Improvement of New Gyro-Sensor and Accelerometer Device Used for Walking and Running

Shinji Takahashi¹, Koya Suzuki², Yudai Yoshida³,
Yosuke Sakairi³ and Tomohiro Kizuka³

¹Department of Regional Management, Faculty of Liberal Arts, Tohoku Gakuin University, Miyagi 981-3193, JAPAN
shinji@izcc.tohoku-gakuin.ac.jp

²Department of Human Science, Faculty of Liberal Arts, Tohoku Gakuin University, Miyagi 981-3193, JAPAN

³Graduate School of Comprehensive Human Science, University of Tsukuba, Ibaraki 305-8571, JAPAN

[Received April 2, 2010; Accepted November 10, 2010; Published online November 23, 2010]

The purposes of this study were to recalibrate a new motion sensor wristband (ViM sports memory: ViM) to metabolic equivalents (METs) during walking and running, and to examine the cross-validity of the recalibration model. Forty-five healthy Japanese (23 males and 22 females) voluntarily participated in this study. The participants were divided into two groups: a recalibration group (16 males and 14 females) and a cross-validity group (7 males and 8 females). The participants performed 5 min of treadmill walking at 3.6, 4.8, and 6.0 km/h, and treadmill running at 7.2, and 9.6 km/h. The ViM was placed on the non-dominant wrist. Simultaneous measurements of the ViM and an indirect calorimeter (IC) were continuously recorded during all exercises. The recalibration models were calculated for males and females from data of the recalibration group using general linear mixed model. The cross-validity was examined by correlation coefficients, repeated ANOVA, and paired t-test adjusted by Bonferroni's method. In the recalibration group, the recalibration models showed high determination coefficients (male: $R^2=0.74$, $SEE=1.51$ METs; females: $R^2=0.84$, $SEE=1.16$ METs). In the cross-validation group, there were significant correlation coefficients at 3.6, 4.8, 6.0, and 9.6 km/h ($r \geq 0.56$, $p < 0.001$). Differences (-4 to 13%) between the IC measurements and the recalibration model estimations were not significant ($p > 0.185$) at any speed. These results showed that the recalibration model improved the validity and accuracy of the ViM.

Keywords: physical activity, motion monitor, validation study, METs, Mixed model

1. Introduction

There is substantial evidence supporting the importance of habitual physical activity in maintaining good health and avoiding lifestyle-related disease (Fujita et al., 2004; Manson et al., 1992; Paffenbarger et al., 1983; Pate et al., 1995). To further develop the evidence of the association between quantity of physical activity and health, accurate and useful methods of physical activity assessment are needed. Although the accuracy of physical activity assessment by accelerometer is lower than that of doubly labeled water, human indirect calorimeter, and heart rate monitoring, accelerometer methods, which do not restrain the daily life of individuals, can be more useful than

other methods. Therefore, researchers can profit from the use of accurate accelerometers in epidemiological physical activity investigations. Recently, many types of accelerometer have been developed, and the validity and accuracy of accelerometers have been reported (Brage et al., 2003; Campbell et al., 2002; Eston et al., 1998; Freedson et al., 1998; King et al., 2004; Kumahara et al., 2004; Rowlands et al., 2004).

In a previous study (Takahashi et al., 2009), we examined the validity and accuracy of a new type of motion monitor, the ViM sports memory wristband (Microstone Co., Saku, Japan). This device includes a uniaxial accelerometer and a gyro-sensor. The ViM classifies physical activity into 10 patterns, based on angular velocity and vertical acceleration

of arm movements and periodicity of these movements, as follows: very slow walking (W1), slow walking (W2), normal walking (W3), brisk walking (W4), jogging (J1), running (J2), light irregular activity (IR1), moderate irregular activity (IR2), heavy irregular activity (IR3), and rest (Rest). This device more accurately estimates energy expenditure (EE, $\text{kcal} \cdot \text{min}^{-1}$) than a conventional uniaxial accelerometer (Lifecorder; Suzuken Co. Ltd., Nagoya, Japan) placed at waist position during running and non-regulated activity (static stretching and hopscotch).

However, the ViM has a serious problem in regards to estimation of physical activity during walking activity. Although outputs of the ViM indicate approximately correct grade changes in walking speed, the device systematically overestimates EE by around 24-74% during walking, compared to measurements made by an indirect calorimeter (Takahashi et al., 2009). The cause of the problem lies in the estimation equation model of EE used by the device. The manufacture's software estimates EE based on the ViM outputs, participant's age, body weight, and gender. The accuracy of the ViM can be improved by correct recalibration of the equation. Thus, the purposes of this study were to recalibrate the estimation equation of the ViM, and to examine the validity and accuracy of the recalibrated equation compared with the manufacture's equation.

2. Materials and Methods

2.1 Participants

Forty-five healthy Japanese (23 males and 22 females) voluntarily participated in this study. The participants were asked to refrain from alcohol use

and strenuous physical activity for 24 h before exercise tests, and from food and caffeine during the 2 h preceding the experiments. Written informed consent was obtained from all participants before exercise tests. The study protocol was approved by the Human Subjects Committee of Tohoku Gakuin University.

In order to examine the cross validity of the recalibrated equation, we randomly divided the participants into two groups: a recalibration group (16 males and 14 females) and a cross-validity group (7 males and 8 females). **Table 1** shows characteristics of the participants. Differences of age, height, weight, and BMI between the two groups were not significant.

