Journal of Physical Activity and Health, 2010, 7, 508-516

Errors in MET Estimates of Physical Activities Using 3.5 ml·kg^{-1}·min^{-1} as the Baseline Oxygen Consumption

Sarah Kozey, Kate Lyden, John Staudenmayer, and Patty Freedson

Purpose: To compare intensity misclassification and activity MET values using measured RMR (measMET) compared with 3.5 ml·kg^{-1}·min^{-1} (standMET) and corrected METs [corrMET = mean standMET × (3.5 ÷ Harris-Benedict RMR)] in subgroups. ***Methods:*** RMR was measured for 252 subjects following a 4-hr fast and before completion of 11 activities. VO_2 was measured during activity using indirect calorimetry (n = 2555 activities). Subjects were classified by BMI category (normal-weight or overweight/obese), sex, age (decade 20, 30, 40, or 50 y), and fitness quintiles (low to high). Activities were classified into low, moderate, and vigorous intensity categories. ***Results:*** The (mean ± SD) measMET was 6.1 ± 2.64 METs. StandMET [mean (95% CI)] was (0.51(0.42, 0.59) METs) less than measMET. CorrMET was not statistically different from measMET (–0.02 (–0.11, 0.06) METs). 12.2% of the activities were misclassified using standMETs compared with an 8.6% misclassification rate for METs based on predicted RMR ($P < .0001$). StandMET differences and misclassification rates were highest for low fit, overweight, and older individuals while there were no differences when corrMETs were used. ***Conclusion:*** Using 3.5ml·kg^{-1}·min^{-1} to calculate activity METs causes higher misclassification of activities and inaccurate point estimates of METs than a corrected baseline which considers individual height, weight, and age. These errors disproportionally impact subgroups of the population with the lowest activity levels.

Keywords: metabolic equivalent, compendium of physical activities, intensity classification

Physical inactivity is a risk factor for obesity, chronic disease, and mortality. As a result, increasing and measuring population levels of physical activity (PA) is a public health priority. A major challenge for researchers is translating the oxygen cost of physical activities into metrics that are easily interpreted by clinicians and the general population. The metabolic equivalent (MET) provides a convenient way to describe and classify PA by expressing activity energy expenditure (EE) as multiples of resting metabolic rate (RMR). One MET is defined as the amount of oxygen consumed at rest per kg of body weight.[1] When RMR is not measured, a reference baseline is assumed where 1 MET = 3.5 ml·kg^{-1}·min^{-1} (4.184 KJ kg^{-1}·hr^{-1}).

The Physical Activity Guidelines Advisory Committee Report describes the dose of activity eliciting health benefits based on METS and classifies intensity of activity as light, moderate, vigorous, and very vigorous (<3, 3–5.99, 6–8.99, and ≥9 METs, respectively).[2] An additional use of METs is in self-report questionnaires to translate the volume of activity into MET hrs per week or to quantify time spent in different MET based intensity categories. The Compendium of Physical Activities was developed to standardize self-reports by assigning MET values for a comprehensive list of activities and transforming METs to EE (1 MET = 1 kcal·kg^{-1}·hr^{-1}).[3,4]

Both applications assume the reference baseline of 1 MET = 3.5 ml·kg^{-1}·min^{-1}, which overestimates RMR for many individuals.[5–11] If the reference baseline is too high for a population or specific subgroups of the population, PA METs will be underestimated.[3,4] For example, RMR has been shown to decline with age, is lower in overweight individuals, and is lower in females compared with males.[12] Furthermore, the effect of a slight difference at rest between measured RMR and the reference baseline is magnified as the oxygen cost of an activity increases. For example, when using the standard reference baseline (3.5 ml·kg^{-1}·min^{-1}) an activity requiring 20 ml·kg^{-1}·min^{-1} is classified as moderate intensity (5.5 METs). However, an individual with a measured RMR of 2.5 ml·kg^{-1}·min^{-1} performing the same activity is actually at a vigorous level of exertion (8 METs). This small 1 ml·kg^{-1}·min^{-1} difference at rest results in a 2.5 MET difference during activity, resulting in an exercise prescription that is too high for an individual, possibly contributing to poor exercise adherence.[13,14] Additionally, activity METs based on the reference baseline, as is done in the Compendium of Physical Activities, leads to errors quantifying PA and EE in surveillance research,[15,16] and in evaluating of prevalence of individuals who are satisfying PA guidelines.[17]

Kozey, Lyden, and Freedson are with the Dept of Kinesiology, University of Massachusetts, Amherst, MA. Staudenmayer is with the Dept of Mathematics and Statistics, University of Massachusetts, Amherst, MA.

Limited data have assessed the impact of the 'standard' reference baseline when it is applied to activity METs. One research group examined 4 moderate intensity activities, but did not provide comparisons across subgroups who have a lower RMR.[7–10] Byrne and colleagues[18] reported a 22% higher measured MET value (measured VO_2/measured RMR) compared with the Compendium MET for walking 5.6 km·hr^{-1}. They reduced the difference between Compendium and measured METs to 0.2% by applying a simple correction based on the Harris-Benedict prediction equation for RMR (corrected MET).[19] If this correction factor is valid for a range of activities, then it will improve MET estimates in survey research and population surveillance without requiring measurement of RMR for an individual.

The purpose of this study was to evaluate the difference in MET estimates when 3.5 ml·kg^{-1}·min^{-1} is used as the reference baseline compared with using measured RMR and determine if the error can be reduced using the Harris-Benedict prediction equation. We assessed MET differences in the 2 main applications of the MET; intensity classification and average activity METs for the total sample. To determine if errors differentially affected subgroups, intensity classification, and average activity METs were compared across subgroups classified on the basis of age, body mass index (BMI), fitness, and sex.

