Validation of Three Alternative Methods to Measure Total Energy Expenditure against the Doubly Labeled Water Method for Older Japanese Men

Hoby Hasina RAfAMANTANANTSOA1, Naoyuki EBINE2, Mayumi YOSHIKAN3, Hiroyuki HIguchi3, Yutaka YOSHITake4, Hiroaki TANAKA3, Shinichi SAT01 and Peter John Harris JONES2

1 Institute of Health and Sport Sciences, University of Tsukuba, Tennoudai 1-1-1, Tsukuba, Ibaraki 305-8574, Japan 2 School of Dietetics and Human Nutrition, McGill University, Quebec H9X 3V9, Canada 3 Department of Health and Sports Sciences, Fukuoka University, Nanakuma 8-19-1, Jonan-ku, Fukuoka 814-0180, Japan 4 National Institute of Fitness and Sports in Kanoya, Shiromizu 1-1, Kanoya, Kagoshima 891-2393, Japan

(Received June 17, 2002)

Summary In a previous study using young Japanese men as subjects, Ebine et al. found that accelerometer (AC) represents a promising technique for measuring free-living total energy expenditure (TEE) when compared to activity records (AR) and heart rate monitoring (HR). Thus, the present study was designed to validate the use of an AC and to determine whether or not the previous findings regarding the three alternative field methods (AC, AR, and HR) could be extended to older Japanese men (n=24; mean±SD age 48±10 y, body mass index 23.1±2.7 kg/m² and body fat 18.7±4.8%). TEE values obtained over a 3 d period by AR, HR, and AC (3dAC), and AC over a 14 d period (14dAC) were simultaneously validated against TEE measured by the doubly labeled water (DLW) method applied within a 14 d period. TEE values obtained by AR, HR, 3dAC, and 14dAC ranged from 1,750 to 3,447 kcal/d, 1,691 to 5,286 kcal/d, 1,716 to 2,765 kcal/d, and 1,700 to 2,855 kcal/d, respectively. Expenditures obtained by HR were similar to those obtained using the DLW method, with a mean difference of 57±603 kcal/d (2%), but those obtained using AR, 3dAC, and 14dAC differed substantially from the DLW method, with mean differences of -335±289 kcal/d (12%), -542±249 kcal/d (-19%), and -566±223 kcal/d (-20%), respectively. AR, HR, 3dAC, and 14dAC were significantly correlated with the DLW method, with r values of 0.76 (p<0.0001), 0.67 (p<0.001), 0.78 (p<0.0001), and 0.83 (p<0.0001), respectively. Intra-individual variation indicated by the coefficient of variation (CV) was significantly higher for HR (15±11%, p<0.001) than for AR (7±4%), 3dAC (7±5%), and 14dAC (8±3%). The same findings were obtained using Bland and Altman plots at the population level. Interestingly, 3dAC and 14dAC were significantly correlated with r=0.97 (p<0.0001), with a lower mean difference of 24 kcal/d. These results suggest that, same as the previous study, AC is superior to HR in estimating TEE, and seems to be satisfactory for estimation at both group and individual levels, particularly for large-scale studies of older individuals when compared to the DLW method. However, some modifications of the AC method may be needed to compensate for the underestimation of TEE.

Key Words energy expenditure, doubly labeled water method, accelerometer, activity record, heart rate monitoring

Recently, there has been growing interest in measuring human total energy expenditure (TEE) under regular living conditions (1, 2). Such studies have characterized energy expenditure largely in sedentary healthy populations (3, 4) and in patients with chronic diseases (5, 6) of different ages and genders. The relationship between TEE and morbidity and mortality has been particularly well established in adult populations (7, 8).

Various techniques have been used to assess habitual TEE, including the use of motion sensors (9), self-reporting (10), physiological markers (11), and calorimetry (12). Most of these conventional methods have methodological problems of accuracy and/or feasibility for estimating TEE in specific populations or circumstances. The use of activity records (AR) requires complete commitment and conscientiousness of the subjects (13) in order to obtain accurate data. Heart rate monitoring (HR) is affected by emotional factors such as mental stress and environmental conditions such as ambient temperature and humidity (14). Thus, the utility of this technique remains controversial and problematic (15).

