J Appl Physiol 103: 1419–1427, 2007.
First published July 19, 2007; doi:10.1152/japplphysiol.00429.2007.

An artificial neural network model of energy expenditure using nonintegrated acceleration signals

Megan P. Rothney,[1] Megan Neumann,[2] Ashley Béziat,[2] and Kong Y. Chen[3]

[1]*Department of Biomedical Engineering, Vanderbilt University* [2]*Department of Medicine, Division of Gastroenterology, Vanderbilt University Medical Center, Nashville, Tennessee; and* [3]*National Institute of Diabetes and Digestive and Kidney Diseases/Clinical Endocrinology Branch, National Institutes of Health, Bethesda, Maryland*

Submitted 19 April 2007; accepted in final form 16 July 2007

Rothney MP, Neumann M, Béziat A, Chen KY. An artificial neural network model of energy expenditure using non-integrated acceleration signals. *J Appl Physiol* 103: 1419–1427, 2007. First published July 19, 2007; doi:10.1152/japplphysiol.00429.2007.—Accelerometers are a promising tool for characterizing physical activity patterns in free living. The major limitation in their widespread use to date has been a lack of precision in estimating energy expenditure (EE), which may be attributed to the oversimplified time-integrated acceleration signals and subsequent use of linear regression models for EE estimation. In this study, we collected biaxial raw (32 Hz) acceleration signals at the hip to develop a relationship between acceleration and minute-to-minute EE in 102 healthy adults using EE data collected for nearly 24 h in a room calorimeter as the reference standard. From each 1 min of acceleration data, we extracted 10 signal characteristics (features) that we felt had the potential to characterize EE intensity. Using these data, we developed a feed-forward/back-propagation artificial neural network (ANN) model with one hidden layer ($12 \times 20 \times 1$ nodes). Results of the ANN were compared with estimations using the ActiGraph monitor, a uniaxial accelerometer, and the IDEEA monitor, an array of five accelerometers. After training and validation (leave-one-subject out) were completed, the ANN showed significantly reduced mean absolute errors (0.29 ± 0.10 kcal/min), mean squared errors (0.23 ± 0.14 kcal^2/min^2), and difference in total EE (21 ± 115 kcal/day), compared with both the IDEEA ($P < 0.01$) and a regression model for the ActiGraph accelerometer ($P < 0.001$). Thus ANN combined with raw acceleration signals is a promising approach to link body accelerations to EE. Further validation is needed to understand the performance of the model for different physical activity types under free-living conditions.

physical activity; actigraph; IDEEA monitor; accelerometer; indirect calorimeter

IN THE PAST FIFTEEN YEARS, portable accelerometers have been used in the research field to characterize the intensity and duration of physical activity (PA), and their output has been used for the estimation of energy expenditure (EE) (8, 12, 14). Accelerometer devices are typically worn at the hip with the aim of capturing displacement of the subject's center of mass, which is generally associated with moderate- to high-intensity activities, and accounts for the largest EE differences from baseline. For data to be collected for more than 1 day, long enough to assess the patterns of PA in free-living individuals, most accelerometery-based PA monitors have historically reported one data point per minute, which represents the summation of acceleration events during the minute. The output of PA monitors is reported to the investigator in units of activity counts (26), which are an arbitrary unit specified by each device manufacturer. To our knowledge, this integration (or summation) process was not designed a priori for estimating EE; rather, it was dictated by memory capacity and battery life in early accelerometers while still giving investigators an intensity scale to rate PA. It is therefore possible that other characteristics of raw acceleration signals may yield better predictive outcomes.

Early modeling approaches relating activity counts and EE typically assumed a linear relationship between the activity count values and EE measured using indirect calorimeters (12, 15, 18, 24). Linear regression fits were used because of their computational simplicity and ability to well characterize the energy costs of moderate intensity, ambulatory activities (walking and jogging). Although models based on this strategy provided an excellent first approximation of the relationship between acceleration signals and EE, they have suffered in their generalization to different PA types and subject populations (27). This is because the models were predominantly developed using short protocols containing set paces of dynamic PA and were developed on homogeneous subject populations. Estimation accuracy of generalized linear models also varies greatly between subjects with different personal characteristics (for example, age, height, body mass) because identical accelerations may not result in the same metabolic costs for these individuals, although the activity count values may be the same.

Several investigators have sought to improve model accuracy by increasing the amount of information gathered during each measurement epoch. This effort has included adding additional acceleration dimensions at the hip (8, 20), adding sensors to the limbs (wrist and ankle) for more complete movement detection (8, 15), and coupling physical and physiological information, such as heart rate, near body temperature, and skin impedance (6, 16, 22). With the use of the additional data collected by these devices, more mathematically sophisticated and in some cases more accurate models relating acceleration and energy expenditure have been developed, such as multiple linear regressions (14) and generalized nonlinear models (8, 21). Recently, a new model for EE estimation was developed that called for recording data in finer time intervals (1 s) using a uniaxial accelerometer rather than collecting more channels or types of sensor data (10). In this model, the minute-by-minute coefficient of variability (CV) was computed with 10-s data segments. This CV was used as

Address for reprint requests and other correspondence: M. P. Rothney, 10 Center Drive MSC 1613, 10 CRC\6-3940, Bethesda, MD 20892 (e-mail: rothneym@niddk.nih.gov).

The costs of publication of this article were defrayed in part by the payment of page charges. The article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Table 1. *Characteristics of study participants*

	All Subjects (n = 102)	Men (n = 46)	Women (n = 56)
Age, yr	38.6±13.1 (19–69)	38.3±12.9 (20–69)	38.9±13.4 (19–67)
Height, m	1.71±0.09 (1.52–1.91)	1.78±0.06 (1.67–1.91)	1.64±0.06 (1.52–1.78)
Weight, kg	75.7±16.4 (48–120)	83.4±14.0 (64–120)	69.3±15.5 (48–114)
BMI, kg/m^2	26.0±5.3 (16.9–42.1)	26.2±4.2 (19.8–38.5)	25.9±6.1 (16.9–42.1)
%Body Fat	29.7±11.9 (6.2–57)	22.6±8.7 (6.2–45.1)	35.7±10.8 (11.7–57)

Values are means ± SD, with total range in parentheses. BMI, body-mass index.

an initial discrimination tool to determine which of two nonlinear models should be applied to the minute of data. This modeling approach was made possible in part because of improvements in the data storage capacity and battery life of modern accelerometers. Increases in the amount of data acquired from each minute of PA open the field to new analytic solution techniques that rely on multiple measurements acquired from each minute of measured activity.