Methods

Subjects

Two hundred and fifty-two healthy participants between the ages of 20 to 60 y were recruited from the Amherst, Massachusetts area. All participants read and signed an informed consent document approved by the University of Massachusetts Institutional Review Board and completed a health history questionnaire, the PA Readiness Questionnaire (PAR-Q) and a questionnaire to evaluate habitual PA status.[20] Before completing the study protocol, female participants over 50 y and male participants over 40 y were screened with a physician-supervised 12-lead ECG stress test where subjects exercised to 90% of age-predicted maximum heart-rate. Participants were excluded if they had a condition impairing their ability to safely exercise, were taking medication altering metabolic rate or if the physician identified cardiovascular abnormalities that potentially prevented safely exercising in the activity trial.

Protocol

RMR was measured using the Med Gem (Microlife USA, Dunedin, FL) a handheld indirect calorimeter, after a 15 minute rest on a bed in a dark, quiet room following a 4-hr restriction of food, caffeine, and exercise. These restrictions are consistent with the recommendations for measurement of RMR.[21] Participants were supine during the measurement period, which lasted between 5 to 10 minutes. The MedGem begins data collection when it detects the first breath, and continues until a steady-state or 10 minutes is reached. The first 2 minutes are not included in the determination of steady-state. The Med Gem measures oxygen concentration in the inspired and expired air via a proprietary fluorescent-quenching sensor. When the active and reference ruthenium cells are excited by an internal light source, they fluoresce. The reaction is quenched in the presence of oxygen and the degree of quenching is inversely proportional to the concentration of oxygen present. The volume of inspired and expired air is measured using ultrasonic sensing technology with transducers at each end of a flow tube that emits sound impulses. The transmission time of the sound waves from the sending to receiving transducer is increased or decreased proportional to the rate and direction of gas flow. The MedGem calculates oxygen consumption based on a modified Weir equation and a fixed respiratory exchange ratio of 0.85. The MedGem has been validated against the "gold standard" Douglas Bag method in 3 studies on adult subjects. The average difference between MedGem and Douglas bag in adults was less than 1% and the intraclass reliability coefficient was $r = .98$.[22–24] A review of 12 studies that validated the MedGem against criterion methods concluded it was an accurate and reliable device for measuring RMR.[25]

Activity Protocol

The activity protocol was divided into 2 sections (Part A and Part B). All activities, except ascending and descending the stairs were performed for 7 minutes with a 4 minute rest between each activity. There was a 15-minute rest period between sections to recalibrate the metabolic measurement system and allow for subjects to rehydrate. The section order (A or B) was balanced across subjects.

Part A. Participants completed 6 treadmill activities at 3 speeds (1.34, 1.56, 2.23 m·sec^{-1}) each at 0% and 3% grade. The order of treadmill conditions was randomized between participants. If HR exceeded the limit based upon the stress test (above 90% of age predicted maximum heart-rate) or the participant could not safely complete the activity, it was eliminated.

Part B. Participants completed 5 self-paced ADLs. Three activities were completed by all participants 1) ascending 16 flights of stairs, 2) descending 16 flights of stairs, 3) moving a 6-kg box back and forth 8 meters between a shelf and the floor. The remaining 2 ADLs were chosen using a blocked randomized design to ensure activities were completed equally among age and sex groups. There were 14 possible household and sporting activities including sweeping, mopping, gardening, trimming, mowing, raking, dusting, laundry, vacuuming, washing dishes, painting, tennis, or basketball.

Oxygen consumption during activity was measured using a portable metabolic system (Oxycon Mobile; Cardinal Healthcare, Yorba Linda, California). This device is a battery-operated, portable, and wireless unit that measures breath-by-breath gas exchange. It is secured to the body using a vest similar to a backpack and weighs 950 g.[26] A face mask (Hans Rudolf, Inc., Kansas City, MO) connects to the flow sensor unit and detects the airflow by the rotation of a low-resistance, bidirectional turbine to measure expired ventilation. The expired air is analyzed for O_2 and CO_2 concentrations using a microfuel O_2 sensor and a thermal conductivity CO_2 sensor. Immediately before data collection, a 2-point (0.2 and 2.0 $L \cdot s^{-1}$) airflow calibration was performed using the automatic flow calibrator, and the gas analyzers were calibrated using a certified gas mixture of 16% O_2 and 4.01% CO_2, and a measurement delay time was determined.[27] This metabolic measurement system has been shown to be a valid device to measure oxygen consumption.[26] Differences between the Oxycon Mobile and Douglas Bag measures during cycling at varying intensities were 1.4%, 1.1%, and 0.5% for V_E, VO_2, and VCO_2, respectively.[28] This is within the 5% acceptable physiological range for variability and comparable with other portable measurement systems.[29]

Analysis

Participant Sub-Groups.

Age: Participants were classified into age groups by decade 20–29, 30–39, 40–49, and 50–60 y. (20, 30, 40, and 50 y, respectively)

Sex: Participants were classified by sex.

BMI: Height (cm) was measured with a stadiometer and body weight (kg) was measured using a balance scale (Detecto, Webb City, MO). Participants were classified into weight categories based on body mass index (BMI). Individuals with a BMI < 25 $kg \cdot m^{-2}$ were considered normal weight (NW) and those with a BMI ≥ 25 $kg \cdot m^{-2}$ were considered overweight/obese (OW).

