0 |
| Slow running ($n$ = 29; 9.1 km·h$^{-1}$) | 8.3 ± 1.3 | 480.8 ± 81.4 | 88.0 ± 15.0 | 1.0 ± 0.3 |
| Fast running ($n$ = 28; 12.1 km·h$^{-1}$) | 10.0 ± 1.4 | 596.7 ± 94.9 | 111.8 ± 17.5 | 1.2 ± 0.3 |
| Left wrist | | | | |
| Supine rest | — | 20.8 ± 17.0 | 1.5 ± 0.7 | 0.0 ± 0.0 |
| Computer work | — | 27.1 ± 12.7 | 6.8 ± 4.8 | 0.2 ± 0.2 |
| Table cleaning | — | 96.9 ± 54.2 | 66.6 ± 20.8 | 0.7 ± 0.7 |
| Sweeping | — | 124.1 ± 37.6 | 88.7 ± 15.7 | 1.0 ± 0.6 |
| Walking | — | 198.1 ± 61.2 | 140.3 ± 43.8 | 0.5 ± 0.3 |
| Stair walking | — | 222.6 ± 65.1 | 141.7 ± 40.7 | 0.8 ± 0.5 |
| Basketball | — | 578.7 ± 167.4 | 226.0 ± 44.2 | 0.8 ± 0.5 |
| Tennis | — | 381.5 ± 93.9 | 177.4 ± 26.6 | 1.2 ± 0.6 |
| Slow run | — | 774.8 ± 198.5 | 192.2 ± 38.1 | 0.6 ± 0.6 |
| Fast run | — | 1009.4 ± 244.8 | 223.3 ± 41.3 | 0.6 ± 0.5 |
| Right wrist | | | | |
| Supine rest | — | 31.9 ± 18.9 | 2.7 ± 0.6 | 0.1 ± 0.2 |
| Computer work | — | 33.7 ± 16.2 | 8.3 ± 6.3 | 0.1 ± 0.2 |
| Table cleaning | — | 131.6 ± 64.3 | 94.2 ± 24.3 | 2.0 ± 0.5 |
| Sweeping | — | 107.4 ± 40.4 | 86.9 ± 18.0 | 2.4 ± 0.5 |
| Walking | — | 134.5 ± 40.1 | 118.4 ± 32.5 | 1.4 ± 0.3 |
| Stair walking | — | 162.5 ± 42.8 | 120.5 ± 33.8 | 2.2 ± 0.3 |
| Basketball | — | 586.5 ± 136.7 | 236.2 ± 35.9 | 3.3 ± 0.2 |
| Tennis | — | 324.6 ± 100.8 | 156.2 ± 21.8 | 3.1 ± 0.3 |
| Slow run | — | 795.2 ± 210.4 | 193.6 ± 31.1 | 2.0 ± 0.4 |
| Fast run | — | 1086.8 ± 247.3 | 232.7 ± 36.0 | 2.1 ± 0.4 |
| Left ankle | | | | |
| Supine rest | — | 3.3 ± 4.3 | 1.7 ± 0.4 | 0.0 ± 0.0 |
| Computer work | — | 19.7 ± 8.2 | 2.4 ± 0.8 | 0.0 ± 0.0 |
| Table cleaning | — | 66.3 ± 27.7 | 43.8 ± 13.2 | 0.1 ± 0.1 |
| Sweeping | — | 95.3 ± 33.5 | 57.7 ± 14.8 | 0.0 ± 0.1 |
| Walking | — | 613.6 ± 126.6 | 171.2 ± 19.0 | 0.0 ± 0.0 |
| Stair walking | — | 413.4 ± 107.1 | 120.6 ± 17.5 | 0.0 ± 0.2 |
| Basketball | — | 594.2 ± 143.1 | 150.8 ± 23.3 | 0.1 ± 0.2 |
| Tennis | — | 593.9 ± 139.5 | 151.2 ± 19.0 | 0.0 ± 0.1 |
| Slow run | — | 1229.4 ± 239.5 | 253.4 ± 30.1 | 0.0 ± 0.1 |
| Fast run | — | 1704.6 ± 297.1 | 310.5 ± 38.3 | 0.0 ± 0.2 |
| Right ankle | | | | |
| Supine rest | — | 6.1 ± 10.7 | 2.1 ± 0.6 | 0.1 ± 0.2 |
| Computer work | — | 12.6 ± 10.0 | 3.1 ± 1.3 | 0.2 ± 0.2 |
| Table cleaning | — | 66.9 ± 27.2 | 43.6 ± 13.1 | 2.1 ± 0.4 |
| Sweeping | — | 98.4 ± 35.2 | 58.2 ± 15.1 | 2.3 ± 0.4 |
| Walking | — | 611.5 ± 140.1 | 171.3 ± 20.3 | 1.2 ± 0.3 |
| Stair walking | — | 420.7 ± 115.7 | 120.5 ± 18.0 | 2.3 ± 0.4 |
| Basketball | — | 590.7 ± 142.9 | 152.0 ± 24.5 | 2.5 ± 0.5 |
| Tennis | — | 599.8 ± 136.3 | 152.5 ± 20.1 | 2.7 ± 0.4 |
| Slow run | — | 1234.4 ± 227.9 | 252.4 ± 30.1 | 1.9 ± 0.4 |
| Fast run | — | 1706.1 ± 320.4 | 313.1 ± 40.1 | 2.3 ± 0.5 |