By increasing the number of acceleration samples per minute, more analytically sophisticated approaches, relying on automated pattern recognition and machine learning, have been applied to several aspects of PA monitoring. The majority of this work has focused on identifying postures (7), locations within a finite space (17), or PA types (3, 19). High probabilities of correct identifications have been shown for several PA types and activity contexts. It has also been shown that the speed and incline of self-paced free-living walking can be estimated with the use of data acquired from treadmill walking using similar analytic frameworks (2). To our knowledge, however, no group has used accelerometer data, coupled with machine-learning algorithms, to predict minute-by-minute EE.

The purpose of this study was to expand on existing EE modeling techniques by capturing raw (32 Hz) acceleration signals from a biaxial accelerometer worn at the hip. We propose a feature extraction scheme where the dense acceleration signals are reduced to a small number of simple to compute statistical parameters (features) that are well correlated with the minute-by-minute EE measured by a whole room indirect calorimeter. The reduced signal information and subject demographics (sex, age, height, weight, body-mass index, and racial/ethnic background) were used to develop an artificial neural network (ANN) model to estimate minute-by-minute EE. Results of the ANN model were compared with both a traditional accelerometer regression equation and the proprietary output of a commercially available accelerometer array.

METHODS

Participants

One hundred and two healthy adults (46 men, 55 women) between the ages of 18 and 70 years completed this study. Subjects were free of both diseases and medications known to alter metabolic rate and major orthopedic limitations and were nonsmokers. The characteristics of these subjects are shown in Table 1.

Experimental Procedures

Volunteers were recruited from the middle Tennessee area using flyers, e-mail distribution lists, and personal contact. Before participation, all subjects signed an informed consent document approved by the Vanderbilt University Committee for the Protection of Human Subjects. Each subject was asked to stay in the room calorimeter for ~24 h while minute-by-minute activity data were acquired with multiple commercially available accelerometery-based PA monitors. Each subject was asked to engage in two structured activity intervals. The morning activity period included self-paced walking and jogging (both in the room and on the treadmill), and the afternoon activity period contained sedentary activities, such as deskwork, along with stationary biking (Fig. 1). Each prescribed activity was performed for 10 min followed by a 10-min rest period to allow the metabolic rate to return to baseline between intervals and to allow post hoc discrimination between activity types. During times when no activity was prescribed, subjects were encouraged to engage in their normal daily PA routine as much as possible. Subject's height and body mass were measured on the morning of the study visit.

Instrumentation

Activity-energy measurement system. EE was computed on a minute-by-minute basis by the Vanderbilt University room calorimeter, which is located within the Vanderbilt General Clinical Research Center. This system measures oxygen consumption and carbon dioxide production with high accuracy (system error of <1%). The room calorimeter is an air-tight environmental room measuring 2.5 × 3.4 × 2.4 m. The calorimeter is equipped with a toilet and sink, desk, chair, telephone, television, DVD player, stereo system, bed, treadmill, and

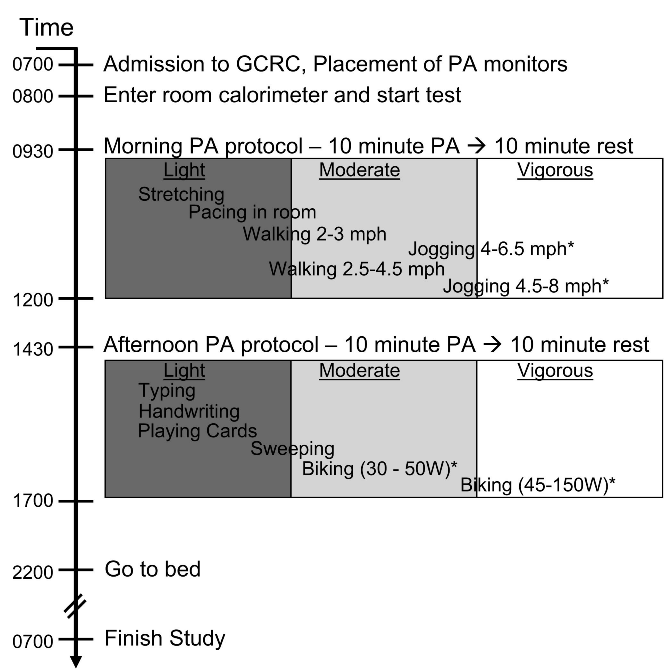


Fig. 1. Protocol for the metabolic chamber stay. *Intervals in which intensity was recommended. However, subjects were asked to self-pace for these intervals. GCRC, Vanderbilt General Clinical Research Center; PA, physical activity.

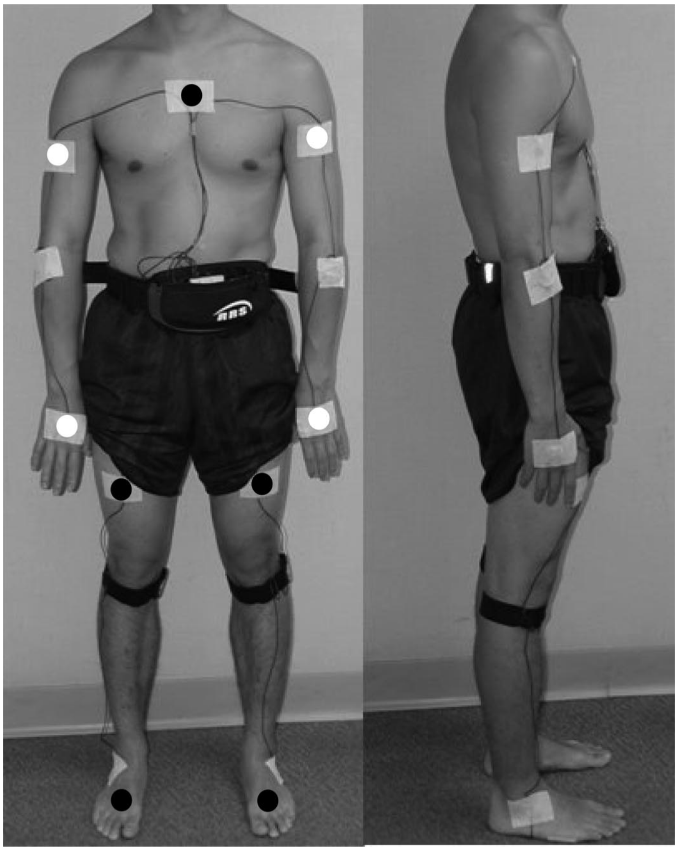


Fig. 2. Schematic of IDEEA sensor locations. Sites shown in black represent the accelerometer sites for the commercial system, and sites in white are the added custom sensor sites.

exercise bike. Although the calorimeter floor contains a force plate and the room has several event markers, information from these systems was not utilized for these experiments. Technical details of the calorimeter have been previously reported (23).