Fitness quintile: A prediction equation based on age, height, sex, body mass, and PA status was used to categorize participants into one of 5 quintiles (Q1 is the lowest fitness group and Q5 is the highest). This equation was developed and validated in a similar sample.[30]

Descriptive Characteristics

The descriptive characteristics for the 252 participants are presented in Table 1. The mean ± SD for age was 38 ± 12.45 y, BMI was 24.5 ± 3.90 $kg \cdot m^{-2}$, and PAS score was 5.2 ± 2.03. A PAS score of 5 corresponds to jogging 1 to 5 miles per week or spending 30 to 60 minutes per week in vigorous activity weekly. The mean and SD for age and BMI are also presented in Table 1 for each subgroup. The differences in RMR and demographic characteristics were assessed using a 4-way ANOVA to verify differences in RMR within subgroups (eg, those with a higher BMI had a lower RMR), that the variable of interest in the subgroup was statistically different (eg, the low BMI group had a lower BMI than the high BMI group), and to examine potential confounding effects between subgroups (eg, that the males and females had the same average age and BMI).

Data Reduction

MET value for each activity was determined by removing the first 120 seconds (nonsteady state) and the last 10 seconds and averaging the remaining 290 seconds of metabolic data. Four different methods to compute MET values were used:

Table 1 Descriptive Characteristics of Participants

Sample	N	Age (±SD) (y)	BMI (±SD) $kg \cdot m^{-2}$	PAS	RMR (±SD) $ml \cdot kg^{-1} \cdot min^{-1}$
Total	252	38.0 (12.45)	24.5 (3.90)	5.2 (2.01)	3.3 (0.49)
20y	70	22.7 (2.81)[a]	24.1 (3.65) [ns]	5.0 (2.03)	3.5 (0.48) [a]
30y	69	33.3 (3.13)[b]	23.8 (3.52) [ns]	5.2 (2.22)	3.3 (0.47) [a]
40y	53	45.0 (3.18)[c]	25.4 (4.90) [ns]	5.5 (1.83)	3.2 (0.51)[a]
50y	59	54.6 (3.08)[d]	24.1 (3.43) [ns]	5.3 (1.89)	3.1 (0.37)[b]
Males	118	37.3 (12.77)[ns]	24.7 (3.32) [ns]	5.7 (1.71)	3.4 (0.50) [a]
Females	134	38.6 (12.18) [ns]	24.4 (4.35) [ns]	4.8 (2.16)	3.2 (0.46) [b]
NW	159	37.1 (12.37) [ns]	22.2 (1.67)[a]	5.5 (1.90)	3.5(0.45) [a]
OW	93	39.4 (12.52) [ns]	28.5 (3.39)[b]	4.8 (2.16)	3.0(0.39) [b]
Q1	20	41.3 (8.78) [a]	31.4 (4.68)[a]	2.5 (1.79)	2.7 (0.28)[a]
Q2	31	37.0 (9.53) [a]	25.8 (4.35) [b]	3.4 (2.36)	3.2 (0.41)[b]
Q3	73	35.1 (11.0) [a]	24.5 (3.21) [b]	5.1 (1.70)	3.3 (0.45)[a]
Q4	88	34.7 (12.02) [a]	23.1 (2.58) [b]	6.1 (1.29)	3.5 (0.45)[c]
Q5	39	50.2 (10.71) [b]	23.3 (2.53) [b]	6.3 (1.11)	3.3 (0.39)[c,d]

Note. Significant differences at $P < .01$ are denoted with a letter representing which groups are different. PAS is physical activity status, a self-report indicator of habitual activity from 0 to 7. Five is equivalent to jogging 30 to 60 minutes per week. Q1 is the lowest fitness quintile and Q5 is the highest fitness quintile. Age categories were defined by decade; 20 y is 20 – <30. BMI classifications were NW <25 $kg \cdot m^{-2}$ OW ≥ 25 $kg \cdot m^{-2}$.

Measured MET = Measured VO_2 ml·kg^{-1}·min^{-1} ÷ Measured RMR,

Standard MET = Measured VO_2ml·kg^{-1}·min^{-1} ÷ 3.5 ml·kg^{-1}·min^{-1}

Predicted MET = Measured VO_2ml·kg^{-1}·min^{-1} ÷ Harris-Benedict predicted RMR

Corrected MET = Mean standard MET per activity × (3.5/Harris-Benedict predicted RMR)

Statistical Analyses

To determine errors in classifying intensity we determined whether the intensity classification of the criterion measured METs was the same as the standard METs and predicted METs. The first hypothesis was that more activities are misclassified using standard METs compared with predicted METs. An activity was considered misclassified if it was placed in a different intensity category than the one established by measured METs. The percentage of activity intensities that were misclassified was analyzed using a repeated measures logistic regression. The rates at which the standard and predicted METs misclassified the MET intensity level relative to measured METs were reported as percentages (100 × number of misclassified activities/total number of activities). The repeated measures logistic regression assesses the statistical significance of the differences between the percentages. These models analyze binary data, and the response is 1 if the MET level is misclassified compared with measured METs and 0 if the classification agrees with measured METs. The covariates in the model are measurement method (standard and predicted) and subject group (fitness status, BMI category, age group, and sex). Additional analyses examined the direction (over or underestimation) of the misclassification.