2.2 Exercise test protocol

Each participant came to the exercise laboratory at 1.5 h before exercise test start. After height and weight measurements, participants were familiarized with a motor-driven treadmill (O2road, Takei Sci. Instruments Co., Niigata, Japan) for 3 min at $3.6 \text{ km} \cdot \text{h}^{-1}$. Participants then rested for over 30 min on a comfortable chair until their oxygen uptake ($\dot{V}\text{O}_2$) returned to baseline (less than $5.0 \text{ ml} \cdot \text{kg}^{-1} \cdot \text{min}^{-1}$).

Participants then completed an exercise test on the treadmill (horizontal slope). The exercise test consisted of walking at 3.6, 4.8, and $6.0 \text{ km} \cdot \text{h}^{-1}$, and running at 7.2 and $9.6 \text{ km} \cdot \text{h}^{-1}$. Participants walked or ran at each speed for 5 min. Intervals of 3 min rest were provided between each speed.

A portable indirect calorimeter (IC) (MetaMax-3B, Cortex, Leipzig, Germany) was connected to participants via a facemask, and $\dot{V}\text{O}_2$ was continuously measured throughout the familiarization and exercise tests. The ViM was placed on the non-dominant side wrist of each

Table 1 Characteristics (mean $\pm$ SD) of the participants

Items (unit)	Gender	Recalibration group (Male, N = 16; Female, N = 14)	Cross-validation group (Male, N = 7; Female, N = 8)
Age (yrs)	Male	23.5 $\pm$ 2.8	22.0 $\pm$ 1.6
	Female	20.5 $\pm$ 1.4	22.0 $\pm$ 3.7
Height (cm)	Male	171.4 $\pm$ 4.8	173.8 $\pm$ 6.6
	Female	161.2 $\pm$ 6.9	158.3 $\pm$ 5.6
Weight (kg)	Male	66.5 $\pm$ 9.1	66.5 $\pm$ 8.7
	Female	51.6 $\pm$ 5.2	52.2 $\pm$ 8.4
BMI (kg/m^2)	Male	22.5 $\pm$ 2.1	22.0 $\pm$ 2.4
	Female	19.9 $\pm$ 1.9	20.7 $\pm$ 2.2

participant, and the IC was held to the chest by a harness.

2.3 Equipment

ViM. The ViM assembly includes an accelerometer, a gyro-sensor, a microprocessor, and communication ports for output to a computer. The device measures 6.5 cm in width, 7.5 cm in height, and 1.5 cm in thickness, and weighs 100 g. **Table 2** shows the algorithm of the ViM provided by the manufacturer to recognize physical activities and classify them into the 10 motion patterns. According to technical specifications provided by the manufacturer, this device samples vertical acceleration and angular velocity of the arm every 50 ms. Dynamic ranges of acceleration and angular velocity are 0.05-2.00 G and 0-300 deg · s⁻¹, respectively. The frequency response is 0.25-20 Hz, and the time period is 2 s. The ViM records the sum of frequencies of each activity pattern every 30 or 180 s by optional setting. A 32 Kb EEPROM is used as an internal memory to store data for 1 day or 7 days, depending on whether 30 or 180 s statistical mode is selected. In this study, 30 s mode was selected. Four ViMs were used in the exercise tests. The reliability and objectivity of the ViMs were reported in the previous study (Takahashi et al., 2009).

The ViM outputs (frequencies · 30 s⁻¹) were downloadable to a computer in which manufacturer's software was installed. The manufacturer's software could calculate EE (kcal · min⁻¹) according to an equation comprising sex,

age, body weight, and the ViM output data. However, the manufacturer did not disclose the detailed equation.

Indirect calorimeter. The MetaMax-3B measurements, described in METs and kilocalories per minute (kcal · min⁻¹), were used as the validity criteria. The IC measures breath-by-breath ventilation ($\dot{V}E$), concentration of expired oxygen (FEO₂), and carbon dioxide (FECO₂). We calibrated the IC before each test. The turbine flow meter (range: 0.05-2.00 L · s⁻¹) was calibrated with a 3.0-L calibration syringe, and the O₂ and CO₂ analyzers were calibrated with room air and a calibration gas of known O₂ (15.94%) and CO₂ (3.97%) composition. The breath-by-breath data were averaged every 30 s. $\dot{V}O_2$ measured by the IC were converted to METs and EE (kcal · min⁻¹) by the following equations (Dale et al, 2002):

$$\text{METs} = \dot{V}O_2 (\text{ml} \cdot \text{kg}^{-1} \cdot \text{min}^{-1}) \div 3.5 \quad (\text{Eq1})$$

$$\begin{aligned} \text{EE (kcal} \cdot \text{min}^{-1}) \\ = 0.0175 \cdot \text{METs} \cdot \text{body weight (kg)} \end{aligned} \quad (\text{Eq2})$$

2.4 Data reduction

Internal clocks of the IC and ViM were initialized and set to standard time before each test. All measurements were converted to synchronize unit time as 1 min: the ViM outputs, recorded every 30 s, were summed for 1 min, and the IC data averaged every 30 s were selected last 30 s data. To ensure steady state of METs, EE, and the ViM outputs, averaged data for the last 2 min (4-5 min) of each speed were calculated.