Accelerometers. Subjects were outfitted with both the ActiGraph (Fort Walton Beach, FL) uniaxial accelerometer and a custom-designed activity monitor, which is a derivative of the commercially available IDEEA (MiniSun, Fresno CA) monitor. The commercial IDEEA monitor consists of an array of five accelerometers (20 × 15 × 4 mm, 2 g) attached to the skin via hypoallergenic tape at the sternum, midthigh, and bottom of each foot. Each sensor is wired to a hip pack that serves to synchronize the signals from each channel and store the data. Although high accuracy for the IDEEA PA type identification routine has been published (29), the study designed to validate the EE estimation routine contained only walking, jogging, and lying down (28), and the EE estimation approach has therefore not been subjected to a rigorous validation in estimation of EE associated with other PA types.

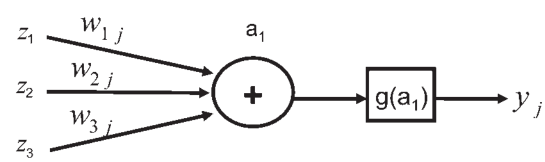
The custom IDEEA monitor used in these experiments includes all of the sensors from the original configuration but adds recording capability at the hip pack (biaxial, anterior/posterior, and medial/lateral), on each upper arm (uniaxial), and on the top of each hand (biaxial). Raw data (32 Hz) are collected at each of the custom sites with data reported separately for each axis of each sensor, and integrated signals are recorded by the original IDEEA sensors (Fig. 2). In this configuration, data can be acquired continuously throughout our study visits (~21 h). To our knowledge, none of the commonly used commercially available PA monitors can record raw data for this length of time. Both the ActiGraph and the IDEEA hip pack were worn on a snug elastic band with both monitors located at the right hip. For this study, we used only the raw data from the hip sensors (biaxial) for analyses. Because most investigators only collect data at the hip and it would be ideal to collect field data from only one site to minimize the inconvenience to the subject, we felt it was important to explore model developments that could be applied to traditional hip-mounted accelerometers before expanding our study goals to include multisite analysis. Stationary biking was removed from the analysis for all monitors.

The ActiGraph is a uniaxial accelerometer that has been in widespread use for more than a decade. Device specifications have been published elsewhere (26). Activity counts were analyzed with a combination of the Freedson equation for moderate- to high-intensity data and the work-energy theorem for low-intensity values (AG_{FW}). This analysis construct is presented in the ActiGraph instruction manual as a way to use the Freedson equation on all intensities of PA data (1, 12). This equation was chosen because it was developed for 1-min epoch data, and the ActiGraph software is equipped to compute these estimations, making them readily available to all researchers.

Modeling Approach

ANN modeling was selected to relate the features of the raw acceleration signal to measured EE on a minute-by-minute basis. ANN modeling is an information-processing paradigm inspired by the way the densely interconnected, parallel structure of the mammalian brain processes information (13). Models are developed using a learning process in which a series of connection weights, analogous to synapses, are tied to a series of processing elements, analogous to neurons. Because the ANN is presented with input-output pairs, the weight values are adjusted until an optimal solution, in our case prediction of minute-by-minute EE, is achieved. ANN is a good candidate model when there are a large number of inputs for a small number of outputs or when the ideal functional form of the solution is not known (4).

To implement ANN, we begin by specifying the number of inputs (acceleration or subject characteristic terms), the number of weight values (interactions between the terms), the architecture of the model, and the number and type of output parameters (EE, a single continuous variable) (Fig. 3). A single neuron receives multiple inputs, which represent characteristics of the acceleration signal or the subjects themselves. The relative importance of each input is specified by a weight value. A single neuron is not capable of solving difficult problems because it may not allow for all required nonlinearities or



Variable	Definition	Meaning
z	Input	Characteristics of acceleration signal
w	Weight	Relative importance of each input
g(a)	Transfer Function	Scales input for next layer
Y	Output	Predicted output (EE, MET, PA Type, etc)

Fig. 3. A single artificial neuron. The inputs, which represent characteristics of the acceleration signals, are multiplied by weight vectors, which represent their strengths. The summation of these quantities is used to estimate an outcome, in our case, energy expenditure (EE). MET, metabolic equivalent.

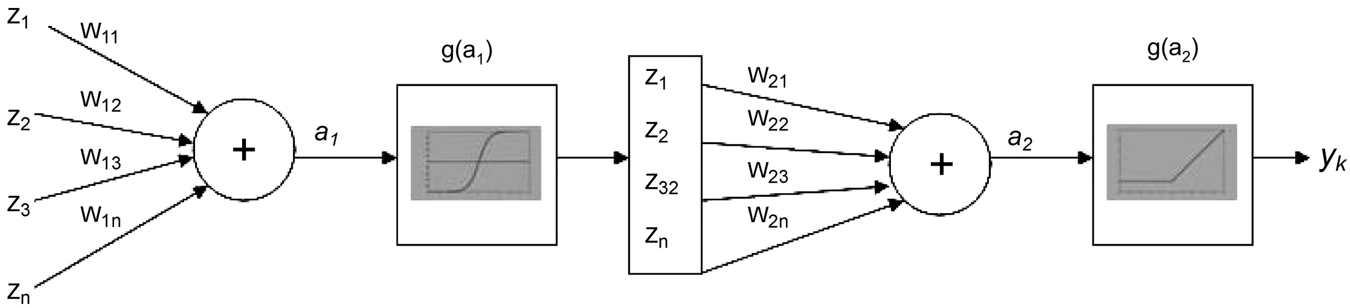


Fig. 4. Schematic for the feed-forward network, which includes multiple neuron units formed into computational layers. Each input (*z*) is multiplied by the layer weights (*w*) and sent through a summation (*a*) and transformation [*g*(*a*)] to the next layer. The process is repeated until the output (*y*) is reached. The first basis function shown is a hyperbolic tangent function, whereas the second is a positive linear basis function.

interaction terms, so multiple neurons are arranged into computational layers linked by transfer functions.