Values are mean ± SD. For each activity, the number of observations, average speed, and K4b$^2$ values are the same for all wear locations.
Epoch lengths: ENMO, 1-s epochs; GVM, 1-s epochs.
dps, degrees per second; dcp5s, direction changes per 5 s.

lower by 6.0%, 2.6%, and 3.9%, respectively. At the left and right wrists, the changes in RMSE were ≤0.06 METs, and MAPE decreased by ≤1.8%. From algorithm 2 to algorithm 3, all locations had differences in RMSE and MAPE of ≤0.03 METs and ≤0.21%, respectively.

## DISCUSSION

This study examined whether the addition of the ActiGraph GT9X's gyroscope and magnetometer improved EE estimates from two-regression algorithms compared with accelerometer-only algorithms. To our knowledge, this is the first study to use data from the GT9X IMU. There were two primary findings from this study. First, individual-level error of EE estimates decreased substantially when the gyroscope was added to the hip and ankle algorithms. Second, the gyroscope was superior to the accelerometer for classifying SB, achieving perfect classification accuracy at all wear locations. The magnetometer did not meaningfully improve any of the algorithms.

**Additive benefit of gyroscope.** The use of multiple sensors in a single device holds promise for improving PA

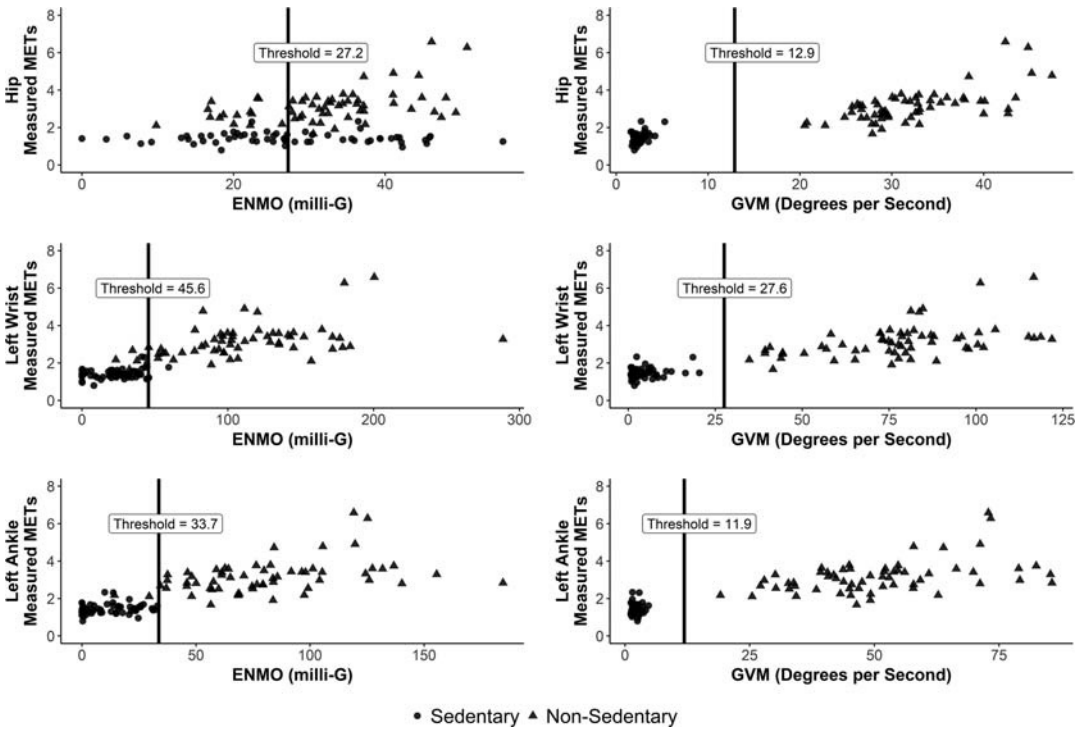


**FIGURE 1—Classification of SB (supine rest and computer work) and non-SB (sweeping and table cleaning) using ENMO (1-s epochs) or GVM (1-s epochs) for the hip, left wrist, and left ankle. Values below the threshold are classified as SB, and those above the threshold are classified as non-SB. The shapes reflect the actual classification.**