In results of the previous study (Takahashi et al., 2009), there were obvious technical biases in the ViM outputs during the treadmill exercise: 1) there were few records of IR1, IR2, and IR3, because the arm swing during walking and running was a periodic rotation movement; and 2) the difference between W2 and W3 became unstable among each individual, because the differences in the arm swing among individuals were large at very low to moderate speeds. These biases probably resulted in problems with the calculation of regression coefficients of the recalibration equation. Thus, we unified W1 and IR1, J1 and IR2, J2 and IR3, W2 and W3 from the information of the acceleration (**Table 2**), respectively.

Table 2 The algorithm of the ViM to classify body motion patterns

Variable (motion pattern)	Acceleration	Angular velocity	Periodicity
W1 (very slow walking)	Low	Very low	Present
W2 (slow walking)	Low	Low	Present
W3 (normal walking)	Low	Moderate	Present
W4 (brisk walking)	Low	High	Present
J1 (jogging)	Moderate	Moderate	Present
J2 (running)	High	Moderate	Present
IR1 (standing working)	Low	Irregular	None
IR2 (light sports)	Moderate	Irregular	None
IR3 (heavy sports)	High	Irregular	None
Rest (rest)	Very low	Irregular	None

This material is provided by the manufacturer (Micrstone Co., Saku, Japan).

2.5 Statistical analysis

Statistical analyses were performed using SPSS 16.0J (SPSS Japan, Tokyo, Japan). General linear mixed model (Mixed model) was used to recalibrate the relationship between the ViM outputs and measured METs over all speeds. In the most calibration studies (Brage et al., 2003; Freedson et al., 1998; Hendelman et al., 2000; Kumahara et al., 2004; Leenders et al., 2003; Nichols et al., 1999), a multiple regression analysis was used to pooled dataset, which included repeated data points for each individual. However, statistically, the approach is not allowed because it violates the independence assumption of multiple regression analysis (Grafen and Hails, 2002; Twisk, 2003). On the contrary, Mixed model, which calculates the repeated nature as a random effect (variance component), can be used on the pooled dataset. The multiple regression model used in the present study can be written as follows:

$$y_{ij} = \alpha + \beta_1 x_{1ij} + \beta_2 x_{2ij} + \beta_3 x_{3ij} + \beta_4 x_{4ij} + \beta_5 x_{5ij} + a_i + e_{ij} \quad (\text{Eq3})$$

where y_{ij} represents response variable (METs) for the j th observation on the i th participant, x_1 (W1 + IR1), x_2 (W2 + W3), x_3 (W4), x_4 (J1 + IR2), and x_5 (J2 + IR3) are explanatory variables, α and β_1 to β_5 are fixed effects that denote the intercept and slopes, a_i represents the variance of the random intercept for each participant, and e_{ij} is the variance of random error. Parameters were estimated by restricted maximal likelihood. The model was fitted to males and females in the recalibration group.

The cross-validity of the recalibration model was examined from the data of the cross-validation group. Estimations by the recalibration model were compared with estimations by the manufacture's software and the IC measurements. Comparisons among the 3 methods were not simple, because the manufacture's software only calculates EE ($\text{kcal} \cdot \text{min}^{-1}$); therefore, METs by the recalibration model and the IC were converted into EE by Eq2.

Estimated EE made by the recalibration model (EE_{REC}), the manufacture's software estimations (EE_{MAN}), and measurements by the IC (EE_{IC}) were compared using two-way repeated ANOVA with within-subjects factors of the methods (EE_{REC} , EE_{MAN} , and EE_{IC}) and the speeds (3.6, 4.8, 6.0, 7.2, and $9.6 \text{ km} \cdot \text{h}^{-1}$). One-way repeated ANOVAs with

a within-subject factor of the methods were performed when the interaction and/or the main effect of the methods factor were significant in results of two-way repeated ANOVA. If main effects were significant in one-way repeated ANOVAs, we examined differences between EE_{IC} and EE_{MAN} , and EE_{IC} and EE_{REC} by paired t -tests. Significance level by two-way repeated ANOVA was $p = 0.050$, and alpha levels of one-way repeated ANOVAs and paired t -tests were adjusted by the Bonferroni inequality. Alpha levels of one-way repeated ANOVA were $p < 0.008$ ($0.05/6$), and those of paired t -tests were $p < 0.004$ ($0.05/12$).

We also calculated correlation coefficients of two estimations of measurements and mean squared error (MSE) at each speed. MSE for both EE_{REC} (MSE_{REC}) and EE_{MAN} (MSE_{MAN}) was calculated using the following equations:

$$\text{MSE}_{\text{REC}} = \frac{1}{N} \cdot \sum (\text{EE}_{\text{REC}i} - \text{EE}_{\text{IC}i})^2 \quad (\text{Eq4})$$

$$\text{MSE}_{\text{MAN}} = \frac{1}{N} \cdot \sum (\text{EE}_{\text{MAN}i} - \text{EE}_{\text{IC}i})^2 \quad (\text{Eq5})$$

3. Results

Recalibration model. Two failures occurred during exercise tests: one case was in a male at $7.2 \text{ km} \cdot \text{h}^{-1}$, and other case was in a female at $9.6 \text{ km} \cdot \text{h}^{-1}$. The reasons for failure were that a sampling tube of the MetaMax broke away from facemask while the male participant was running, while the other case was of a female who could not run due to exhaustion. Both the male and female were in the recalibration group.

