



Published in final edited form as:

*J Sci Med Sport*. 2016 January ; 19(1): 35–40. doi:10.1016/j.jsams.2014.10.002.

## Comparisons of prediction equations for estimating energy expenditure in youth

Youngwon YK Kim<sup>a</sup>, Scott E. SEC Crouter<sup>b</sup>, Jung-Min JML Lee<sup>c</sup>, Phillip M. PMD Dixon<sup>d</sup>, Glenn A. GAG Gaesser<sup>e</sup>, and Gregory J. GJW Welk<sup>a</sup>

<sup>a</sup>235 Forker Building, Ames, Iowa, IA, 50011, USA; Department of Kinesiology, Iowa State University

<sup>b</sup>1914 Andy Holt Avenue, Knoxville, TN, 37996, USA; Department of Kinesiology, Recreation, and Sport Studies, University of Tennessee

<sup>c</sup>6001 Dodge Street, Omaha, NE, 68182, USA; School of Health, Physical Education and Recreation, University of Nebraska – Omaha

<sup>d</sup>Snedecor Hall, Ames, Iowa, IA, 50011, USA; Department of Statistics, Iowa State University

<sup>e</sup>500 N 3<sup>rd</sup> Street, Phoenix, AZ, 85004, USA; School of Nutrition & Health Promotion, Arizona State University

### Abstract

**Objectives**—The purpose of this study was to compare the validity of Actigraph 2-regression models (2RM) and 1-regression models (1RM) for estimation of EE in children.

**Design**—The study used a cross-sectional design with criterion estimates from a metabolic cart

**Methods**—A total of 59 children (7–13yrs) performed 12 activities (randomly selected from a set of 24 activities) for 5min each, while being concurrently measured with an Actigraph GT3X and indirect calorimetry. MET<sub>RMR</sub> (MET considering one's resting metabolic rate) for the GT3X was estimated applying 2RM with vector magnitude (VM2RM) and vertical axis (VA2RM), and four standard 1RMs. The validity of the 2RMs and 1RMs was evaluated using 95% equivalence testing and mean absolute percent error (MAPE).

**Results**—For the group-level comparison, equivalence testing revealed that the 90% confidence intervals for all 2RMs and 1RMs were outside of the equivalence zone (range: 3.63, 4.43) for indirect calorimetry. When comparing the individual activities, VM2RM produced smaller MAPEs (range: 14.5 to 45.3%) than VA2RM (range, 15.5 to 58.1%) and 1RMs (range, 14.5 to 75.1%) for most of the light and moderate activities.

---

Corresponding Author: Name: Youngwon Kim, Mailing Address: Department of Kinesiology, Iowa State University, 235 Forker Building, Ames, IA, 50011, Telephone: 1 - 515 - 294 - 8009, Fax: 1 - 515 - 294 - 8740, myungshingo@gmail.com.

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

**Conclusions**—None of the 2RMs and 1RMs were equivalent to indirect calorimetry. The 2RMs showed smaller individual-level errors than the 1RMs.

### Keywords

Public health; Physical fitness; Accelerometer; Children; Calibration; Validation Studies

---

## Introduction

The Actigraph (Actigraph, Pensacola, FL) is the most widely used activity monitoring device but there is still debate about the most appropriate way to reduce and process the data.<sup>1</sup> A particularly challenging issue is to convert accelerometer counts into more useable units (a process typically referred to as calibration). Calibration equations are typically developed using regression analyses that link counts with corresponding energy expenditure (EE) values (metabolic equivalents: METs) measured from a criterion measure.<sup>2</sup> Four independent calibration studies<sup>3-6</sup> have proposed equations to estimate children's EE but cross-validation studies<sup>7, 8</sup> have not supported their accuracy for assessing free living physical activity behavior.