measurement. The improvements for the two-regression algorithms are attributable to the gyroscope's usefulness in both the initial classification step and the subsequent prediction step, which together make up the two-regression algorithms. In this study, compared with predictions obtained using only an accelerometer, the combined use of an accelerometer and gyroscope led to reductions in RMSE of up to 14.4% (algorithm 2, right ankle) and reductions in MAPE up to 6.0% (algorithm 2, hip).

For the classification step specifically, the GVM was able to differentiate SB from non-SB with 100% accuracy at all five wear locations. This is highly relevant to the growing body of research attempting to measure SB in free-living populations. For CWR, the gyroscope again led to improved classification, reaching 100% accuracy for the hip. For the other sites, sensitivity averaged 97.7% (SD, ±2.3%), and specificity averaged 96.9% (SD, ±2.7%). This may suggest that classification-focused research, such as that which has been commonly performed with gyroscopes (6–9) or accelerometers (32–34), could benefit from simultaneous measurement with both sensors, particularly if machine learning algorithms are used to distinguish between more activity classes.

**Additive benefit of magnetometer.** Adding magnetometer data to algorithms with the accelerometer and gyroscope resulted in changes in RMSE of ≤0.03 METs and changes in MAPE of ≤0.21% in this study. Previous research has suggested that magnetometers may be of little value for classifying activities (10,11), and the results of the present study suggest a similar lack of usefulness for predicting EE using regression-based approaches. Furthermore, the magnetometer is subject to several notable limitations.

A primary limitation of the magnetometer is that ascertaining cardinal direction is dependent on the monitor's orientation in space. For the present study, cardinal direction was coded using the manufacturer's scheme for monitors worn in a vertical position (i.e., *y*-axis of magnetometer aligned with the body's vertical axis in the anatomical position, when worn on the hip). The alternative scheme was for monitors worn in a horizontal position (i.e., when worn on the hip, the magnetometer's *y*-axis is aligned with the body's anteroposterior axis in the anatomical position). This is problematic during periods of frequent orientation changes in space. For example, an ankle-worn magnetometer changes orientation almost continuously during running. Thus, whether the vertical or horizontal classification scheme is used, it will be incorrect most of the time. It is possible that changes in direction are meaningful simply from a variability standpoint (i.e., more movement leads to more variability in the magnetometer signal, which translates to more direction changes, regardless of orientation), but this does not intuitively support the use of the magnetometer for predicting EE.

The mechanical properties of the magnetometer may also be problematic in free-living. It is sensitive to things like microwaves and magnets, which means that the interpretability of magnetometer data in free-living is questionable without knowledge of a participant's environment. The magnetometer can track position changes with respect to the earth's magnetic field, but this information likely overlaps

TABLE 3. Thresholds and regression equations for each algorithm developed for the hip, left and right wrists, and left and right ankles.