Means and standard deviations (SD) of the ViM outputs except Rest for the recalibration group are shown in **Table 3**. The ViM recorded seated rest as Rest completely ($30 \text{ frequencies} \cdot \text{min}^{-1}$, 100%) when participants rested on the comfortable chair; there were almost no counts of Rest during walking and running. **Table 4** represents the results of Mixed model. The recalibration models described in **Table 4** are also presented separately for males and females as follows:

Males model:

$$\begin{aligned} \text{METs} = & 1.22 + 0.07 \cdot (\text{W1} + \text{IR1}) \\ & + 0.07 \cdot (\text{W2} + \text{W3}) + 0.13 \cdot \text{W4} \\ & + 0.22 \cdot (\text{J1} + \text{IR2}) + 0.32 \cdot (\text{J2} + \text{IR3}) \quad (\text{Eq6}) \end{aligned}$$

Table 3 Means $\pm$ SD for METs and the ViM in the recalibration group; results are shown for each speed separately

	Variables	0 km/h	3.6 km/h	4.8 km/h	6.0 km/h	7.2 km/h	9.6 km/h
Males (N = 16)	METs	1.2 $\pm$ 0.3	3.0 $\pm$ 0.6	3.7 $\pm$ 0.8	4.9 $\pm$ 1.1	7.1 $\pm$ 1.3	9.3 $\pm$ 1.8
	W1	0 $\pm$ 0	10 $\pm$ 9	2 $\pm$ 5	0	3 $\pm$ 5	1 $\pm$ 3
	W2	0 $\pm$ 0	17 $\pm$ 8	7 $\pm$ 7	1 $\pm$ 4	0 $\pm$ 1	1 $\pm$ 2
	W3	0 $\pm$ 0	2 $\pm$ 3	12 $\pm$ 6	2 $\pm$ 3	0 $\pm$ 0	0 $\pm$ 0
	W4	0 $\pm$ 0	0 $\pm$ 1	9 $\pm$ 8	25 $\pm$ 7	6 $\pm$ 10	3 $\pm$ 8
	J1	0 $\pm$ 0	0 $\pm$ 1	1 $\pm$ 1	2 $\pm$ 3	13 $\pm$ 9	11 $\pm$ 9
	J2	0 $\pm$ 0	0 $\pm$ 0	0 $\pm$ 0	0 $\pm$ 0	7 $\pm$ 6	13 $\pm$ 9
	IR1	0 $\pm$ 0	0 $\pm$ 0	0 $\pm$ 0	0 $\pm$ 1	0 $\pm$ 1	0 $\pm$ 0
	IR2	0 $\pm$ 0	0 $\pm$ 0	0 $\pm$ 0	0 $\pm$ 0	0 $\pm$ 0	0 $\pm$ 0
	IR3	0 $\pm$ 0	0 $\pm$ 0	0 $\pm$ 0	0 $\pm$ 0	0 $\pm$ 0	0 $\pm$ 0
Females (N = 14)	METs	1.3 $\pm$ 0.4	3.4 $\pm$ 0.8	4.1 $\pm$ 0.9	5.7 $\pm$ 1.2	7.7 $\pm$ 1.4	9.4 $\pm$ 1.3
	W1	0 $\pm$ 0	6 $\pm$ 9	0 $\pm$ 1	0 $\pm$ 0	3 $\pm$ 4	2 $\pm$ 4
	W2	0 $\pm$ 0	13 $\pm$ 9	6 $\pm$ 9	0 $\pm$ 0	0 $\pm$ 0	0 $\pm$ 1
	W3	0 $\pm$ 0	6 $\pm$ 6	7 $\pm$ 7	1 $\pm$ 2	0 $\pm$ 0	0 $\pm$ 0
	W4	0 $\pm$ 0	5 $\pm$ 9	14 $\pm$ 12	23 $\pm$ 7	1 $\pm$ 3	0 $\pm$ 0
	J1	0 $\pm$ 0	0 $\pm$ 1	2 $\pm$ 7	6 $\pm$ 7	14 $\pm$ 7	8 $\pm$ 8
	J2	0 $\pm$ 0	0 $\pm$ 0	0 $\pm$ 0	0 $\pm$ 0	12 $\pm$ 7	19 $\pm$ 8
	IR1	0 $\pm$ 0	0 $\pm$ 0	0 $\pm$ 0	0 $\pm$ 0	1 $\pm$ 1	0 $\pm$ 0
	IR2	0 $\pm$ 0	0 $\pm$ 0	0 $\pm$ 0	0 $\pm$ 0	0 $\pm$ 0	0 $\pm$ 0
	IR3	0 $\pm$ 0	0 $\pm$ 0	0 $\pm$ 0	0 $\pm$ 0	0 $\pm$ 0	0 $\pm$ 0

Units of W1, W2, W3, W4, J1, J2, IR1, IR2, and IR3 are frequencies $\cdot$ min⁻¹. Each ViM variable is a relative frequency scale (30 frequencies $\cdot$ min⁻¹ is 100%).