The process by which the model derives a solution is referred to as training because the model is presented with examples for which a known measurement value exists. To begin the training process, each of the network weight values is assigned a small, random value. To improve the estimation error, ANN models "learn" using a feed-forward/back propagation approach. The feed-forward process consists of computing the transformations between the inputs and outputs as specified by the current weights and transfer functions (Fig. 4) for each minute of data. The estimation from the forward pass through the network is compared with the measured value, and an error signal is computed. This gradient (rate of change) of the error is used to iteratively adjust the weights. The process of updating the weights constitutes the back-propagation step because, procedurally, it occurs from the output end to the input end. Back propagation is an optimization procedure and can be governed by most standard optimization approaches (9). Because the same inputs are shown to the network many times and the weights are allowed to adapt, the error signal decreases, and an optimal set of weights can be realized. Training is allowed to continue until a specified error tolerance has been achieved. Once optimized weight values have been realized, EE estimation for novel acceleration inputs consists of a forward pass through the network. Because the forward step consists only of matrix multiplications and summations, the end used is spared the bulk of the computational costs. The computational time for estimating a day of EE values using a small to moderately sized network would be $<$1 s.

For this problem, a 12-node input layer, followed by a 20-node hidden layer, and a single node output layer architecture were used. Hyperbolic tangent basis functions were chosen for the input and hidden layer because our data were standardized (mean = 0; variance = 1) before training to reduce magnitude bias in our data. The number of nodes in each layer was empirically determined by a pilot sample of 10 subjects. Values of 5–25 were tested for the input layer, and 5–30 were tested for the hidden layer. Error reductions past the chosen size (12 input, 20 hidden) did not seem to confer significant benefits in error reduction in our pilot sample and were therefore not used in the final development. Training was performed with a gradient descent training function with a learning rate of 0.01.

Validation was performed by leave-one-subject-out cross validation. In this approach, the total data were divided into a training set (n = 101) and a testing set (n = 1). The training set data were used to optimize the model to the estimation of minute-by-minute EE, whereas the test set data, which were not used to derive the model, were used to assess the performance of the model on new data (25). This process was repeated 102 times so that the model performance on each subject's data could be assessed. For each validation step, training ended when the error on the validation set failed to decrease by more than 1*e*-6 per iteration (after an initial drop), the error gradient fell below 1*e*-6, or 5,000 iterations were reached.

Feature extraction. Feature extraction is the key step in preparing raw data for ANN modeling. The purpose of this step is data reduction. In this study, 1,920 (32 samples/s $\times$ 60 s/min) data points are collected by each IDEEA sensor channel for each minute of study data collected. These values all correspond to a single measurement made by the indirect calorimeter. This amount of information quickly becomes cumbersome to analyze; however, more importantly, redundant information is likely contained in the acceleration signals. It is therefore vital that the raw data are reduced into a small number of parameters that carry the most relevant information. We chose to reduce the data into a series of parameters that we felt were both statistically relevant and physically meaningful. Eleven parameters were extracted from each channel of raw data [median, integral, peak intensity, interquartile interval, skew, kurtosis, peak CV over any 10 s of data, lowest 10 s CV, mean absolute error (MAE), and the summation of signal power above 0.7 Hz, and sum of signal power below 0.7 Hz]. The signal power cutoff of 0.7 Hz was determined based on optimizing the division in the power spectral density between walking and sedentary tasks in a sample of 10 subjects not used for model development. The 11 computed acceleration parameters were then analyzed based on their correlations with one another to eliminate redundant information. This step reduced the inputs to five for each hip sensor channel. These consisted of the peak value, the interquartile interval, the lowest coefficient of variability when each minute of data was analyzed in 10-s increments, the sum of the signal power below 0.7 Hz, and the sum of the signal power above 0.7 Hz.

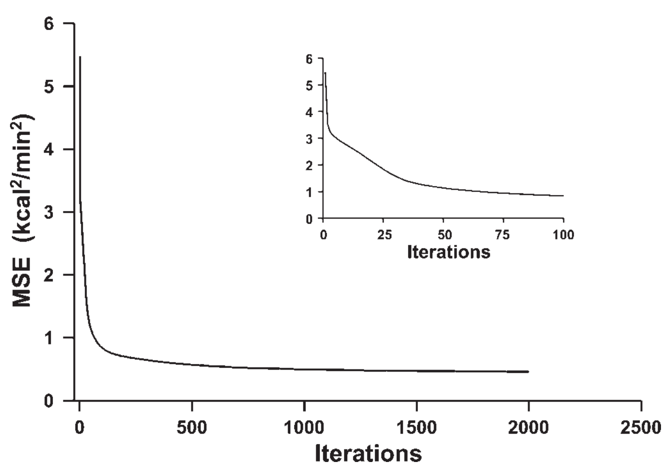


Fig. 5. Error decay for a representative model training for the first 2,000 iterations. MSE, mean square error. *Inset*: error over the first 100 iterations. Error begins at a high value because weight values are randomly assigned and reduced quickly as the model learns the correct estimation for frequently represented data.

These data were joined in the input set by the subjects' sex, age, height, body mass, and ethnic background because these features have been shown to impact resting metabolic rate and can be easily measured or self-reported (11).

Feature extraction was designed such that the final inputs to the model are quantities researchers are generally familiar with (at least conceptually). Additionally, by using a small number of easily computed data features, the storage requirements for any future activity monitors would be minimized because raw data would not need to be stored, only the relevant computed parameters. This effectively minimizes the amount of internal storage capacity required of the accelerometer while maintaining the quality of information derived from the raw signal. Model development and feature extraction were performed with Matlab 7.01 (Mathworks, Natick, MA).