Despite the inadequacies of these alternative approaches, they have gained considerable popularity among researchers. In a previous validation study,
Ebine et al. compared AR, HR, and accelerometer (AC) methods against the doubly labeled water (DLW) method as a reference methodology over 14 consecutive days for young Japanese men (24.2±1.8 y), and found that, compared with HR, AR and AC were superior for estimating TEE (2). However, the subjects used were not of the age group most likely to need activity-related lifestyle interventions to prevent chronic diseases and the assessment period in that study was longer than the 3 to 7 d that are usually considered to be better tolerated by subjects for time intervals representative of daily life (16, 17). Yet, regarding the assessment period of each method examined in this study, the DLW method and AC are non-invasive and easier to be applied, whereas AR and HR demand much from subjects and are so tedious that they might interfere with the subjects’ daily activities, especially when they are occupationally involved individuals (14) like our participants. Therefore, our objectives were firstly to validate AC and secondly to determine whether the previous findings about the three alternative field methods of assessing TEE (AR, HR, and AC) against the DLW method could be extended to Japanese men of age 30 to 69 y under free-living conditions.

METHODS

Subjects
Twenty-four healthy Japanese male volunteers, ages 30–69 y old, gave their informed consent and were monitored over the experimental period after comprehension of the nature and purpose of the study. Subjects were recruited mainly from the University of Fukuoka and the surrounding communities. Anthropometric measurements are presented in Table 1. All subjects were screened for thyroid status (triiodothyronine, thyroxine, and thyrotropin) and were found to have values within the normal ranges. This study was approved by the ethical committee of the National Institute of Health and Nutrition in Japan.

Table 1. Physical characteristics and body composition of subjects (n=24).

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Mean±SD (y)</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>48±10</td>
<td>30–69</td>
</tr>
<tr>
<td>Body weight</td>
<td>65.5±10.9</td>
<td>49.3–89.6</td>
</tr>
<tr>
<td>Height</td>
<td>167.8±7.2</td>
<td>155.0–189.0</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>23.1±2.7</td>
<td>19.6–29.5</td>
</tr>
<tr>
<td>Body fat (%)</td>
<td>18.7±4.8</td>
<td>10.6–28.3</td>
</tr>
<tr>
<td>Fat-free mass (kg)</td>
<td>53.0±8.7</td>
<td>37.6–73.2</td>
</tr>
<tr>
<td>Fat mass (kg)</td>
<td>12.3±4.0</td>
<td>6.1–19.3</td>
</tr>
</tbody>
</table>

BMI: body mass index.

Experimental design
The experimental period lasted for 16 d, as shown in Fig. 1. The measurements of TEE with the DLW method were conducted over the last 14 d. AC was used over a 3 d period (3dAC), and recordings of AR, HR, and dietary records were performed during the 3 d. AC was then applied extensively over the same period when the DLW method was applied (14dAC). The two days prior to the 14 d period of TEE measurement were spent for familiarizing the subjects with the different methods applied in this study and collecting physiological samples (urine) as baseline (day −1). On the next day (day 0), measurements of the subjects' physical characteristics, body composition, and basal metabolic rate (BMR) were performed, followed by the administration of a single oral dose of DLW (2H₂18O). Urinary samples were collected 3 h and 4 h after administration for the determination of dilution space. Urinary samples for isotopic elimination analysis of ²H and ¹⁸O were collected on the first (day 1) and last (day 15) days of the 14 d period. On the last day after the last collection of urinary sampling, each participant performed a treadmill exercise to establish the relation between HR and VO₂. Subjects were asked not to change their daily activity patterns and not to alter their habitual dietary intake over the 14 d period of measurements.
Measurements

Anthropometry and body composition. Body mass index (BMI) was calculated by dividing body weight in kilograms by the square of body height in meters (kg/m²). Body composition was measured according to Brozek et al. using an underwater weighing method (18) (Table 1).