| Algorithm | Sedentary Threshold (1-s epoch) | CWR Threshold (CV) | Regression Equations |
|---|---|---|---|
| Hip | | | |
| 1 | ENMO ≤ 27.2 | ENMO ≤ 15.8% | CWR METs = 1.48 + 0.01405(ENMO) [$R^2$ = 0.90, SEE = 1.15]<br>IA METs = 1.76 + 0.04844(ENMO) − 0.00015(ENMO$^2$) + 0.0000003(ENMO$^3$) [$R^2$ = 0.76, SEE = 1.16] |
| 2 | GVM ≤ 12.9 | GVM ≤ 14.7% | CWR METs = 2.32 + 0.0141(ENMO) − 0.00893(GVM) [$R^2$ = 0.83, SEE = 1.21]<br>IA METs = 1.28 + 0.06305(ENMO) − 0.00025(ENMO$^2$) + 0.0000004(ENMO$^3$) + 0.00357(GVM) [$R^2$ = 0.85, SEE = 0.96] |
| 3 | GVM ≤ 12.9 | GVM ≤ 14.7% | CWR METs = 1.88 + 0.01419(ENMO) − 0.0158(GVM) + 1.01(Direction) [$R^2$ = 0.83, SEE = 1.19]<br>IA METs = 1.30 + 0.06295(ENMO) − 0.00025(ENMO$^2$) + 0.0000004(ENMO$^3$) + 0.00415(GVM) − 0.04(Direction) [$R^2$ = 0.85, SEE = 0.96] |
| Left wrist | | | |
| 1 | ENMO ≤ 45.6 | ENMO ≤ 19.4% | CWR METs = −12.13 + 3.1381(log(ENMO)) [$R^2$ = 0.83, SEE = 1.21]<br>IA METs = 0.81 + 0.03033(ENMO) − 0.00005(ENMO$^2$) + 0.00000002(ENMO$^3$) [$R^2$ = 0.67, SEE = 1.41] |
| 2 | GVM ≤ 27.6 | GVM ≤ 14.1% | CWR METs = −14.65 + 3.5685(log(ENMO)) − 0.00179(GVM) [$R^2$ = 0.85, SEE = 1.13]<br>IA METs = 0.58 + 0.02589(ENMO) − 0.00004(ENMO$^2$) + 0.00000002(ENMO$^3$) + 0.01(GVM) [$R^2$ = 0.71, SEE = 1.34] |
| 3 | GVM ≤ 27.6 | GVM ≤ 14.1% | CWR METs = −14.65 + 3.5666(log(ENMO)) − 0.00176(GVM) + 0.02(Direction) [$R^2$ = 0.85, SEE = 1.14]<br>IA METs = 0.27 + 0.02787(ENMO) − 0.00004(ENMO$^2$) + 0.00000002(ENMO$^3$) + 0.00823(GVM) + 0.34(Direction) [$R^2$ = 0.71, SEE = 1.33] |
| Right wrist | | | |
| 1 | ENMO ≤ 60.2 | ENMO ≤ 21.2% | CWR METs = −8.86 + 2.6564(log(ENMO)) [$R^2$ = 0.85, SEE = 1.13]<br>IA METs = 0.82 + 0.03423(ENMO) − 0.00004(ENMO$^2$) + 0.00000004(ENMO$^3$) [$R^2$ = 0.59, SEE = 1.58] |
| 2 | GVM ≤ 37.8 | GVM ≤ 15.6% | CWR METs = −8.99 + 2.5917(log(ENMO)) + 0.00256(GVM) [$R^2$ = 0.84, SEE = 1.15]<br>IA METs = 0.67 + 0.03014(ENMO) − 0.00005(ENMO$^2$) + 0.00000004(ENMO$^3$) + 0.00707(GVM) [$R^2$ = 0.61, SEE = 1.54] |
| 3 | GVM ≤ 37.8 | GVM ≤ 15.6% | CWR METs = −8.98 + 2.6013(log(ENMO)) + 0.00254(GVM) − 0.04(Direction) [$R^2$ = 0.84, SEE = 1.16]<br>IA METs = −0.36 + 0.02963(ENMO) − 0.00006(ENMO$^2$) + 0.00000004(ENMO$^3$) + 0.00519(GVM) + 0.57(Direction) [$R^2$ = 0.62, SEE = 1.52] |
| Left ankle | | | |
| 1 | ENMO ≤ 33.7 | ENMO ≤ 18.1% | CWR METs = 2.44 + 0.00444(ENMO) [$R^2$ = 0.70, SEE = 1.52]<br>IA METs = 1.94 + 0.01749(ENMO) − 0.00002(ENMO$^2$) + 0.00000002(ENMO$^3$) [$R^2$ = 0.81, SEE = 1.07] |
| 2 | GVM ≤ 11.9 | GVM ≤ 12.9% | CWR METs = −0.28 + 0.00299(ENMO) + 0.01697(GVM) [$R^2$ = 0.83, SEE = 1.20]<br>IA METs = 1.99 + 0.01711(ENMO) − 0.00002(ENMO$^2$) + 0.00000001(ENMO$^3$) − 0.00105(GVM) [$R^2$ = 0.84, SEE = 1.00] |
| 3 | GVM ≤ 11.9 | GVM ≤ 12.9% | CWR METs = −0.29 + 0.00295(ENMO) + 0.01714(GVM) + 1.27(Direction) [$R^2$ = 0.83, SEE = 1.20]<br>IA METs = 2.02 + 0.01877(ENMO) − 0.00002(ENMO$^2$) + 0.00000001(ENMO$^3$) − 0.00554(GVM) + 1.42(Direction) [$R^2$ = 0.84, SEE = 0.98] |
| Right ankle | | | |
| 1 | ENMO ≤ 28.6 | ENMO ≤ 18.0% | CWR METs = 2.16 + 0.00460(ENMO) [$R^2$ = 0.74, SEE = 1.48]<br>IA METs = 1.74 + 0.02013(ENMO) − 0.00003(ENMO$^2$) + 0.00000002(ENMO$^3$) [$R^2$ = 0.81, SEE = 1.10] |
| 2 | GVM ≤ 14.8 | GVM ≤ 12.3% | CWR METs = −1.13 + 0.00217(ENMO) + 0.02405(GVM) [$R^2$ = 0.83, SEE = 1.18]<br>IA METs = 1.78 + 0.0182(ENMO) − 0.00002(ENMO$^2$) + 0.00000002(ENMO$^3$) + 0.00168(GVM) [$R^2$ = 0.83, SEE = 1.00] |
| 3 | GVM ≤ 14.8 | GVM ≤ 12.3% | CWR METs = −1.50 + 0.00164(ENMO) + 0.02349(GVM) + 0.65(Direction) [$R^2$ = 0.84, SEE = 1.16]<br>IA METs = 1.33 + 0.01896(ENMO) − 0.00002(ENMO$^2$) + 0.00000002(ENMO$^3$) − 0.00021(GVM) + 0.22(Direction) [$R^2$ = 0.83, SEE = 1.00] |