Table 4 Estimated coefficients of recalibration models

Term (unit)	Males			Females		
	β	SE	p value	β	SE	p value
Intercept	1.22	0.38	0.003	1.30	0.32	< 0.001
x_1 (frequencies $\cdot$ min ⁻¹)	0.07	0.02	0.001	0.08	0.02	0.001
x_2 (frequencies $\cdot$ min ⁻¹)	0.07	0.01	< 0.001	0.08	0.01	< 0.001
x_3 (frequencies $\cdot$ min ⁻¹)	0.13	0.01	< 0.001	0.13	0.01	< 0.001
x_4 (frequencies $\cdot$ min ⁻¹)	0.22	0.02	< 0.001	0.19	0.02	< 0.001
x_5 (frequencies $\cdot$ min ⁻¹)	0.32	0.02	< 0.001	0.31	0.01	< 0.001
R^2		0.74			0.84	
SEE (METs)		1.51			1.16	

x_1 : W1 + IR1, x_2 : W2 + W3, x_3 : W4, x_4 : J1 + IR2, and x_5 : J2 + IR3.

β : regression coefficient.

SE: standard error of the coefficient.

SEE: standard error of estimate.

Females model:

$$\begin{aligned} \text{METs} = & 1.30 + 0.08 \cdot (W1 + IR1) + \\ & 0.08 \cdot (W2 + W3) + 0.13 \cdot W4 \\ & + 0.19 \cdot (J1 + IR2) + 0.31 \cdot (J2 + IR3) \quad (\text{Eq7}) \end{aligned}$$

Goodness of fit for these models, corresponding to a determination coefficient (R^2) and standard error of estimate (SEE), were $R^2 = 0.74$ (SEE = 1.51 METs) for males and $R^2 = 0.84$ (SEE = 1.16 METs) for females, respectively.

Cross-validation. Figure 1 shows the relationships between the measurements made by the IC and estimations by two estimation methods for the cross-validation group data. Determination coefficients of EE_{MAN} and EE_{REC} to EE_{IC} were $R^2 = 0.59$ (SEE = 2.07 kcal $\cdot$ min⁻¹) and $R^2 = 0.83$ (SEE = 1.26 kcal $\cdot$ min⁻¹), respectively.

EE_{IC} , EE_{MAN} , and EE_{REC} were compared the two-way repeated ANOVA. The results showed a significant interaction of methods and speeds ($F_{10, 140}$

Figure 1 Two plots show the relationships between measurements by the IC and two estimations for the cross-validation group: (A) represents the relationship between the IC and the manufacture's software, and (B) represents the relationship between the IC and the recalibration model. For both plots, straight line is identity line ($y=x$), and dash line is regression line.

=22.8, $p<0.001$). **Table 5** presents mean EE_{IC} , EE_{MAN} , and EE_{REC} at each speed. Three methods were analyzed by one-way repeated ANOVAs and paired t -tests conducted for every speed. ANOVAs detected significant main effects at $0.0 \text{ km} \cdot \text{h}^{-1}$ ($F_{2,28} = 10.6$, $p<0.001$), $4.8 \text{ km} \cdot \text{h}^{-1}$ ($F_{2,28} = 38.4$, $p<0.001$), $6.0 \text{ km} \cdot \text{h}^{-1}$ ($F_{2,28} = 81.6$, $p<0.001$), and $9.6 \text{ km} \cdot \text{h}^{-1}$ ($F_{2,28} = 7.0$, $p=0.003$). Main effects at $3.6 \text{ km} \cdot \text{h}^{-1}$ ($F_{2,28} = 5.0$, $p=0.015$) and $7.2 \text{ km} \cdot \text{h}^{-1}$

($F_{2,28} = 0.0$, $p=1.000$) were not significant. For the manufacture's software estimations, significant overestimations compared with the IC measurements were at $4.8 \text{ km} \cdot \text{h}^{-1}$ ($p<0.001$) and $6.0 \text{ km} \cdot \text{h}^{-1}$ ($p<0.001$). Although main effects were significant at $0.0 \text{ km} \cdot \text{h}^{-1}$ and $9.6 \text{ km} \cdot \text{h}^{-1}$, paired t -tests did not detect significant differences between EE_{IC} and EE_{MAN} ($0.0 \text{ km} \cdot \text{h}^{-1}$: $p=0.006$; $9.6 \text{ km} \cdot \text{h}^{-1}$: $p=0.008$). Differences between EE_{IC} and EE_{REC}

Table 5 Means $\pm$ SD for energy expenditure for cross-validation group; results are shown for each speed and measurement method separately (N = 15)

