Towards valid estimates of activity energy expenditure using an accelerometer: searching for a proper analytical strategy and big data

Alberto G. Bonomi
Personal Health Solutions, Philips Research, Eindhoven, The Netherlands
Submitted 10 September 2013; accepted in final form 10 September 2013

Estimating energy expenditure and the metabolic response to physical activity are complex areas of research, manifested by the increasingly sophisticated instruments and methodologies being proposed in the literature. Wearable motion sensors, such as accelerometers, are among the most widely studied tools to assess activity-related energy expenditure (AEE) by quantifying body movement. These two fundamental aspects of physical activity—body movement and the metabolic response—although dependent, deserve specific evaluation methods. Body movement has been conventionally represented by the so-called activity counts, a summary metric of the acceleration signal variability, while estimates of AEE have been obtained by combining activity counts with subject characteristics using linear-regression models. Many consider this approach obsolete or intuitively subject to limitations, leading to inevitable methodological error. Indeed, the variability in activity counts between different activities and workloads does not always reflect a corresponding variation in AEE. Thus interest grew in processing raw accelerometry data to better describe body movement, for example, by limiting influences of noise and gravity (10), gathering advanced statistical features about human motion, and discriminating between categories of activities (1). Concurrently, interpretative algorithms of body movement evolved from simple regression to models developed using machine-learning techniques (8, 9). Among these attempts to estimate AEE accurately, establishing which one yielded the largest improvements has never been an easy mission.

Recently, two original analytical methods have been presented in the literature to achieve accurate AEE estimates from accelerometer data. The first, based on the “divide-and-conquer” strategy, implies a two-stage process, in which body movement is initially analyzed to recognize the type of activity carried out and then estimate energy expenditure using activity-specific prediction algorithms (2). The second analytical approach is “machine-learning driven” and provides AEE estimates by training sophisticated algorithms, such as artificial neural network (9), polynomial or spline interpolation models (3), and support vector machines (8). The divide-and-conquer strategy specifically addresses the problem of a strong activity-type dependency in the relationship between body movement and AEE. On the other hand, the machine-learning-driven method delegates to the learning algorithm the task of representing the nonlinear patterns in the training data and accurately estimating the metabolic cost.

Comparing these analytical strategies for estimating AEE was the aim of a study published in the current issue of the Journal of Applied Physiology. Ruch and colleagues (7) tested under laboratory conditions the agreement between objective measurements of AEE and the outcome of several predictive algorithms based on data captured by a hip-mounted tri-axial accelerometer. Body movement was expressed as activity counts/second, number of steps, and trunk-inclination angle. A large group of children was recruited to follow a protocol, including sedentary, recreational, ambulatory, and sport activities. Results showed that measurements and estimates of AEE were in larger agreement using activity-specific equations rather than a single regression equation, representing the conventional analytical method. The application of activity-specific AEE prediction equations, according to the outcome of a decision-tree activity classifier, showed a lower bias and lower root mean squared error than an artificial neural network. The results confirmed previous data on the low accuracy of an artificial neural network to predict AEE during sedentary occupations (4). Furthermore, the study by Freedson and colleagues (4) showed that already a simple algorithm based on the divide-and-conquer strategy (Creuter 2-regression model) had only slightly worse accuracy than an artificial neural network. This suggests that deploying advanced information on activity types may result in more correct AEE estimates, but free-living validation is required to substantiate this hypothesis. Only the divide-and-conquer strategy has been validated in both laboratory and free-living settings (2, 7). AEE estimation models derived by machine-learning techniques have not yet been compared with an objective reference in daily life.

A valid method to estimate AEE should be unbiased, with small error variability and capable of providing researchers with a tool to evaluate accurately the overall physical activity level, as well as the intensity level for each daily activity. The design of an appropriate analytical strategy to process and interpret accelerometer data represents the way to achieve this goal. Activity classification should become a central element of AEE estimation algorithms, given the documented advantages compared with both traditional linear regression as well as algorithms developed using sophisticated machine-learning techniques (2, 7). Undoubtedly, more effort should be oriented toward the creation of activity-type classification methods with robust accuracy in free-living conditions (5), for example, by reducing the likelihood of ambiguous acceleration features derived from short activities and activity transition. With this target in mind, preprocessing techniques capable of isolating intervals with a stationary acceleration signal may lead to fewer classification errors. Furthermore, classification of both upper and lower limbs’ motion could be achieved by collecting data using distributed wearable sensors, fostered by recent advantages in sensor miniaturization, low-power consumption, and wireless communication. However, the range of activity categories identified should complement the possibility to explain the variability in AEE using summary metrics of body movement for each activity type. In this respect, machine-learning techniques are a powerful tool to combine multiple...
features of body movement and model AEE variability. Yet, the improvement of machine-learning models, such as an artificial neural network, without prior activity classification or with only basic discrimination between activity and inactivity periods (6) could be investigated further by using large training datasets. Sophisticated algorithms to estimate AEE without activity-type information may be developed using large databases and tools provided by big data analytics. Nevertheless, the effort for creating large and shared training datasets is perhaps unjustifiable given the severe lack of standardization in collecting and processing accelerometer data.

In conclusion, algorithms based on the divide-and-conquer strategy have shown better results when validated to estimate AEE under laboratory conditions rather than machine-learning-driven ones. A proper strategy for achieving valid AEE estimates should include activity classification, since it simplifies the development of models aimed at explaining the variability in AEE from body movement data. Notably, the achievement of precise classification among activity categories using wearable and unobtrusive accelerometers still requires further work, especially concerning validation in free-living conditions. A composite solution that implies AEE estimates from a comprehensive machine-learning model when activity-type classification is achieved with poor confidence may still represent a valid intermediate step to improve AEE assessments. Big data may revolutionize the way AEE estimation algorithms are developed, yet establishing infrastructures to collect large datasets and standardizing physical activity measurements are key issues challenging the creation of sustainable and valid solutions to estimate AEE accurately.

DISCLOSURES

The author is employed by Philips Research Laboratories.

AUTHOR CONTRIBUTIONS

Author contributions: A.G.B. drafted manuscript; A.G.B. edited and revised manuscript; A.G.B. approved final version of manuscript.

REFERENCES