Algorithm 1, ENMO (m*g*, 1-s epochs); Algorithm 2, ENMO and GVM (°·s$^{-1}$, 1-s epochs); Algorithm 3, ENMO, GVM, and Direction (direction changes per 5 s); CWR, continuous walking/running; CV, per 10 s; IA, intermittent activity; SEE, standard error of the estimate.

substantially with what the gyroscope provides. It is possible that using direction changes per 5 s is not the most effective way to use the magnetometer, but it seems unlikely that other treatments would have much effect on its usefulness in these algorithms.

**Practical considerations.** It is important to emphasize that this study used the same calibration data set at each stage of the modeling process and analysis, including cross-validation. Thus, all results reflect performance at the data set's level of aggregation (i.e., average data for each participant and activity). This is appropriate for the development stage of the algorithms, because it maximizes regression to the mean. However, it creates important differences between the aggregated data in the calibration data set and estimates that would be obtained from applying the algorithms to second-by-second data in free-living samples. Thus, there are several issues to address before using the algorithms in free-living or drawing comparisons to methods presented in previous studies.

A key issue that needs to be addressed is how to smooth the second-by-second EE predictions generated by the

TABLE 4. RMSE and MAPE for each algorithm and placement location, across all activities combined.

| | Algorithm 1 | | Algorithm 2 | | Algorithm 3 | |
|---|---|---|---|---|---|---|
| | RMSE | MAPE, % | RMSE | MAPE, % | RMSE | MAPE, % |
| Hip | 1.14 | 20.32 | 1.00 | 14.35 | 1.00 | 14.52 |
| Left wrist | 1.24 | 19.55 | 1.18 | 17.78 | 1.20 | 17.99 |
| Right wrist | 1.29 | 20.72 | 1.29 | 20.61 | 1.28 | 20.76 |
| Left ankle | 1.16 | 18.26 | 1.01 | 15.66 | 1.04 | 15.86 |
| Right ankle | 1.18 | 19.15 | 1.01 | 15.28 | 1.01 | 15.48 |

Measured EE (mean ± SD) across all activities was 5.29 ± 3.13 METs.
Algorithm 1, ENMO; Algorithm 2, ENMO and GVM; Algorithm 3, ENMO, GVM, and direction changes per minute.

algorithms. That is, the algorithms produce estimates based on second-by-second device data, which is an important part of promptly detecting and accounting for transitions between SB, CWA, and intermittent activity. However, a consequence in free-living is that normal variation in device signal from 1 s to another (e.g., during cyclic motions such as running) may lead to unrealistic variability in the predicted METs. This is especially concerning given the level of error observed in LOOCV (i.e., RMSE $\geq$ 1.0 MET, MAPE $\geq$ 14.35%), which could translate to rapid divergence from true EE in free-living. Thus, a smoothing approach (e.g., averaging each minute) is likely necessary to prevent this behavior. It was not within the scope of this study to address smoothing, because the focus was on examining the additive effect of the IMU sensors. Future studies should address this issue by examining smoothing techniques to identify the best approach. In the meantime, a suitable recommendation is to average the predictions over each minute, as Crouter et al. (13) did in the refined two-regression algorithm.