Activities	Method	Mean $\pm$ SD (kcal/min)	Percent Difference (%)	Correlation (r)
0.0 km/h (Seating rest)	IC	1.3 $\pm$ 0.4		
	MAN	1.0 $\pm$ 0.2	-18 $\pm$ 30	0.32
	REC	1.3 $\pm$ 0.2	9 $\pm$ 37	0.36
3.6 km/h	IC	3.4 $\pm$ 0.9		
	MAN	4.0 $\pm$ 0.8	23 $\pm$ 33	0.50
	REC	3.7 $\pm$ 0.5	13 $\pm$ 26	0.57 [†]
4.8 km/h	IC	4.1 $\pm$ 0.9		
	MAN	6.1 $\pm$ 1.4	57 $\pm$ 48*	0.36
	REC	4.3 $\pm$ 0.7	9 $\pm$ 24	0.65 [†]
6.0 km/h	IC	5.4 $\pm$ 1.3		
	MAN	8.2 $\pm$ 1.4	56 $\pm$ 29*	0.51
	REC	5.3 $\pm$ 1.1	1 $\pm$ 15	0.79 [†]
7.2 km/h	IC	7.6 $\pm$ 1.7		
	MAN	7.6 $\pm$ 1.9	4 $\pm$ 29	0.01
	REC	7.6 $\pm$ 1.5	4 $\pm$ 27	0.04
9.6 km/h	IC	9.8 $\pm$ 2.0		
	MAN	8.4 $\pm$ 1.5	-13 $\pm$ 15	0.51
	REC	9.2 $\pm$ 1.2	-4 $\pm$ 15	0.56 [†]

IC: indirect calorimeter; MAN: manufacture's software; REC: recalibration model;

Percent Difference (%) = (Estimations - EE_{IC}) / EE_{IC} $\times$ 100;

* Significantly different from the EE_{IC} ($p < 0.004$);

[†] Significant correlation coefficient between the IC and estimations ($p < 0.05$).

were not significant at any speed ($p > 0.185$).

Table 5 also shows correlation coefficients between EE_{IC} and the two estimations. No significant correlation between EE_{IC} and EE_{MAN} was noted at any speed ($r \leq 0.51$, $p \geq 0.050$). EE_{REC} significantly correlated with EE_{IC} at all speeds ($r \geq 0.56$, $p \leq 0.030$), except 0.0 km $\cdot$ h⁻¹ ($r = 0.36$, $p = 0.187$) and 7.2 km $\cdot$ h⁻¹ ($r = 0.04$, $p = 0.878$).

Figure 2 presents a box plot of MSE of the manufacture's software and the recalibration model at each speed. At 0.0 km $\cdot$ h⁻¹, the median (interquartile range) of MSE_{MAN} was 0.2 (0.3), and MSE_{REC} was 0.1 (0.2). In walking and running, MSE_{MAN} vs MSE_{REC} were 1.0 (2.4) vs 0.7 (0.9) at 3.6 km $\cdot$ h⁻¹, 4.4 (8.6) vs 0.2 (0.7) at 4.8 km $\cdot$ h⁻¹, 7.7 (5.3) vs 0.3 (0.8) at 6.0 km $\cdot$ h⁻¹, 1.5 (7.7) vs 0.6 (6.9) at 7.2 km $\cdot$ h⁻¹, and 1.0 (6.9) vs 1.6 (3.5) at 9.6 km $\cdot$ h⁻¹.

4. Discussion

Major findings. We recalibrated the estimation model of the ViM, a device used to measure physical activity. We then examined the cross-validity of the recalibration model. In the recalibration group, the determination coefficients of the recalibration model used to measure METs (males: $R^2 = 0.74$; females: $R^2 = 0.84$) demonstrated that the recalibration model

Figure 2 The box plot shows medians, 25th and 75th percentiles, and range of mean squared error for the manufacture's software and the recalibration model at each treadmill speed for the cross-validation group.

could explain the change of METs during walking and running. In the cross-validation group, R^2 of the recalibration model ($R^2=0.83$) was approximately equal to those of the recalibration group. Furthermore, percent differences and MSE of the recalibration model were smaller than those of the manufacture's software. Therefore, we conclude that the recalibration model can improve the accuracy of PAEE estimation by the ViM.

Recalibration model. The recalibration models were calculated on the assumption that several activity patterns of the ViM can be unified respectively (W1 and IR1, J1 and IR2, J2 and IR3, W2 and W3). The occasion for the assumption is that biases inherent in the ViM outputs that stem from the exercise protocol may result in problems in the recalibration analysis (see 2.4 Data reduction). Justification for the assumption is not provided; however, the goodness of fit of the recalibration model suggests that the procedure is probably correct.

Interestingly, regression coefficients of x_1 (W1 + IR1) and x_2 (W2 and W3) yielded by Mixed model equally contributed to estimated METs in both males (coefficients of x_1 and $x_2=0.07$) and females (coefficients of x_1 and $x_1=0.08$). Coefficients of W2 and W3 should be higher than that of W1, provided that relationship between METs and walking speed is linear. Equal coefficients of W1, W2, and W3 appear to suggest that the gradation among W1, W2, and W3 is out of accord with increases of METs.

During walking, the ViM classifies walking intensity into 4 categories (W1, W2, W3, and W4). The classification algorithm is based on the relationship between the amplitude of arm swing increases and walking speed (Donker et al., 2001; Ford et al., 2007; Wagenaar and van Emmerik, 2000). Takahashi et al. (2009) showed that the ViM could distinguish walking speeds approximately correctly, whereas Takahashi and coworkers also reported that the accuracy of the ViM classification was inferior to that of a conventional accelerometer placed on waist. Because differences in arm swing motion among individuals were larger than those of acceleration at waist, standard deviations of W1, W2, and W3 became large in the previous study. There are similar results in this study (Table 3). The equal coefficients of x_1 and x_2 may be attributed to the large individual differences of arm swing in gait.