To determine the error in average activity METs we compared the mean measured MET for each activity to the mean standard MET and mean corrected MET. The second hypothesis was that the average difference between standard METs and measured METs is greater than the average difference for corrected METs and measured METs. The differences in the average activity METs were analyzed with a repeated measures linear regression model. The responses in this model were differences between each of the alternative MET calculation methods and measured METs: standard METs minus measured METs, and corrected METs minus measured METs for each activity. Covariates were the type of difference (of the 2 types above) and subject groups (fitness status, BMI category, age group, and sex). Significance levels were set at $P < .05$ unless otherwise indicated.

Results

The mean RMR values for the total sample and for each category are presented in the Table 1. There were significant main effects for age, BMI, sex, and fitness on RMR. Specifically, RMR was 7% lower in 50 year olds than all other age groups, 6% lower for women than men, and 15% lower in OW individuals compared with NW. RMR for Q1 was 15% to 23% lower than all other quintiles.

The total possible number of activities was 2772 (252 subjects × 11 activities). One hundred and 10 activities were not completed due to participant or researcher decisions (eg, uncomfortable running or HR above 90% of age predicted maximum), 46 activities were eliminated because participants did not perform the activity for > 200 seconds (eg, stopped stair ascent due to volitional fatigue), and 61 activities were eliminated due to data collection problems (eg, mask leaking). The final sample of 2555 activities was used in subsequent analyses.

The misclassification rates are presented in Figure 1. Overall, 12.2% of the activity intensities were misclassified using standard METs and 8.6% of the activity intensities were misclassified using predicted METs ($P < .0001$). For standard METs, 34.8% of the activity intensities were misclassified for low fit individuals compared with 11.3% of the activity intensities for high fit individuals, while there were no differences across fitness groups for predicted METs (13.2% of activity intensities were misclassified for low fit, 14.8% of the activities were misclassified for the high fit). Results were similar for BMI category with lower misclassification in NW vs. OW (8.5% and 19.1%) men compared with women (9.7% and 14.8%) and older compared with younger individuals (vs. 16.4% for 50 y and 7.6% for 20 y) for standard METs. No significant differences in misclassification rates among groups were observed for predicted METs (see Figure 1). Of the activities that were misclassified using each method, the proportion that overestimated and underestimated METs are presented in Figure 2. Overall 89.5% of the misclassified standard METs were underestimated compared with measured METs, and 59.0% were underestimated using predicted METs. The mean MET values for each activity are shown in Figure 3. In support of the first hypothesis, standard METs were significantly lower than measured METs for all activities ($P < .05$ for all activities except for gardening, $P = .06$, and painting, $P = .06$). Predicted METs were not statistically different from measured METs for any activity.

The mean MET differences for each subgroup are presented in Table 2. The mean measured MET value across all activities and individuals was (mean ± SD) 6.1 ± 2.64. The MET difference [mean (± 95% CI)] for measured MET-corrected MET was [–0.02 (–0.11, 0.06) METs] across all observations. MET difference for measured MET – standard MET was [0.51 (–0.59, –0.42) METs], which was significantly different than corrected METs. The difference between measured and standard METs was greatest for individuals who were the least fit [–1.04 (–1.38,–0.71) METs] compared with [–0.32 (–0.48, –0.16 METs med-high fit)] and [–0.57(–0.81, –0.32) METs high fit], There were no significant differences in mean MET values between women, BMI classifications, and age groups. There were no differences in mean MET values within subgroups for corrected METs.

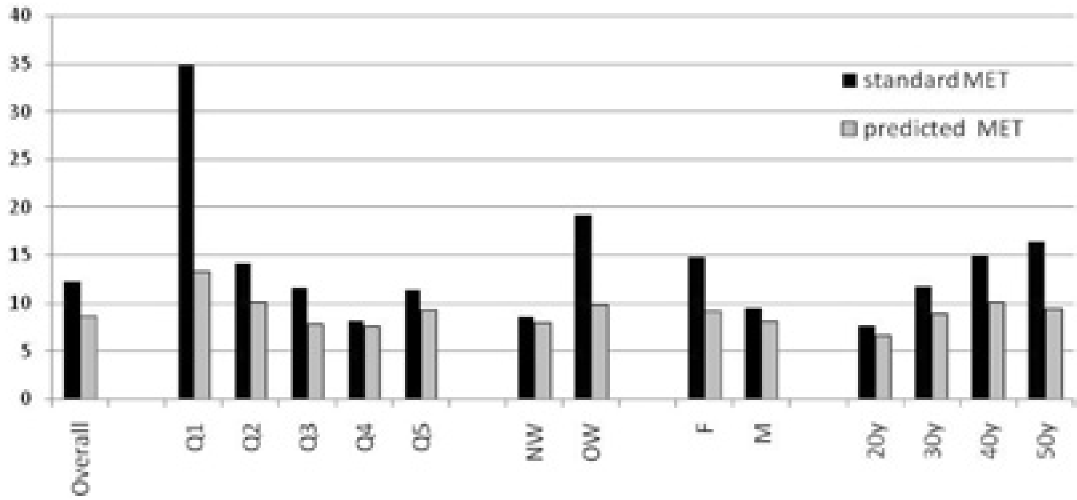

Figure 1 — Misclassification of activity intensity using standard MET (3.5 ml·kg^{-1}·min^{-1}) or predicted MET (Harris-Benedict equation) compared with measured RMR overall and within subgroups. Q1 is the lowest fitness quintile and Q5 is the highest fitness quintile. Age categories were determined by decade (eg, 20 y is 20 – <30 y). BMI classifications were NW BMI < 25 kg·m^{-2} OW BMI ≥ 25 kg·m^{-2}