Another issue in free-living is correctly implementing the algorithms, which requires reducing the data to 1-s epochs, merging primary accelerometer and IMU data, calculating CV per 10 s, and applying the classifiers and regression equations to obtain the MET estimates. Of particular concern is the calculation of CV per 10 s using the technique presented in this study, which was based on the technique of Crouter et al. (13). It is essential to calculate this correctly, to enable transitions into and out of CWR to be detected freely. To facilitate the correct use of these algorithms, R code is available from https://github.com/paulhibbing/IMU-2Regression/blob/master/Demo_of_Algorithms.zip, which also includes sample data and an extensively documented tutorial. For non-R users, the output from the sample code can be used as a reference to ensure that processing has been performed correctly using another platform.

Although the algorithms need further validation (especially in free-living), they offer several benefits. A primary benefit is that they improve EE estimation without the use of "black box" elements such as artificial neural networks. This makes the algorithm components (e.g., regression coefficients) visible and interpretable, which puts valuable information in the hands of researchers instead of obscuring it. Another benefit is that there are algorithms available for both wrists. Previous methods have typically been developed for just one wrist when using raw acceleration data from ActiGraph monitors. This limits their applicability in some cases, of which a key example is the nondominant wrist data from the most recent cycle of the National Health and Nutrition Examination Survey (NHANES). To our knowledge, there are only three other processing methods (two linear models and a random forest) that are potentially applicable to this raw acceleration data (35,36), and each method is subject to critical limitations. The linear model and random forest from Staudenmayer et al. (36) were developed using monitors on the dominant wrist, and thus, they are not applicable to the NHANES data. The linear model from Hildebrand et al. (35) is for data from monitors on the nondominant wrist, but the equation is unable to estimate SB and performed poorly in a recent independent validation (37). Thus, none of the currently available methods seem suitable for the NHANES data, and the present study's accelerometer-only wrist algorithms are relatively advantageous, despite their need for further validation.

As focus shifts to implementing the present study's algorithms in free-living, it is also necessary to address current limitations of the GT9X monitors. A primary concern is the battery life when the IMU is enabled. The device manual estimates that a full charge can last 24 h when all IMU sensors are enabled, assuming other configuration settings (e.g., display enabled, no wireless communication with device, primary accelerometer sampling at 30 Hz). Our own informal observations are consistent with these estimates. It is also possible to selectively enable the IMU sensors during device initialization. However, disabling all IMU components except the gyroscope only increases battery life by an average of 18 h (unpublished observations). Thus, there are currently barriers to using the IMU for long-term collection, but hardware and programming changes to extend the battery life are feasible. Lowering the IMU sampling rate from 100 Hz to a lower rate (e.g., 30 Hz) would be a simple yet fruitful change to make along this line, and further innovations in the sensor technology could eventually support a week of data collection.

Another concern when using the IMU is the volume of data it generates. This already presents a challenge with primary accelerometer data, but the effect is multiplied when adding the extra sensors in the IMU. With the full IMU enabled, more than 1 GB of data can accumulate in less than 24 h, on top of the data collected by the primary accelerometer. Activating only the gyroscope yields approximately 800–900 MB over the course of approximately 42 h, on top of the primary accelerometer data (unpublished observations). This presents obvious challenges for storing data, and it also affects the computational power required to process the data. For approximately 80 min of data, applying the algorithms can take several minutes when performed in R, most of which is spent reading in the data files, filtering the IMU data, and collapsing to 1-s epochs. Although a number of factors could cause processing time to vary widely, it will be an issue of continued importance as the GT9X becomes more widely used in free-living.

**Future work.** The present study was not focused on assessing the relative advantage of the different monitor attachment sites. However, the results demonstrate that ankle-worn devices can predict EE with a level of accuracy similar to hip-worn devices. This has important implications for continued work, because ankle-worn monitors are understudied compared with hip- and wrist-worn monitors. In this study, for either ankle, RMSE estimates (algorithms 1–3) were within 4.0% of the corresponding hip algorithm, and all ankle RMSE and MAPE were lower than the corresponding wrist

algorithms. Thus, greater attention to ankle placement is warranted in future studies.

Continued research should also focus on optimizing the algorithms for free-living by replicating this study's protocol in a more robust way. That is, the algorithms were developed in this study using a small set of structured activities that were performed in a laboratory setting with a small and homogenous sample. Studies involving larger sample sizes and more activities will be helpful in making the algorithms more generalizable, which will contribute to better performance in free-living.