Basal metabolic rate. Basal metabolic rate (BMR) was measured in the early morning of day 0 by collecting the subjects’ expired air using a ventilated hood. Prior to this measurement, subjects had fasted for 10 h after their last meal and rested supine for 20 min. Expired air was analyzed by an ARCO-1000 respiratory mass spectrometer (ARCO System Inc., Japan). Energy expenditure during rest was computed from oxygen (VO₂) and carbon dioxide (VCO₂) production according to Weir’s equation (19).

Energy intake. Energy intake was evaluated from self-administered dietary records accompanied with multimedia equipment (camera: Fuji Nexia F100 color-APS film) over 3 d and analyzed by a well-trained dietitian using a computer program (Basic-4 Version 2.0; Kagawa Nutrition University) based on Standard Tables of Food Composition in Japan (20) to provide food quotient values.

Individual calibrations (treadmill exercise). Subjects exercised incrementally on a treadmill to establish the relation between HR and VO₂. Before commencing the treadmill exercise, subjects rested for 20 min, during which time simultaneous measurements were made of HR and VO₂ while subjects were lying supine, sitting, and standing quietly with minimal movement. Next, participants exercised by walking, jogging, and running according to a step-wise increase in intensity gradient, where every 4 min the treadmill speed was increased 20 m/min. The slope of the treadmill was kept at 0% during the exercise test. Each 4 min measurement period consisted of a 2 min period of equilibration followed by 2 min for measurement of HR and VO₂. During the treadmill exercise, VO₂ was measured using a nose clip and mouthpiece connected to a respiratory valve that sensed the expired air, which was simultaneously analyzed by an ARCO-1000 respiratory mass spectrometer.

Total energy expenditure

Doubly labeled water method: Details of the method used have been described by Ebine et al. elsewhere (21). TEE from the DLW method was determined from calculation of the carbon dioxide production rate, known as rCO₂ (mol/d), and the food quotient, which was derived from dietary records (22) using Weir’s formula (19). On day 0, after collection of a urinary sample for baseline, each subject ingested a dose of approximately 0.12 g/kg estimated total body water (eTBW) of ²H₂O (99.8 atom %) and 0.25 g/kg eTBW of H₂¹⁸O (10.0 atom %), where eTBW was estimated as being 60% of each individual’s body weight (23). After dose administration, subjects refrained from eating and drinking for a 4 h equilibration period (3 h and 4 h urine sampling). Then, the second void urines in the mornings of day 1 and day 15 were collected for the isotope elimination analysis. ²H enrichment was analyzed using a 903D dual-inlet isotope ratio mass spectrometer (IRMS, VG Isogas, Cheshire, UK) calibrated using Vienna Standard Mean Ocean Water, 302B, and the Greenland Ice Sheet Precipitation Standard; ¹⁸O enrichment was determined using SIRA 12 IRMS (VG Isogas). Each measurement of all samples and the corresponding reference was performed in triplicate. The average standard deviation through the analysis was 1.22‰ for ²H and 0.20‰ for ¹⁸O. TEE was expressed as the mean TEE over the 14 d period of assessment.

Activity records: The AR consists of logging subject daily activity over a 3 d period within the 14 d interval of DLW method assessment. Participants used a specially designed form to record each activity minute-by-minute to facilitate maintaining the diary. They were instructed to record only when there was change in activity, by drawing a line at the end of one activity under the corresponding indicator of time. They were also instructed to state the nature of each activity and fractionate it in terms of the percentage of time spent in lying, sitting, standing, and/or walking position and whether multiple tasks were done at once. Detailed demonstration and a completed sample as a model were given before the recording. Soon after the achievement of 3 d of activity logging, the recording sheet was collected and checked to decrease under-recording. Then, data from the diary were converted into energy expenditures by estimating and multiplying the time period in which the subject was engaged in the activity by the energy cost of each activity in terms of METs (1 MET=1 kcal/kg/h) based on the compendium of physical activities (24). Twenty-four-hour energy expenditure was calculated by summing the total of energy expended by each activity during a day. Hence, TEE from AR was expressed as the mean TEE over a 3 d period.