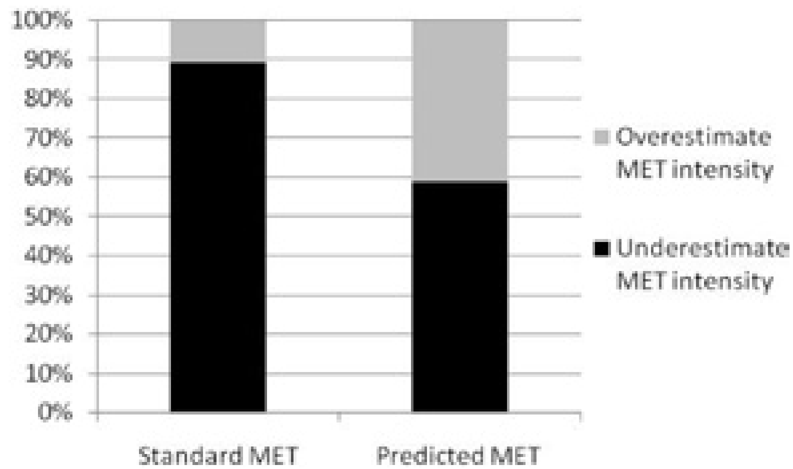

Figure 2 — Proportion of over- and under-estimation of MET intensity misclassifications using Standard MET (3.5 ml·kg^{-1}·min^{-1}) or Predicted MET (Harris-Benedicted equation) compared with Measured RMR.

Discussion

The major finding of this study was that use of the reference baseline 3.5 ml·kg^{-1}·min^{-1} to define METs leads to misclassification of activity intensity and underestimation of average MET values compared with measured METs. These errors disproportionately affect subgroups of the population who are the least likely to be meeting PA recommendations. The use of a simple correction factor based on the Harris-Benedict equation reduced the misclassification rate, improved average MET estimates, and eliminated the differences within subgroups compared with the standard METs.

The first application of this error is seen in exercise prescription based on MET intensity cut-offs. The PA guidelines refer to the Compendium of Physical Activities to identify activities within the defined intensity ranges.[2] However, the Compendium of Physical Activities uses MET values based on the reference baseline, and therefore, based on the results of this study, individuals are often prescribed exercise at an absolute intensity higher than what is recommended. In this investigation when misclassification of intensity occurred, the standard MET was lower than the measured MET value 89.5% of the time. When a predicted MET was used the error was more evenly distributed, with 59% of MET values underestimated. The underestimation of intensity that occurs when standard METs are used could contribute to the high level of attrition in individuals starting exercise programs considering that an inverse relationship

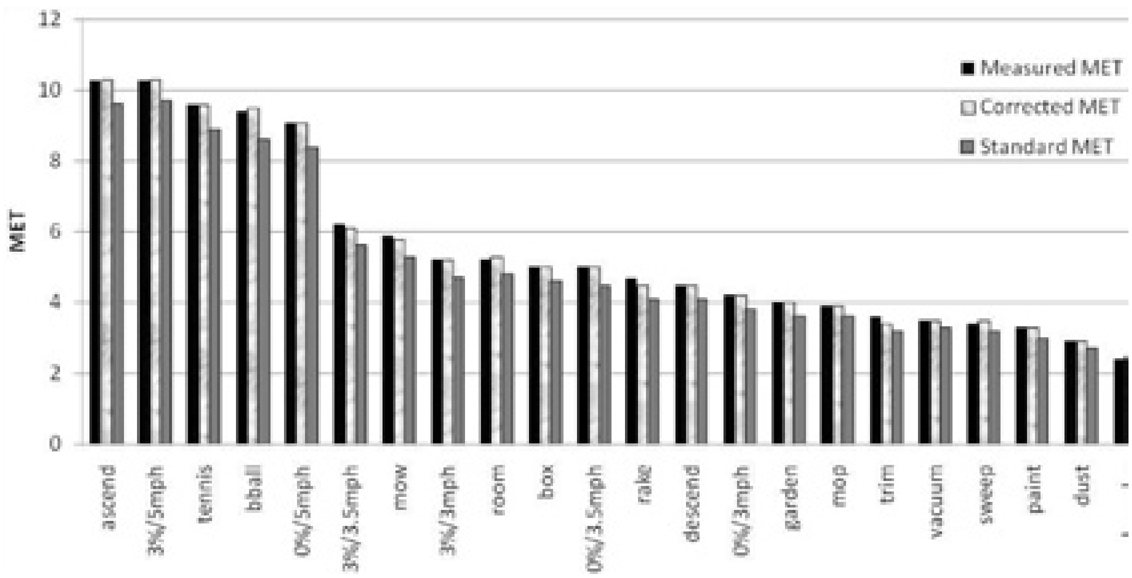

Figure 3 — Average MET values for each activity. All standard METs were significantly lower than measured METs ($P < .05$) except gardening and painting. Corrected METs and measured METs were not significantly different.