The large differences in arm swing among individuals may also influence the goodness of fit of the model. In females, determination coefficient ($R^2=0.84$) and SEE (1.16 METs) of the recalibration model are comparable to results in previous calibration studies (Freedson et al., 1998; Kumahara et al., 2004; Leenders et al., 2003; Nichols et al., 1999); however, it is likely that the goodness of fit in males ($R^2=0.74$; SEE=1.51 METs) is inferior to those of the previous studies. Most of the calibration studies that were based on treadmill exercise reported higher determination coefficients than $R^2=0.80$ and smaller SEE than 1.5 METs (Freedson et al., 1998; Leenders et al., 2003; Nichols et al., 1999). In the calibration studies, object devices were conventional accelerometers placed on waist. Because conventional uniaxial or triaxial devices (e.g. CSA, TriTrac, RT3, Lifecorder) are superior to the ViM in terms of the reliability, low goodness of fit of the recalibration model in males may reflect the characteristics of the ViM.

Cross-validation and accuracy. In cross-validation group, the determination coefficient between EE_{IC} and EE_{REC} was higher than that between EE_{IC} and EE_{MAN} , and was not lower than those of the recalibration group. When the recalibration model was compared with the manufacture's software for each speed, correlations of EE_{REC} were higher than that of EE_{MAN} . Percent differences and mean squared errors of the recalibration model were smaller than those of the manufacture's software. In particular, differences of MSE_{REC} and MSE_{MAN} were remarkable at 4.8 and 6.0 $km \cdot h^{-1}$ (Figure 2). The correlation coefficients and differences between EE_{IC} and EE_{REC} showed equally high correlation and smaller difference in comparison with other accelerometers (CSA, TriTrac, RT3, Lifecorder) (King et al., 2004; Takahashi et al., 2009), except at 7.2 $km \cdot h^{-1}$. These results demonstrate that the recalibration model solves the problem inherent in the manufacture's software, and is an accurate and valid physical activity energy expenditure estimation method.

On the other hand, although the difference between EE_{IC} and EE_{REC} were small (4%), the correlation between EE_{IC} and EE_{REC} is low ($r=0.04$) at 7.2 $km \cdot h^{-1}$. The likely cause is in the manufacture's software as well (percent difference = 4%; $r=0.01$). In the previous study (Takahashi et al., 2009), for males and females, percent differences

and correlations between the IC measurements and manufacture's software were 2 and 14% and $r=0.54$ and 0.76 at $7.2 \text{ km} \cdot \text{h}^{-1}$, respectively. To consider whether the disagreements in results of the present study and the previous study were systematic, we calculated correlation coefficients and percent differences between EE_{IC} ($7.8 \pm 1.9 \text{ kcal} \cdot \text{min}^{-1}$) and EE_{REC} ($8.0 \pm 1.6 \text{ kcal} \cdot \text{min}^{-1}$), and EE_{IC} and EE_{MAN} ($8.2 \pm 1.6 \text{ kcal} \cdot \text{min}^{-1}$) for the recalibration group data ($N=30$) at $7.2 \text{ km} \cdot \text{h}^{-1}$. EE_{REC} and EE_{MAN} were significantly correlated with EE_{IC} (EE_{REC} : $r=0.44$, $p=0.0022$; EE_{MAN} : $r=0.58$, $p<0.001$). There were no significant differences in two estimations and EE_{IC} , and the values were small (EE_{REC} : 4%; EE_{MAN} : 4%). The results of the post hoc analyses were similar to those of the previous study (Takahashi et al., 2009). Therefore, we consider that reasons that the correlations of EE_{REC} and EE_{MAN} to EE_{IC} were low at $7.2 \text{ km} \cdot \text{h}^{-1}$ in the cross-validation group were not systematic but random. It is possible that the cause of low correlations can be attributed to decrease of variance, due to characteristics of the participants in the cross-validation group.

Acknowledgments

The authors are grateful to the participants in this study. The study was supported by Grant-in-Aid for Young Scientists (B) from the Ministry of Education, Culture, Sports, Science, and Technology of Japan (2007-2009 Project number: 19700548).

Conflict of interest: The results of the present study do not constitute endorsement of the product by the authors or Japan Society of Physical Education, Health and Sport Sciences. No financial support for this study was obtained from Microstone Co.