Heart rate monitoring: A cardio-frequency meter “Accurex plus” (Polar electro, Kempele, Finland), was used to record the heart rate of each subject minute-by-minute over a 3 d period within the 14 d interval of DLW method assessment to evaluate the mean TEE from HR. The device consists of an electrode-belt transmitter and wrist microcomputer receiver that store the pulse, recorded at intervals of every 1 min, in memory. Recorded data were retrieved via an interface unit and computerized to calculate the TEE from an established calibration line. The relationship between HR and VO₂ was determined by assigning the derivation of calibration line and identification of the FLEX-HR point and based on the treadmill exercise. FLEX-HR is empirically defined as the mean of the lowest heart rate during the lightest imposed exercise and the highest heart rate during rest (12). The procedures applied in this study were adopted from those fully described by Ceesay et al. elsewhere (25). Using this instrument, TEE was computed by summing estimated energy expenditure established by HR monitoring and combining the energy cost of sleep. The latter was assumed to be equal to the measured BMR (26).

Accelerometer: The LifeCorder (Suzuken Co., Japan)
is a newly developed uni-dimensional AC in a lightweight microcomputer (6.2 × 4.6 × 2.6 cm; weight = 42 g) designed to detect steps and acceleration rates along the vertical axis at the waist level with an operating part to automatically estimate energy expenditure. According to the instructions of the manufacturer, the device was programmed initially to estimate TEE by registering, summing, and then computerizing the data information sensed from body movement using the following equation:

\[ \text{TEEAC} = \text{BMR} + \text{PA} + \left( \frac{1}{10} \right) \text{TEE} \]

where the TEE value is obtained by calculating the average of TEE over a period of 2 wk during the same interval of DLW method assessment. Before use, the AC was programmed with the age, gender, height, and body weight of each subject for estimating adjusted BMR. Participants were instructed to firmly clip the device to a belt at the right side of the waist during waking time except during bathing.

PA represents the energy of physical activity in kcal. According to the number of body movements sensed within intervals of every 4 s, a profile of that PA was presented in a graph, which was generated after downloading recorded data information into a computer.

\[ \text{BMR (kcal)} = K_b \times \text{BSA} \times \frac{T}{10,000} \]

where \( K_b \) is a standard value for Japanese men (kcal/m²/h) (27); BSA is the body surface area (cm²); and \( T \) is time (h).

\[ \text{BSA (cm²)} = \text{body weight}^{0.444} (\text{kg}) \times \text{height}^{0.663}(\text{cm})^{-88.83} \]

\( (1/10)\text{TEE} \) in kcal refers to the energy spent for diet-induced thermogenesis.

Statistical analysis. Values in the text and tables are presented as means ± SD. Differences in TEE among the three alternative methods and the DLW method were examined using one-way ANOVA followed by post-hoc test (Tukey-Kramer test), whereas differences between the AC and DLW methods applied over a 14 d period and between 3dAC and 14dAC were tested for significance by Student’s paired t-test. Pearson correlation coefficients were used to determine the linear relationship between AR, HR, and AC and the DLW method where values were specified to be significant at \( p<0.05 \). To evaluate the relative within-subject variability of each method, which is expressed as a percentage, we calculated the intra-individual CV by dividing the individual standard deviations by the individual mean TEE obtained over the respective assessment period. Agreement between each alternative method and DLW was plotted using the approach of Bland and Altman (28). All statistical analyses were performed using the statistical package StatView 5.01 (SAS Institute Inc., Cary, NC, USA, 2000–2001).

**RESULTS**

Following the criteria of the International Association for the Study of Obesity (IASO), subjects were classified as non-obese (29) (Table 1), and weight stable individuals with no significant difference between initial and final body weight at 65.3 ± 10.8 kg and 65.6 ± 11.0 kg, respectively.