Statistical Analysis

Data are presented as means, standard deviation, and total range. Models (AG_{FW}, proprietary IDEEA, ANN) were compared on a per subject basis according to the MAE (*Eq. 1*), the mean squared error (MSE) (*Eq. 2*), the absolute percent difference between each model and the measured total energy expenditure (TEE), and the squared Pearson's correlation coefficient (r^2) for each subject over the entire study duration, using ANOVA with post hoc Tukey tests. Bland-Altman plots (5) were used to examine trends in total EE estimation relative to the calorimeter.

$$\mathrm{MAE} = \frac{1}{N}\sum_{i=1}^{N}\left|(x_i - \mu)\right| \tag{1}$$

$$\mathrm{MSE} = \frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2 \tag{2}$$

RESULTS

Approximately 112,000 minutes of data were used to train the ANN model. Convergence of the training sets occurred after an average of 3,328 iterations with a range of 475–5,000 iterations (5,000 was the maximum number of iterations allowed for this experiment). The training error as a function of the number iterations showed an exponential decay profile. An error profile from a randomly selected subject is shown in Fig. 5. Although the rates of decay may change between training

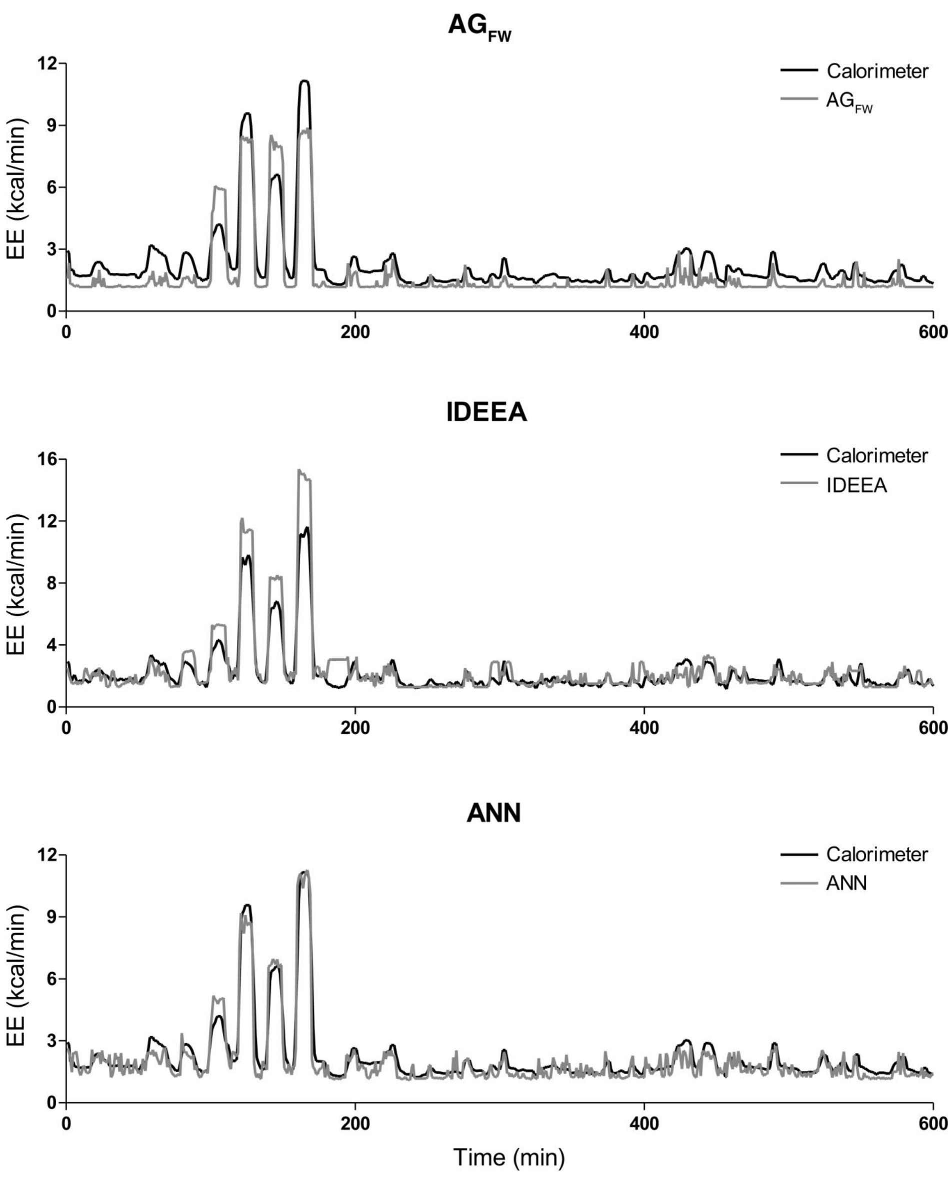


Fig. 6. Minute-by-minute EE estimation using ActiGraph (AG_{FW}), IDEEA, and artificial neural network (ANN). This subject was a 43-yr-old male (height of 1.69 m, weight of 76.9 kg, body-mass index of 26.7 kg/m^2). The ANN has improved estimation in the high-intensity intervals (walking and jogging on a treadmill) and good detection of the lower-intensity features of the data.

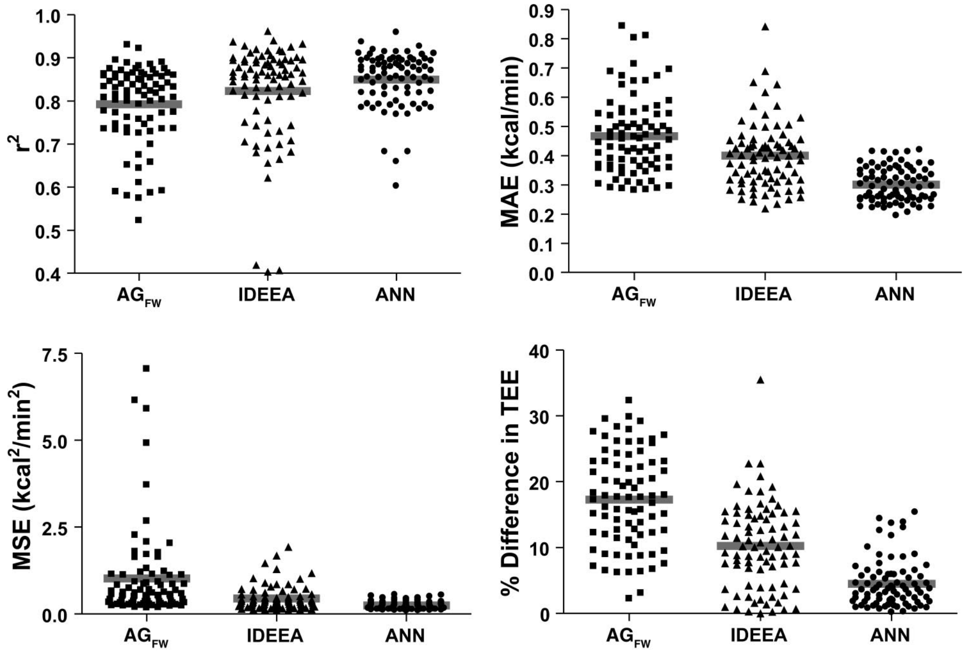


Fig. 7. Results of summary analysis of ActiGraph (AG_{FW}), IDEEA, and ANN models. The correlation coefficient (r^2) was higher in the ANN than in the IDEEA or AG_{FW} model, whereas mean absolute error (MAE), MSE, and absolute percent difference in total energy expenditure (TEE) was reduced. Improvements were noted in both the mean value and the variance of the estimations.