References

- Brage, S., Wedderkopp, N., Franks, P. W., Andersen, L. B., and Froberg, K. (2003). Reexamination of validity and reliability of the CSA monitor in walking and running. *Med. Sci. Sports. Exerc.*, 35(8), 1447-1454.
- Campbell, K. L., Crocker, P. R., and McKenzie, D. C. (2002). Field evaluation of energy expenditure in women using Tritrac accelerometers. *Med. Sci. Sports. Exerc.*, 34(10), 1667-1674.
- Dale, D., Welk, G. J., and Matthews, C. E. (2002). Methods for assessing physical activity and challenges for research. in *Physical activity assessments for health-related research*. In Welk, G. J. (Ed.) (pp. 25). Champaign: Human Kinetics Publishers, Inc.
- Donker, S. F., Beek, P. J., Wagenaar, R. C., and Mulder, T. (2001). Coordination between arm and leg movements during locomotion. *J. Mot. Behav.*, 33(1), 86-102.
- Eston, R. G., Rowlands, A. V., and Ingledew, D. K. (1998). Validity of heart rate, pedometry, and accelerometry for predicting the energy cost of children's activities. *J. Appl. Physiol.*, 84(1), 362-371.
- Ford, M. P., Wagenaar, R. C., and Newell, K. M. (2007). Arm constraint and walking in healthy adults. *Gait Posture*, 26(1), 135-141.
- Freedson, P. S., Melanson, E., and Sirard, J. (1998) Calibration of the Computer Science and Applications, Inc. accelerometer. *Med. Sci. Sports. Exerc.*, 30(5), 777-781.
- Fujita, K., Takahashi, H., Miura, C., Ohkubo, T., Sato, Y., Ugajin, T., et al. (2004). Walking and mortality in Japan: the Miyagi Cohort Study. *J. Epidemiol.*, 14 Suppl 1, S26-S32.
- Grafen, A., and Hails, R. (2002). Random effects. in *Modern Statistics for the Life Science* (pp. 232-254). New York: Oxford University Press.
- Hendelman, D., Miller, K., Baggett, C., Debold, E., and Freedson, P. (2000). Validity of accelerometry for the assessment of moderate intensity physical activity in the field. *Medicine and Science in Sports and Exercise*, 32(9 Suppl), S442-S449.
- King, G. A., Torres, N., Potter, C., Brooks, T. J., and Coleman, K. J. (2004). Comparison of activity monitors to estimate energy cost of treadmill exercise. *Med. Sci. Sports. Exerc.*, 36(7), 1244-1251.
- Kumahara, H., Schutz, Y., Ayabe, M., Yoshioka, M., Yoshitake, Y., Shindo, M., et al. (2004). The use of uniaxial accelerometry for the assessment of physical-activity-related energy expenditure: a validation study against whole-body indirect calorimetry. *Br. J. Nutr.*, 91(2), 235-243.
- Leenders, N. Y., Nelson, T. E., and Sherman, W. M. (2003). Ability of different physical activity monitors to detect movement during treadmill walking. *Int. J. of Sports Med.*, 24(1), 43-50.
- Manson, J. E., Nathan, D. M., Krolewski, A. S., Stampfer, M. J., Willett, W. C., and Hennekens, C. H. (1992). A prospective study of exercise and incidence of diabetes among US male physicians. *JAMA*, 268(1), 63-67.
- Nichols, J. F., Morgan, C. G., Sarkin, J. A., Sallis, J. F., and Calfas, K. J. (1999). Validity, reliability, and calibration of the Tritrac accelerometer as a measure of physical activity. *Med. Sci. Sports. Exerc.*, 31(6), 908-912.
- Paffenbarger, R. S., Jr., Wing, A. L., Hyde, R. T., and Jung, D. L. (1983). Physical activity and incidence of hypertension in college alumni. *Am. J. Epidemiol.*, 117(3), 245-257.
- Pate, R. R., Pratt, M., Blair, S. N., Haskell, W. L., Macera, C. A., Bouchard, C., et al. (1995). Physical activity and public health. A recommendation from the Centers for Disease Control and Prevention and the American College of Sports Medicine. *JAMA*, 273(5), 402-407.
- Rowlands, A. V., Thomas, P. W., Eston, R. G., and Topping, R. (2004). Validation of the RT3 triaxial accelerometer for the assessment of physical activity. *Med. Sci. Sports. Exerc.*, 36(3), 518-524.
- Takahashi, S., Suzuki, K., and Kizuka, T. (2009). Evaluation of New Gyro-Sensor and Accelerometer Device to Estimate Physical Activity. *Int. J. Sport Health Sci.*, 7, 59-68.
- Twisk, J. W. R. (2003). Continuous outcome variables-relationships with other variables. in *Applied Longitudinal Data Analysis for Epidemiology: A Practical Guide* (pp. 77-88). New York: Cambridge University Press.
- Wagenaar, R. C., and van Emmerik, R. E. (2000). Resonant frequencies of arms and legs identify different walking patterns. *J. Biomech.*, 33(7), 853-861.

Name:
Shinji Takahashi

Affiliation:
Department of Regional Management,
Faculty of Liberal Arts, Tohoku Gakuin
University

Address:
2-1-1 Tenjinzawa, Izumi-ku, Sendai, Miyagi, Japan

Brief Biographical History:
2001-2004 Doctoral Program in Health and Sports Science,
University of Tsukuba
2004-2005 Researcher, Graduate School of Comprehensive
Human Sciences, University of Tsukuba
2005-2007 Lecturer, Department of Regional Management,
Faculty of Liberal Arts, Tohoku Gakuin University
2007- Associate Professor, Department of Regional
Management, Faculty of Liberal Arts, Tohoku Gakuin
University

Main Works:

- Validity of Dynamic Prediction Model for Oxygen Uptake during Supra Maximal Intermittent Load Exercise. International Journal of Sport and Health Science, Vol. 3: 68-74, 2005
- Expired gas kinetics during 20 m shuttle running test. Human Performance Measurement, Vol. 4, 9-16, 2007
- Evaluation of New Gyro-Sensor and Accelerometer Device to Estimate Physical Activity. International Journal of Sports and Health Science, Vol. 7, 59-68, 2009

Membership in Learned Societies:

- Japan Society of Physical Education, Health and Sports Science
 - Japan Society of Physical Fitness and Sports Medicine
 - Japanese Society of Test and Measurement in Health and Physical Education
 - American College of Sports Medicine
-