All 24 subjects were able to complete all four of the assessments. Mean TEE values obtained from the AR, HR, 3dAC, 14dAC, and DLW methods are given in Table 2. TEE values obtained from HR did not significantly differ from those obtained from DLW, with a mean difference of 57±603 kcal/d, but AR, 3dAC, and 14dAC underestimated TEE by a mean of 335±289 kcal/d (\( p<0.001 \)), 542±249 kcal/d (\( p<0.001 \)), and 566±223 kcal/d (\( p<0.001 \)), respectively. AR (\( r=0.76, p<0.0001 \)) and 3dAC (\( r=0.78, p<0.0001 \)) were more highly correlated with the DLW method than HR (\( r=0.67, p<0.001 \)). The 14dAC measurements confirmed the strong relationship between the AC and DLW methods, with a significant increase of up to 0.83 (\( p<0.0001 \)) in the correlation coefficient. Examined for intra-individual variability, AR and 3dAC showed similar small deviations with CVs of 7±4% (\( p<0.001 \)) and 7±5% (\( p<0.001 \)), respectively, whereas HR showed

<table>
<thead>
<tr>
<th>Method</th>
<th>Mean±SD (kcal/d)</th>
<th>Range (kcal/d)</th>
<th>Correlation coefficient</th>
<th>Intra-individual CV (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AR</td>
<td>2.389±430**</td>
<td>1.750–3.447</td>
<td>0.76**</td>
<td>7±4</td>
</tr>
<tr>
<td>HR</td>
<td>2.781±791</td>
<td>1.691–5.286</td>
<td>0.67*</td>
<td>15±11**</td>
</tr>
<tr>
<td>3dAC</td>
<td>2.182±249**</td>
<td>1.716–2.765</td>
<td>0.78**</td>
<td>7±5</td>
</tr>
<tr>
<td>14dAC</td>
<td>2.158±259**</td>
<td>1.700–2.855</td>
<td>0.83**††</td>
<td>8±3</td>
</tr>
<tr>
<td>DLW</td>
<td>2.724±396</td>
<td>2.045–3.769</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**\( p<0.0001 \), * \( p<0.001 \) significance compared to DLW method.

**\( p<0.0001 \) significance among AR, HR, and 3dAC.

†† \( p<0.0001 \) significance between 3dAC and 14dAC.

Intra-individual CV represents the average within-subject deviations using each alternative method, which is calculated by dividing individual standard deviations by the individual mean of TEE obtained over a respective assessment period.
large deviations with a CV of $15\pm 11\%$ ($p<0.001$). Interestingly, although applied over a longer period of TEE assessment, the CV of 14dAC ($8\pm 3\%$) did not differ from the CV of 3dAC. The same findings were obtained when the mean difference of each alternative and DLW methods was plotted against the mean of the DLW method considered as a reference method using Bland and Altman plots (28) for method agreement at the population level (Fig. 2).

**DISCUSSION**

In a previous study, Ebine et al. reported that although AC underestimated TEE, it was significantly correlated and had good agreement with the DLW method, thereby suggesting that AC is more accurate in estimating TEE than HR at the group level. These findings were confirmed in the present study. In addition, intra-individual validation of AC is smaller than that of HR. Thus, AC is superior to HR even at the individual level.

Various motion-sensing instruments (ACs) for measuring energy expenditure in validation studies have been introduced, such as the Caltrac accelerometer (30), triaxial accelerometer (31), and calorie counter method (32), but to the best of our knowledge this is the first study to evaluate the newly developed uni-dimen-

sional AC LifeCorder against the DLW method in free-living older individuals. More precisely, first, we examined the accuracy of AC over a longer period (14 d). Then, we also evaluated it against the DLW method over a shorter period (3 d) of application parallel with two other field measures.