Table 2 Differences Between Corrected METs and Measured METs, and Calculated METs and Measured METs

Group (n observations)		Average measured MET (SD)	Corrected–measured		Standard–measured	
			Average difference (METs)	95% CI	Average difference (METs)	95% CI
Overall	(n = 2555)	6.1 (2.64)	–0.02	–0.11,0.06	–0.51	–0.59, –0.42
Fitness level	Q1 (n = 181)	6.1 (2.31)	0.29	–0.08, 0.66	–1.04	–1.38, –0.71
	Q2 (n = 305)	5.9 (2.43)	0.15	–0.15, 0.44	–0.48	–0.75, –0.21
	Q3 (n = 785)	6.2 (2.73)	–0.10	–0.29,0.11	–0.57	–0.75, –0.40
	Q4 (n = 912)	6.0 (2.70)	–0.07	–0.24,0.11	–0.32	–0.48, –0.16
	Q5 (n = 399)	6.1 (2.61)	0.05	–0.21,0.31	–0.57	–0.81, –0.32
Sex	Females (n = 1340)	6.0 (2.50)	0.10	–0.04, 0.23	–0.55	–0.68, –0.41
	Males (n = 1215)	6.1 (2.79)	–0.12	–0.26, 0.03	–0.46	–0.60, –0.32
Age	20 y (n = 748)	6.0 (2.76)	–0.18	–0.37, 0.01	–0.34	–0.53, –0.13
	30 y (n = 689)	6.1 (2.63)	–0.17	–0.36, 0.02	–0.49	–0.68, –0.31
	40 y (n = 546)	6.1 (2.64)	0.20	–0.02, 0.42	–0.57	–0.79, –0.36
	50 y (n = 574)	6.0 (2.49)	0.21	0.00, 0.42	–0.67	–0.87, –0.47
BMI	NW (n = 1662)	6.1 (2.67)	–0.17	–0.29, –0.04	–0.43	–0.54, –0.30
Class	OW (n = 893)	6.0 (2.60)	0.29	–0.13, –0.46	–0.65	–0.81, –0.49

Note. Average measured MET value is the mean of all observations in a category. Q1 is the lowest fitness quintile and Q5 is the highest fitness quintile. Age categories were determined by decade 20y is 20 – <30. BMI classifications were NW BMI <25 kg·m^{-2} OW BMI ≥ 25 kg·m^{-2}.

between intensity of PA and adherence has been reported in the literature.[13,14,31]

The second application of MET errors is in translation of data collected with PA questionnaires. According to our results, self-report instruments that transform activity into MET metrics (eg, MET hrs·day^{-1}) using the Compendium of Physical Activities which again, assumes the 3.5 ml·kg^{-1}·min^{-1} baseline, will result in inaccurate estimates of PA particularly in individuals who are overweight and low fit. This is consistent with previous investigations that concluded questionnaires based on METs, underestimate EE for obese women compared with doubly labeled water.[15,16,32,33] Racette et al[33,37] showed that correcting the reference baseline for measured RMR reduces the underestimation of EE for obese women. Similarly, in this study a simple correction factor based on predicted RMR corrected the difference between average measured and standard METs. This suggests the underestimation of METs reported in free-living situations can be attributed to an unrepresentative reference baseline. In our study, the reduction in error when the correction factor was used was similar to what Byrne et al[18] reported. In their sample, the difference between corrected and measured MET values for walking at 1.56 m·sec^{-1} was 0.2% compared with 0.06% in our investigation.

The Compendium of Physical Activities is an invaluable resource that has standardized the way PA survey research is reported. It is impossible for any self-report measure to accurately assign MET values for every individual and activity and there will always be individual differences in population level surveillance. However, the consistent underestimation of standard METs, is concerning because there is an uneven distribution of error between subgroups. The simple correction factor first proposed by Byrne and colleagues can be applied to existing Compendium estimates to improve MET estimates overall and within subgroups. The correction factor considers individual demographic characteristics including age, height, weight, and sex that are easily obtained and may improve individual estimates of activity levels within a population sample.

This study had numerous strengths including the large sample size (n = 252), and a diverse range of activities including treadmill, household, and sporting activities. Participants completed these activities following minimal instruction to ensure they were as representative of daily living as possible. Oxygen consumption was directly measured for over 2500 activities. The metabolic system is lightweight and portable and should not significantly affect the way the participant completes the activities. RMR was also measured directly, using the MedGem, a portable indirect calorimeter. The average RMR for our sample (3.3 ml·kg^{-1}·min^{-1}) was higher than previous investigations in middle-aged adults,[7–10,18] older adults,[11] and coronary heart disease patients.[5] This is potentially a result of differences in pretesting conditions. Our participants fasted and refrained from exercise for 4 hours, consistent with recommendations for RMR,[5] while in other investigations participants fast for 12 hours and obtain the measurement in the early morning, consistent with recommendations for basal metabolic rate. On average, there is an additional 61 kcal·day^{-1} cost associated with holding the MedGem device. However, by having participants supine rather than seated (average 70 kcal·day^{-1} less), the values offset one another, resulting in a valid resting value consistent with the definition of a MET.[25] Given that 1 MET is considered resting EE while seated, the protocol for RMR and use of the MedGem is an appropriate method to establish a baseline value. Our results were consistent with Byrne et al[18] who showed RMR was lowest in overweight and older individuals. A unique feature of our study was the examination of the effect of aerobic fitness on RMR and corresponding MET estimates. The nonexercise prediction measure we used to generate fitness categories is a simple, cost effective way to estimate fitness and identifies individuals who are likely to have activity intensity misclassified. However, future research should examine fitness as determined by VO_2max or treadmill time to further explore the relationship between fitness and RMR. Both the fitness prediction equation and the correction factor use variables that are easily measured in surveillance research, suggesting it is feasible to correct for errors in RMR in large epidemiologic studies.