**Strengths and limitations.** As noted, this study was limited by the small sample and the inclusion of only a few structured activities. However, continued research can address these issues. A key strength of this study is that it presents the first techniques for estimating EE from high-frequency IMU data.

## CONCLUSIONS

Two-regression algorithms that incorporate both accelerometer and gyroscope data produce substantially better EE estimates than algorithms that use only accelerometer data, when using data from hip- or ankle-worn monitors. The gyroscope is superior to the accelerometer for classifying SB and CWR. Magnetometer data do not contribute substantially to EE prediction. Future research should examine the effect of gyroscope data on PA classification and EE estimation using machine learning algorithms. In addition, greater attention to ankle-worn monitors is warranted.

This study was not funded.

The authors declare no conflicts of interest. The results of this study are presented clearly, honestly, and without fabrication, falsification, or inappropriate data manipulation. This study does not constitute endorsement by the American College of Sports Medicine.

## REFERENCES

1. Chen KY, Janz KF, Zhu W, Brychta RJ. Redefining the roles of sensors in objective physical activity monitoring. *Med Sci Sports Exerc*. 2012;44(1 Suppl 1):S13–23.
2. Troiano RP, McClain JJ, Brychta RJ, Chen KY. Evolution of accelerometer methods for physical activity research. *Br J Sports Med*. 2014;48(13):1019–23.
3. Aminian K, Najafi B. Capturing human motion using body-fixed sensors: outdoor measurement and clinical applications. *Comput Animat Virtual Worlds*. 2004;15(2):79–94.
4. Pärkkä J, Ermes M, Antila K, van Gils M, Manttari A, Nieminen H. Estimating intensity of physical activity: a comparison of wearable accelerometer and Gyro sensors and 3 sensor locations. In: *29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*; 2007: Lyon (France); 2007. pp. 1511–4.
5. Vathsangam H, Emken B, Schroeder E, Spruijt-Metz D, Sukhatme GS. Energy estimation of treadmill walking using on-body accelerometers and gyroscopes. In: *32nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society*; 2010 Sep 1–4: Buenos Aires (Argentina); 2010. pp. 6497–501.
6. Aminian K, Najafi B, Büla C, Leyvraz PF, Robert P. Spatio-temporal parameters of gait measured by an ambulatory system using miniature gyroscopes. *J Biomech*. 2002;35(5):689–99.
7. Najafi B, Aminian K, Loew F, Blanc Y, Robert PA. Measurement of stand-sit and sit-stand transitions using a miniature gyroscope and its application in fall risk evaluation in the elderly. *IEEE Trans Biomed Eng*. 2002;49(8):843–51.
8. Najafi B, Aminian K, Paraschiv-Ionescu A, Loew F, Bula CJ, Robert P. Ambulatory system for human motion analysis using a kinematic sensor: monitoring of daily physical activity in the elderly. *IEEE Trans Biomed Eng*. 2003;50(6):711–23.
9. Zhang Y, Markovic S, Sapir I, Wagenaar RC, Little TD. Continuous functional activity monitoring based on wearable tri-axial accelerometer and gyroscope. In: *5th International Conference on Pervasive Computing Technologies for Healthcare*; 2011 May 23–26: Dublin (Ireland); 2011. pp. 370–3.
10. Shoaib M, Bosch S, Incel OD, Scholten H, Havinga PJ. Fusion of smartphone motion sensors for physical activity recognition. *Sensors (Basel)*. 2014;14(6):10146.
11. Shoaib M, Scholten H, Havinga PJM. Towards physical activity recognition using smartphone sensors. In: *IEEE 10th International Conference on Autonomic and Trusted Computing*; 2014 Dec 18–20: Vietri sulMare (Italy); 2013. pp. 80–7.
12. Albinali F, Intille S, Haskell W, Rosenberger M. Using wearable activity type detection to improve physical activity energy expenditure estimation. In: *Proceedings of the 12th ACM International Conference on Ubiquitous Computing*; 2011 Sep 26–29: Copenhagen (Denmark); 2011. pp. 311–20.
13. Crouter SE, Kuffel E, Haas JD, Frongillo EA, Bassett DR. Refined two-regression model for the actigraph accelerometer. *Med Sci Sports Exerc*. 2010;42(5):1029–37.
14. Heil DP. Predicting activity energy expenditure using the Actical® activity monitor. *Res Q Exerc Sport*. 2006;77(1):64–80.
15. Crouter SE, Bassett DR. A new 2-regression model for the Actical accelerometer. *Br J Sports Med*. 2008;42(3):217–24.
16. Crouter SE, Clowers KG, Bassett DR. A novel method for using accelerometer data to predict energy expenditure. *J Appl Physiol*. 2006;100(4):1324–31.
17. Bonomi AG, Plasqui G, Goris AH, Westerterp KR. Improving assessment of daily energy expenditure by identifying types of physical activity with a single accelerometer. *J Appl Physiol (1985)*. 2009;107(3):655–61.
18. Dongwoo K, Kim HC. Activity energy expenditure assessment system based on activity classification using multi-site triaxial accelerometers. In: *29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*; 2007 Aug 23–26: Lyon (France); 2007. pp. 2285–7.
19. Ruch N, Joss F, Jimmy G, Melzer K, Hanggi J, Mader U. Neural network versus activity-specific prediction equations for energy expenditure estimation in children. *J Appl Physiol (1985)*. 2013;115(9):1229–36.
20. Thomas S, Reading J, Shephard RJ. Revision of the Physical Activity Readiness Questionnaire (PAR-Q). *Can J Sport Sci*. 1992;17(4):338–45.
21. McLaughlin JE, King GA, Howley ET, Bassett DR Jr, Ainsworth BE. Validation of the COSMED K4 $b^2$ portable metabolic system. *Int J Sports Med*. 2001;22(4):280–4.
22. Van Hees VT. *GGIR: Raw Accelerometer Data Analysis*. 2016 [cited 2017 Nov 15]. Available from: https://CRAN.R-project.org/package=GGIR.
23. Van Hees VT, Gorzelniak L, León ECD, et al. Separating movement and gravity components in an acceleration signal and