sets and some training sets may have long plateaus early in the training if the initial weight values place the model in an area of space where the gradient is small, all eventually exhibit an exponential profile where the bulk of the error decay occurs over ~50 iterations (Fig. 5, *inset*). The number of iterations to solution convergence can be altered through the model learning rate, which was set at 0.01 for this model. A linear regression was performed on the MSE vs. number of iterations to convergence. The slope of this regression was not significantly different from zero ($P = 0.9214$), suggesting that the number of iterations to model convergence did not significantly impact the final error.

Although the ANN was developed on 102 subjects and all subjects wore both the IDEEA monitor and an ActiGraph, only 81 subjects had both IDEEA data (used for ANN development and IDEEA model results) and ActiGraph data as a result of a software problem with the ActiGraph monitors, which has subsequently been corrected. All results shown are based on paired comparisons of the subjects who had both IDEEA and ActiGraph data. A representative subject's minute-by-minute EE estimation showed that the ANN was able to characterize the baseline EE. Additionally, most of the programmed activity intervals and spontaneous PA were well characterized, although some small errors did occur (Fig. 6).

Summary statistical measures, the r^2, MAE, MSE (using minute-by-minute EE estimations), and absolute percent difference between the chamber measured and model estimated TEE were computed for the AG_{FW}, IDEEA, and ANN models on each of the 81 subjects (Fig. 7). All models showed, on average, high correlation with measured minute-by-minute EE. ANOVA revealed that the r^2 was higher ($P < 0.001$) in the ANN model relative to the other models and that both the mean of the testing set MAE and MSE were significantly reduced in the IDEEA model relative to AG_{FW} ($P < 0.001$), with further reductions ($P < 0.01$) also seen when the ANN was compared with the IDEEA monitor proprietary model. The absolute percent difference between TEE measured by the room calorimeter and that estimated by each model was computed to account for differences in the number of measurement minutes and PA intensities represented in each subject's data. Analysis of the percent difference revealed a drop in percent error in the IDEEA relative to AG_{FW}, and a further decrease in error when the ANN was compared with the IDEEA model (Table 2).

Bland-Altman plots were used to characterize the ability of each model to estimate TEE during the measurement period (Fig. 8). AG_{FW} showed an average difference of -355 ± 240 kcal/day. There was also a significant trend ($P = 0.0351$) toward underestimation of TEE as the absolute value of TEE increased. The IDEEA model demonstrated a mean difference of 230 ± 209 kcal/day. There was a trend ($P < 0.001$) toward overestimation as TEE increased. The ANN shows a mean difference of 21 ± 115 kcal/day. No significant trend ($P = 0.86$) was observed as a function of magnitude of daily EE.

Table 2. *Comparison of ActiGraph, IDEEA, and ANN model performance assessed using summary error statistics for 81 subjects*

	r^2	MAE	MSE	%Difference
AG_{FW}	0.79±0.09 (0.55–0.93)	0.47±0.13 (0.28–0.84)	1.02±1.31 (0.18–7.04)	17.27±7.20 (2.20–32.32)
IDEEA	0.82±0.11 (0.40–0.96)	0.40±0.11 (0.22–0.84)	0.45±0.35 (0.11–1.93)	10.23±6.54 (0.02–35.53)
ANN	0.85±0.06 (0.60–0.96)	0.30±0.06 (0.19–0.42)	0.25±0.11 (0.09–0.55)	4.47±3.62 (0.13–15.35)

Values are means ± SD, with total range in parentheses. ANN, artificial neural network; MAE, mean absolute error; MSE, mean square error. ANOVA for repeated measurements was significant for all error metrics ($P < 0.05$) except for MSE between IDEEA and ANN model.

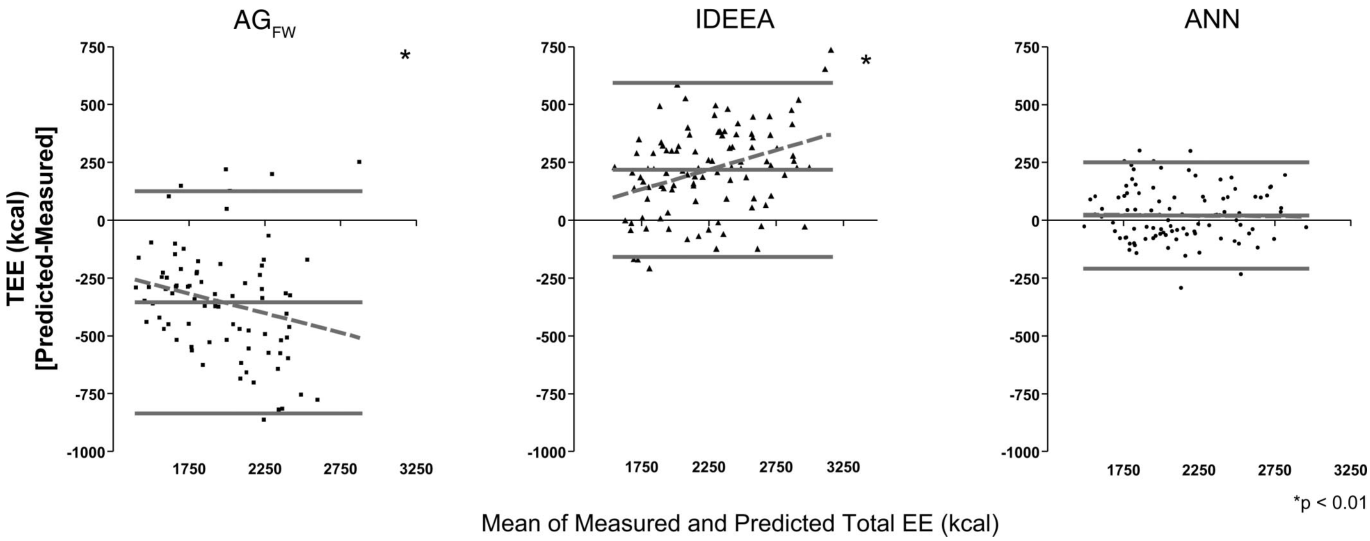


Fig. 8. Bland-Altman plots of TEE using AG_{FW}, IDEEA, and ANN models. Both the mean difference between measures and the standard deviation over the subject population show improvements in the ANN relative to the IDEEA and AG_{FW} models. Both the AG_{FW} and the IDEEA models also have significant slope trends ($^*P < 0.05$), whereas no significant trend was detected for the ANN model.