Although this study has numerous strengths, there are limitations to note as well. The MedGem is valid for measurement of RMR however, it is not without limitations. The MedGem does not measure VCO_2 directly and assumes a constant respiratory exchange ratio (RER) of 0.85. Analyses done comparing computed oxygen consumption using the actual RER compared with computed oxygen consumption using an RER of 0.85 revealed an average difference of 0.14% (range –2.4 to 1.7%), which is within the accepted range of physiologic variability.[29] The Douglas Bag method is the gold-standard for measurement of RMR, however it requires is a cumbersome technique that requires an extensive time period for collecting and analyzing data and calibration. The MedGem self-calibrates, is inexpensive and easy to use, reducing the burden on participants who completed a lengthy protocol. Using separate devices to measure RMR and activity oxygen consumption introduces a source of error. However, the Oxycon Mobile which we used for the activity measures has been validated during activity and not at rest.[28] Although the metabolic measurement system is lightweight (950 g) and is worn like a backpack, the facemask and additional weight may alter the way participants complete activities. Another potential limitation of this study is that the participants were a convenience sample of adults volunteering for a study involving PA. Significant efforts were made to recruit participants of all fitness levels and BMI categories however, it is likely our sample was more fit and active than the general population. Inclusion of less fit, active subjects would likely increase the amount of misclassification and errors in MET estimates, thus the prevalence of these errors may be underestimated in our sample. Activities were performed in a laboratory setting, which may be different than a free-living environment. All participants were instructed to "do the activities as they would in everyday life" and previous research suggests there is not a systematic difference in energy cost of activities performed in the laboratory compared with at home.[7–10] Future research should examine the efficacy of a correction factor based on measurement error modeling of anthropometric variables to improve the simple correction factor used in this study. In addition, researchers should compare EE estimates from self-report instruments with and without the correction factor to determine if the correction factor is valid in a free-living setting.

It is important to consider the public health impact of underestimating MET values. Incorrect MET estimates may lead to errors in population surveillance and quantifying PA pattern changes following intervention. When self-report measures are converted into PA metrics such as MET hrs·week^{-1} and kcal kg·hr^{-1}, an additional source of error is introduced when the reference baseline is used. The following example illustrates the discrepancy between measured METs and standard METs in PA assessment. An average activity score from the 7 day PAR using standard METs is 263 MET hrs·week^{-1}.[34] If measured METs were used the score would be 284 MET hrs·week^{-1} for an "average" individual, 291 MET hrs·week^{-1} for an overweight individual and 308 MET hrs·week^{-1} for a low fit individual, resulting in a discrepancy between actual EE and what is captured by the 7-day PAR. For the "average" individual, this is 3 MET hrs·day^{-1} that are expended and not captured by the 7-day PAR, an amount exceeding the current PA guidelines. We are not suggesting that the average individual is currently achieving 30 minutes of moderate activity 5 days per week above what they report. There are additional factors affecting the accuracy of self-report measures of PA include social desirability and recall bias.[35,36] and strong objective data show the general population is currently not sufficiently active.[17,37] However, this significant difference between measured METs and standard METs could lead to errors quantifying changes in activity levels following an intervention and errors when self-report measures are used to estimate energy balance. Investigations comparing individuals from different subgroups (eg, normal weight vs. overweight) will result in a skewed distribution of the error between groups. Furthermore, the subgroups of the population who typically are the least active[17] are the most likely groups to be misclassified and have the greatest average difference in MET estimates

Conclusion

In conclusion, the results of this study suggest that care should be taken to consider the characteristics of an individual (weight status, fitness, sex, and age) when prescribing activity based on MET intensity classifications to ensure an appropriate intensity level. Future survey research should consider application of this correction factor to ensure accurate estimates of EE and improve population surveillance estimates of PA exposure.

Acknowledgments

Thank you to the graduate and undergraduate students who assisted in data collection and the study participants. Funding: NIH R01 CA121005.

References

1. Jette M, Sidney K, Blumchen G. Metabolic equivalents (METS) in exercise testing, exercise prescription, and evaluation of functional capacity. *Clin Cardiol*. 1990;13(8):555–565.
2. Physical Activity Guidelines Advisory Committee report, 2008. To the Secretary of Health and Human Services. Part A: executive summary. *Nutr Rev*. 2009;67(2):114–120.
3. Ainsworth BE, Haskell WL, Leon AS, et al. Compendium of physical activities: classification of energy costs of human physical activities. *Med Sci Sports Exerc*. 1993;25(1):71–80.
4. Ainsworth BE, Haskell WL, Whitt MC, et al. Compendium of physical activities: an update of activity codes and MET intensities. *Med Sci Sports Exerc*. 2000;32(9, Suppl):S498–S504.
5. Savage PD, Toth MJ, Ades PA. A re-examination of the metabolic equivalent concept in individuals with coronary heart disease. *J Cardiopulm Rehabil Prev*. 2007;27(3):143–148.
6. Howley ET. You asked for it: question authority. *ACSM Health Fitness Journal.* 2000;4:6–8.
7. Brooks AG, Withers RT, Gore CJ, Vogler AJ, Plummer J, Cormack J. Measurement and prediction of METs during household activities in 35- to 45-year-old females. *Eur J Appl Physiol*. 2004;91(5-6):638–648.
8. Gunn SM, Brooks AG, Withers RT, et al. Determining energy expenditure during some household and garden tasks. *Med Sci Sports Exerc*. 2002;34(5):895–902.
9. Gunn SM, Brooks AG, Withers RT, Gore CJ, Plummer JL, Cormack J. The energy cost of household and garden activities in 55- to 65-year-old males. *Eur J Appl Physiol*. 2005;94(4):476–486.
10. Withers RT, Brooks AG, Gunn SM, Plummer JL, Gore CJ, Cormack J. Self-selected exercise intensity during household/garden activities and walking in 55- to 65-year-old females. *Eur J Appl Physiol*. 2006;97(4):494–504.
11. Kwan M, Woo J, Kwok T. The standard oxygen consumption value equivalent to one metabolic equivalent (3.5 ml/min/kg) is not appropriate for elderly people. *Int J Food Sci Nutr*. 2004;55(3):179–182.
12. Powers SK, Howley ET. *Exercise Physiology: Theory and Application to Fitness and Performance*. 4 ed. Boston, MA: McGraw-Hill; 2001.
13. Dishman RK, Buckworth J. Increasing physical activity: a quantitative synthesis. *Med Sci Sports Exerc*. 1996;28(6):706–719.
14. Ekkekakis P, Lind E. Exercise does not feel the same when you are overweight: the impact of self-selected and imposed intensity on affect and exertion. *Int J Obes (Lond)*. 2006;30(4):652–660.
15. Conway JM, Seale JL, Jacobs DR, Jr, Irwin ML, Ainsworth BE. Comparison of energy expenditure estimates from doubly labeled water, a physical activity questionnaire, and physical activity records. *Am J Clin Nutr*. 2002;75(3):519–525.
16. Howell WC, Earthman, P., Reid, K., Greaves, J., Delaney, J., Houtkooper, L. Doubly-labeled water validation of the Compendium of Physical Activities in lean and obese college women. *Med Sci Sports Exerc.* 1999;31(S):142.
17. Troiano RP, Berrigan D, Dodd KW, Masse LC, Tilert T, McDowell M. Physical activity in the United States