implications for the assessment of human daily physical activity. *PLoS One*. 2013;8(4):e61691.
24. Catalfamo P, Ghoussayni S, Ewins D. Gait event detection on level ground and incline walking using a rate gyroscope. *Sensors (Basel)*. 2010;10(6):5683–702.
25. ActiGraph, LLC. ActiGraph IMU White Paper. [Internet]. 2015 [cited 2017 Nov 15]. Available from: http://actigraphcorp.com/wp-content/uploads/2015/06/ActiGraph_IMU_White_Paper.pdf.
26. Ainsworth BE, Haskell WL, Herrmann SD, et al. 2011 Compendium of physical activities: a second update of codes and MET values. *Med Sci Sports Exerc*. 2011;43(8):1575–81.
27. Tremblay MS, Aubert S, Barnes JD, et al. Sedentary Behavior Research Network (SBRN)—Terminology Consensus Project process and outcome. *Int J Behav Nutr Phys Act*. 2017;14(1):75.
28. Matthews CE, Chen KY, Freedson PS, et al. Amount of time spent in sedentary behaviors in the United States, 2003–2004. *Am J Epidemiol*. 2008;167(7):875–81.
29. Chau JY, Merom D, Grunseit A, Rissel C, Bauman AE, van der Ploeg HP. Temporal trends in non-occupational sedentary behaviours from Australian Time Use Surveys 1992, 1997 and 2006. *Int J Behav Nutr Phys Act*. 2012;9(1):76.
30. Robin X, Turck N, Hainard A, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinform*. 2011;12:77.
31. Froud R, Abel G. Using ROC curves to choose minimally important change thresholds when sensitivity and specificity are valued equally: the forgotten lesson of Pythagoras. Theoretical considerations and an example application of change in health status. *PLoS One*. 2014;9(12):e114468.
32. Bao L, Intille SS. Activity recognition from user-annotated acceleration data. *Pervasive Comput*. 2004:1–17.
33. Mannini A, Intille SS, Rosenberger M, Sabatini AM, Haskell W. Activity recognition using a single accelerometer placed at the wrist or ankle. *Med Sci Sports Exerc*. 2013;45(11):2193–203.
34. Mathie MJ, Celler BG, Lovell NH, Coster AC. Classification of basic daily movements using a triaxial accelerometer. *Med Biol Eng Comput*. 2004;42(5):679–87.
35. 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.
36. 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 (1985)*. 2015;119(4):396–403.
37. Ellingson LD, Hibbing PR, Kim Y, Frey-Law LA, Saint-Maurice PF, Welk GJ. Lab-based validation of different data processing methods for wrist-worn ActiGraph accelerometers in young adults. *Physiol Meas*. 2017;38(6):1045–60.