DISCUSSION

In this study, we investigated a procedure of extracting features from raw acceleration data collected by hip-worn accelerometers and development of a minute-by-minute EE estimation model based on ANN techniques. This experiment was developed as a proof of concept to show that, by using features of the acceleration signal other than the integral and combining these features with a flexible, high-dimensional modeling approach, estimation errors for minute-by-minute EE can be reduced, while also reducing the intersubject variability in estimation accuracy, which tends to be high in generalized models. The resulting network model showed a reduction in MAE, MSE, and percent difference in TEE when compared with both a uniaxial accelerometer using integrated signals (AG_{FW}) and the proprietary model used in the IDEEA monitor, a five-sensor array. In addition to reductions in mean errors for the group, the variance was also reduced, which further suggests the robustness of the ANN modeling approach.

Although the ActiGraph, proprietary IDEEA model, and the ANN model all had high correlation with the TEE on a minute-by-minute basis, this is more reflective of the capability of the accelerometer to determine whether any motion is present, rather than the model accurately reflecting minute-by-minute EE intensities. When measures that reflect the magnitude of the EE differences observed on a minute-by-minute basis (MAE or MSE) are considered, significant reductions in error were observed in the IDEEA model relative to AG_{FW} and the ANN relative to both of the other models. The cumulative effect of these error reductions can be observed in the absolute percent difference between measured and estimated TEE. There is reduction of nearly 13% between AG_{FW} and the ANN and 5% between the IDEEA monitor and the ANN. The mean of the difference in TEE was also greatly reduced by the ANN relative to the IDEEA and AG_{FW} with the mean difference in the ANN model being only 21 kcal, suggesting that the model has corrected for some of the baseline offset (resting EE) problems that have been previously observed using the IDEEA model. In the future, the ANN approach should be compared with other nonlinear EE estimation approaches to test the capabilities of the ANN relative to more sophisticated modeling approaches using the ActiGraph.

One of the biggest challenges involved in generalized modeling with accelerometers has been the large standard deviation of estimations between subjects. In this study, we found high individual estimation errors by AG and IDEEA compared with the measured EE [95% CI (−835 to 125) and (−187 to ∼647) kcal/day, respectively]. These values are beyond the treatment effect that we typically target in sustainable weight loss interventions, which is 100–250 kcal/day. Thus reducing this measurement error has crucial clinical implications. Large variability in performance across subjects may be related to the fact that standard regression approaches do not have sufficient flexibility to alter estimations when the same acceleration count value is achieved by a subject whose personal characteristics are different from those used for model development. The standard deviation in the TEE observed with the ANN, which allows interactions between characteristics and acceleration terms, showed a reduction of ∼50% relative to AG_{FW} and nearly 45% relative to the proprietary IDEEA model. The IDEEA uses similar characteristics to those used in the ANN and still exhibited a higher variability.

ANN has a number of attractive features for the energy expenditure estimation problem such as the flexibility of estimations across subjects, allowing interactions between all input terms, and its ability to map multiple inputs (acceleration terms) to a single output (EE) without prespecifying a functional form (linear, logistic, and so forth). The major disadvantages of ANN approaches are the computational complexity of the models that may require long training times and a relatively large number of free parameters. Additionally, model training requires a large number of labeled examples, acceleration data from a diverse sample of PA types, for which the EE is known, which requires a long data collection period before models can be developed. These models are also more difficult to disseminate to potential users, which would require developers to create macros or other simple-to-use tools for distribution. This is in contrast to generalized linear and nonlinear models, which can easily be presented in a manuscript and implemented by most researchers.

Although a classical ANN may seem like a black box solution technique, we attempted to minimize this appearance by carefully selecting model inputs that make sense in the context of the EE estimation problem. We chose terms that represent the magnitude and frequency of movements and the variability in motion patterns, which may be characteristic of certain PA types. This feature extraction process does require the model developer to make decisions about what data features may be of interest. Alternately, feature extraction can be performed by a standard data reduction technique such as principal component analysis, which achieves data reduction by combining parameters that are linearly related. This process maximizes the amount of the information from the data while eliminating repetitious measurements. The advantage of this technique is its capability to succinctly and consistently reduce data to the desired proportion of the total data variance. The disadvantage is that the reduced data are not composed of characteristics that would be familiar to researchers; rather, the data contain features representing agglomerations of measurements.

Perhaps the most challenging aspects of model development are collecting appropriate model training data and validation. Because there are literally hundreds of modes of PA that individuals may engage in and at least that many profiles for subjects' metabolic response to exercise, models will tend to generalize best to data sets composed of activities similar to those that were used for the original model development. We have attempted to mitigate this factor by *1*) asking subjects to self-pace activities and *2*) capturing spontaneous bouts of PA. These two steps allow for the collected data to be both diverse in intensity composition and representative of the activity patterns our subjects would normally engage in. To attempt to minimize the potential error increases associated with applying our model to new subjects (generalization errors), a leave one subject out cross-validation procedure was selected. This technique allows the bulk of the collected data to be used in model development relative to a split sample validation where a much larger percentage of the total data is withheld. The data from the validation sample may include unique features that would have affected the model development had they been available, so it is desirable to use as much data as possible for model training. The model presented here, however, is meant only to prove that, in principle, a high dimensional modeling approach such as ANN coupled with feature extraction from raw (32 Hz) acceleration signals can be used for EE estimation on a minute-by-minute basis, and the specific weighting coefficients should not be viewed as final. A split sample validation should also be implemented once a larger data sample has been collected to more independently characterize the model performance.