measured by accelerometer. *Med Sci Sports Exerc*. 2008;40(1):181–188.
18. Byrne NM, Hills AP, Hunter GR, Weinsier RL, Schutz Y. Metabolic equivalent: one size does not fit all. *J Appl Physiol*. 2005;99(3):1112–1119.
19. Harris JA, Benedict FG. A biometric study of human basal metabolism. *Proc Natl Acad Sci USA*. 1918;4(12):370–373.
20. Ross RM, Jackson, A.S. *Exercise Concepts, Calculations, and Computer Applications*: Carmel, IN: Brown and Benchmark; 1990.
21. Compher C, Frankenfield D, Keim N, Roth-Yousey L. Best practice methods to apply to measurement of resting metabolic rate in adults: a systematic review. *J Am Diet Assoc*. 2006;106(6):881–903.
22. Storer TWKK, Dabari B, Tran Z, Rozneck R. Validation of the MedGem device for measurement of resting metabolic rate. [abstract] *Med Sci Sports Exerc*. 2004;36(5, Suppl):S247.
23. Nieman DC, Trone GA, Austin MD. A new handheld device for measuring resting metabolic rate and oxygen consumption. *J Am Diet Assoc*. 2003;103(5):588–592.
24. Murphy OF, Kearny JT. Validation of the MedGem device for measuring resting metabolic rate. [abstract] *Med Sci Sports Exerc*. 2004;36(5, Suppl):S247.
25. McDoniel SO. Systematic review on use of a handheld indirect calorimeter to assess energy needs in adults and children. *Int J Sport Nutr Exerc Metab*. 2007;17(5):491–500.
26. Perret C, Mueller G. Validation of a new portable ergospirometric device (Oxycon Mobile) during exercise. *Int J Sports Med*. 2006;27(5):363–367.
27. HeltheTech. *MedGem User Manual*. CO: Golden; 2003.
28. Bacharach DW, Seifert JG, Harmon J, Burckhard J, Hendrickson N. *Validity of the Oxycon Mobile Open-Circuit Spirometry System*. St. Cloud, MN: Human Performance Lab, St Cloud State University; 2009.
29. Atkinson G, Davison RC, Nevill AM. Performance characteristics of gas analysis systems: what we know and what we need to know. *Int J Sports Med*. 2005;26(Suppl 1):S2–S10.
30. Matthews CE, Heil DP, Freedson PS, Pastides H. Classification of cardiorespiratory fitness without exercise testing. *Med Sci Sports Exerc*. 1999;31(3):486–493.
31. Perri MG, Anton SD, Durning PE, et al. Adherence to exercise prescriptions: effects of prescribing moderate versus higher levels of intensity and frequency. *Health Psychol*. 2002;21(5):452–458.
32. Irwin ML, Ainsworth BE, Conway JM. Estimation of energy expenditure from physical activity measures: determinants of accuracy. *Obes Res*. 2001;9(9):517–525.
33. Racette SB, Schoeller DA, Kushner RF. Comparison of heart rate and physical activity recall with doubly labeled water in obese women. *Med Sci Sports Exerc*. 1995;27(1):126–133.
34. Sallis JF, Haskell WL, Wood PD, et al. Physical activity assessment methodology in the Five-City Project. *Am J Epidemiol*. 1985;121(1):91–106.
35. Durante R, Ainsworth BE. The recall of physical activity: using a cognitive model of the question-answering process. *Med Sci Sports Exerc*. 1996;28(10):1282–1291.
36. Adams SA, Matthews CE, Ebbeling CB, et al. The effect of social desirability and social approval on self-reports of physical activity. *Am J Epidemiol*. 2005;161(4):389–398.
37. Matthews CE, Chen KY, Freedson PS, et al. Amount of time spent in sedentary behaviors in the United States, 2003–2004. *Am J Epidemiol*. 2008;167(7):875–881.