Studies in literature have revealed agreement and/or differences between ACs and the DLW method, but Ekelund et al. suggested that discrepancies between studies may be due to the different activity monitors used since the devices may differ in their sensitivity in measuring vertical acceleration during free-living conditions (4). Regardless of which formula the manufacturer is using to convert body acceleration into TEE, our results demonstrated that, in spite of the underestimation of 3dAC of 19% with a 95% confidence interval (CI) of $-437$ to $-647$ kcal/d compared to the DLW method, there was a high correlation with it ($r=0.78$, $p<0.0001$). Yet, the limits of agreements from $-1.039$ to $-44$ kcal/d are small enough to demonstrate concordance of the two methods (Fig. 2) for use at the group level. Furthermore, the small percentage of intra-individual deviation (Table 2) indicates that 3dAC may also be useful at the individual level. For validation, 14dAC confirmed most of the findings with 3dAC, such as the
underestimation of TEE by 20% with 95% CI of −472 to −660 kcal/d where the limits of agreement (28) were narrower, from −1,012 to −120 kcal/d. Intra-individual deviations were also small (Table 2). Interestingly, when AC was monitored over a longer period, a significant relationship with the DLW method was confirmed by an increase in r value, up to 0.83 (p<0.0001). Additionally, 3dAC and 14dAC correlated significantly with r=0.97 (p<0.0001), with a lower mean difference of 24 kcal/d. Nevertheless, our data indicate a need to improve the instrument to reduce the degree of underestimation of TEE by AC. The underestimation by AC may be partly explained by its non-application during sleeping time and during any activity in contact with water. However, other factors may also be involved.

In relation to these factors, a previous report has emphasized that part of the underestimation of TEE from body movement might be related to the fact that during some activities involving static exercise, accelerometer output is not proportional to the increase in energy spent during their assessment (33). Moreover, other studies have stated that accelerometers do not detect the increase in energy expenditure associated with an increase in the grade of walking such as climbing stairs (34) or carrying loads (31). The body does not accelerate in proportion to energy expenditure during such activities, and this factor may contribute to the underestimation of energy expenditure with this tool. In another study, Caltrac appeared to be related to physical activity, so that the more the person moves, the higher the activities count (30). In contrast to these limitations, motion sensors appear to be especially useful in studies of children because they provide continuous, objective assessments of energy expenditure without observer or self-report bias (30). ACs have been validated specifically for use by non-athletes in free-living activities (32) rather than athletes. Indeed our results suggest that an AC would be a meaningful predictor of TEE and its components in real-life situations if some modifications are made to its algorithm.

Concerning the validity of AR against the DLW method, we found that the results in the present study seem to show a different tendency from the earlier findings of Ebine et al. (2). The previous study reported neither significant difference nor correlation between the AR and DLW methods. However, the present study demonstrates that TEE using AR is significantly different from that obtained by the DLW method, by −12% with 95% CI of −457 to −213 kcal/d at the group level, but had a significant correlation at r=0.97, p<0.0001. Despite this, the limits of agreement from −913 to 243 kcal/d are small enough to show the concordance of the two methods (Fig. 2) analyzed at the group level. Furthermore, the small percentage of intra-individual deviations (Table 2) indicates that AR may also be useful for assessments at the individual level. However, as we mentioned earlier (14), AR applied over a long period seems to be limited for use by occupationally involved individuals.

A previous validation study of HR applied to young Japanese men (2) demonstrated no discrepancy between HR (3%) and DLW methods for TEE measurement in free-living conditions. This is similar to the present study (2%) with 95% CI of −198 to 312 kcal/d. However, when plotted by Bland and Altman approaches (28) for assessing population data, the limits of agreement were very wide, from −1,150 to 1,264 kcal/d (Fig. 2). This was supported by the intra-individual deviations of 15±11%, which are about two fold larger than those of AR and AC.

In conclusion, the same as the previous study, AC applied over a 14 d period yielded a stronger relationship with the DLW method, but also had a significant correlation to AC applied over a 3 d period. AC applied over a 3 d period, and also AR, seemed to perform somewhat better in showing a smaller extent of inter-individual deviation and smaller width of limits of agreement than HR. Although HR did not underestimate TEE compared to the DLW method, its value is limited to TEE measurement at the group level. Thus, the introduction of AC for epidemiological studies conducted on older individuals may be appropriate.

Acknowledgments

We gratefully acknowledge Makoto Ayabe for his valuable assistance with the data processing of information from the LifeCorder. Special recognition must also go to all the participants for enabling this project to be conducted. Finally, we wish to express special appreciation for the support of the Japanese Ministry of Health, Labour, and Welfare and Medical Frontier, Strategy Research (H11-KENKOU-018) for funding this study.

REFERENCES


Comparison of Various Field Methods to Measure Total Energy Expenditure

523