In conclusion, accelerometers have long been considered a promising tool for estimating EE due to their relatively low price, ease of use, and ability to record for many days at a time. This potential has not been fully met to date because of limitations in our ability to relate the output variables from the monitors to EE. Collecting raw acceleration data has the capability of improving the precision of EE estimation by allowing researchers the flexibility to identify relevant parameters during the feature extraction phase as well as opening the field to high dimensional modeling techniques such as ANN, which have the capability of generating more flexible estimations than more traditional modeling techniques. This study has shown a proof of concept that, by applying feature extraction and ANN models to biaxial acceleration data acquired at the hip, minute-by-minute and total EE estimations can be improved. Additional subjects and modes of PA should be acquired to both validate the current model and for use in developing a more robust algorithm.

ACKNOWLEDGMENTS

The authors thank Jessica White and Sarah Johnson for contributions to the data collection and Dr. Maciej Buchowski for continued support of this project.

GRANTS

This work was supported by National Institutes of Health Grants DK-069465, HL-082988, DK-02973, and RR-00095.

REFERENCES

1. **ActiGraph.** *Actisoft Analysis Software 3.2 User's Manual*. Fort Walton Beach, FL: MTI Health Services, 2005.
2. **Aminian K, Robert P, Jequier E, Schutz Y.** Incline, speed, and distance assessment during unconstrained walking. *Med Sci Sports Exerc* 27: 226–234, 1995.
3. **Bao L, Intille S.** Activity recognition from user-annotated acceleration data. *Lecture Notes Comp Sci* 3001: 1–17, 2004.
4. **Bishop C.** *Neural Networks for Pattern Recognition*. New York: Oxford, 1995.
5. **Bland JM, Altman DJ.** Regression analysis. *Lancet* 1: 908–909, 1986.
6. **Brage S, Brage N, Franks PW, Ekelund U, Wareham NJ.** Reliability and validity of the combined heart rate and movement sensor Actiheart. *Eur J Clin Nutr* 59: 561–570, 2005.
7. **Bussman JB, Martens WL, Tulen JH, Schasfoort FC, Berg-emons HJ, Stam HJ.** Measuring daily behavior using ambulatory accelerometry: the activity monitor. *Behav Res Methods Instr Comp* 33: 349–356, 2001.
8. **Chen KY, Sun M.** Improving energy expenditure estimation by using a triaxial accelerometer. *J Appl Physiol* 83: 2112–2122, 1997.
9. **Chong E, Zak S.** *An Introduction to Optimization*. New York: Wiley, 2001.
10. **Crouter SE, Clowers KG, Bassett DR Jr.** A novel method for using accelerometer data to predict energy expenditure. *J Appl Physiol* 100: 1324–1331, 2006.
11. **Donahoo WT, Levine JA, Melanson EL.** Variability in energy expenditure and its components. *Curr Opin Clin Nutr Metab Care* 7: 599–605, 2004.
12. **Freedson PS, Melanson E, Sirard J.** Calibration of the Computer Science and Applications, Inc., accelerometer. *Med Sci Sports Exerc* 30: 777–781, 1998.
13. **Haykin S.** *Neural Networks: A Comprehensive Foundation*. Upper Saddle River, NJ: Prentice Hall, 1999.
14. **Heil DP.** Predicting activity energy expenditure using the Actical activity monitor. *Res Q Exerc Sport* 77: 64–80, 2006.
15. **Hendelman D, Miller K, Baggett C, Debold E, Freedson P.** Validity of accelerometry for the assessment of moderate intensity physical activity in the field. *Med Sci Sports Exerc* 32: S442–S449, 2000.
16. **Jakicic JM, Marcus M, Gallagher KI, Randall C, Thomas E, Goss FL, Robertson RJ.** Evaluation of the SenseWear Pro Armband to assess energy expenditure during exercise. *Med Sci Sports Exerc* 36: 897–904, 2004.
17. **Lee SW, Mase M.** Activity and location recognition using wearable sensors. *Pervasive Computing* 1: 24–32, 2002.
18. **Nichols JF, Morgan CG, Chabot LE, Sallis JF, Calfas KJ.** Assessment of physical activity with the Computer Science and Applications, Inc., accelerometer: laboratory versus field validation. *Res Q Exerc Sport* 71: 36–43, 2000.
19. **Pober DM, Staudenmayer J, Raphael C, Freedson PS.** Development of novel techniques to classify physical activity mode using accelerometers. *Med Sci Sports Exerc* 38: 1626–1634, 2006.
20. **Rowlands AV, Thomas PW, Eston RG, Topping R.** Validation of the RT3 triaxial accelerometer for the assessment of physical activity. *Med Sci Sports Exerc* 36: 518–524, 2004.

21. **Schmitz KH, Treuth M, Hannan P, McMurray R, Ring KB, Catellier D, Pate R.** Predicting energy expenditure from accelerometry counts in adolescent girls. *Med Sci Sports Exerc* 37: 155–161, 2005.
22. **Strath SJ, Brage S, Ekelund U.** Integration of physiological and accelerometer data to improve physical activity assessment. *Med Sci Sports Exerc* 37: S563–S571, 2005.
23. **Sun M, Reed GW, Hill JO.** Modification of a whole room indirect calorimeter for measurement of rapid changes in energy expenditure. *J Appl Physiol* 76: 2686–2691, 1994.
24. **Swartz AM, Strath SJ, Bassett DR Jr, O'Brien WL, King GA, Ainsworth BE.** Estimation of energy expenditure using CSA accelerometers at hip and wrist sites. *Med Sci Sports Exerc* 32: S450–S456, 2000.
25. **Theodordis S, Koutroumbas K.** *Pattern Recognition.* New York: Elsevier Academic, 2003.
26. **Tryon WW, Williams R.** Fully proportional actigraphy: a new instrument. *Behav Res Meth Instr Comp* 28: 392–403, 1996.
27. **Welk GJ, Blair SN, Wood K, Jones S, Thompson RW.** A comparative evaluation of three accelerometry-based physical activity monitors. *Med Sci Sports Exerc* 32: S489–S497, 2000.
28. **Zhang K, Pi-Sunyer FX, Boozer CN.** Improving energy expenditure estimation for physical activity. *Med Sci Sports Exerc* 36: 883–889, 2004.
29. **Zhang K, Werner P, Sun M, Pi-Sunyer FX, Boozer CN.** Measurement of human daily physical activity. *Obes Res* 11: 33–40, 2003.