The inherent limitation of these standard single regression models (referred to as 1RM hereafter) is that they are specific for the activities used in the calibration.<sup>9</sup> Previous research by Crouter et al.<sup>10, 11</sup> demonstrated the utility of a two-regression model (2RM) that allows for the use of separate regression equations for locomotive (i.e. running, walking) and non-locomotive activities. This enables separate predictions for these two distinct patterns. In subsequent independent validation studies, this method was shown to provide comparable EE values relative to doubly-labeled water<sup>12</sup> and indirect calorimetry<sup>13</sup> in adults. A set of 2RMs was developed to estimate EE in youth using the latest, tri-axial Actigraph GT3X model,<sup>14</sup> but a subsequent cross-validation study<sup>7</sup> yielded equivocal findings. This study<sup>7</sup> reported significant underestimation of EE for the seven structured activities (performed for 5 min each) and considerable error across the more extended, free-living measurement session. The discrepant findings are difficult to reconcile considering the promising findings in adults<sup>13</sup> and the favorable preliminary validation results.<sup>14</sup> It is widely recommended that validity of accelerometry-based monitors should be examined using independent samples<sup>15, 16</sup> and under various settings.<sup>2</sup> The present study systematically evaluates the relative validity of two 2RM models compared to four different 1RM equations in an independent sample of youth.

## Method

A total of 59 children between the ages of 7 and 13 volunteered to participate in the study. All of the participants and their parents signed an assent form and written informed consent, respectively, prior to participation. In Table 1, the demographic profiles of the participants were summarized. This study was approved by the Institutional Review Board at our institution.

Each participant performed 12 activities randomly selected from a set of 24 activities designed to mimic free-living activities in children. The 24 activities were classified into

four distinct intensity categories according to MET-derived criteria based on criterion MET values from the OM. The criteria were  $\text{MET}_{\text{RMR}} \geq 2.0$  for sedentary,  $2.0 < \text{MET}_{\text{RMR}} < 3.9$  for light,  $4.0 \leq \text{MET}_{\text{RMR}} < 5.9$  for moderate, and  $\text{MET}_{\text{RMR}} \geq 6.0$  for vigorous intensity. We used adjusted  $\text{MET}_{\text{RMR}}$  values of 2.0 and 4.0 for the sedentary and moderate intensity, respectively, instead of traditional  $\text{MET}_{\text{RMR}}$  values of 1.5 and 3.0. The higher threshold of 4 METS for MVPA<sup>8</sup> is well established based on the established evidence for higher RMR values in youth ( $\sim 5.0 \text{ ml}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$ ) for children compared with adults ( $\sim 3.5 \text{ ml}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$ ).<sup>17</sup> The lower threshold of 2 METS to define Sedentary behavior preserves the same 2:1 ratio between MVPA and Sedentary in adults (i.e. 3.0 METS: 1.5 METS). Recent findings from our lab (manuscript in review) have supported the high classification accuracy of the 2.0 and 4.0  $\text{MET}_{\text{RMR}}$  for classifying sedentary and moderate intensity in children.

The activities were completed in random order and in a semi-structured manner to capture natural variation in physical activity. The characteristics of the included activities by intensity along with corresponding means and standard deviations are presented (see Supplementary Table 1). Each activity was performed for 5 min, with a 1-min resting period between them. Participants wore an Actigraph GT3X for measurement of physical activity and were fitted with a portable metabolic analyzer (i.e. Oxycon Mobile (OM); CareFusion Corp, San Diego, CA) for measurement of pulmonary ventilation and gas exchange (for determination of oxygen uptake). The funding organization (i.e. National Institutes of Health) did not play any role in collecting data (including the analyses and interpretation) and in approving or disapproving the publication of this study.

The Actigraph GT3X is a compact ( $3.8 \times 3.7 \times 1.8 \text{ cm}$ ) and lightweight (27g) tri-axial accelerometer that records acceleration ranging from 0.05g to 2.00g. The Actigraph GT3X was initialized at 30Hz and placed on the right hip at the level of iliac crest. Data were downloaded at 1s epoch and then reintegrated into 10s epochs and 60s for analyses. The ActiLife software (firmware version 6.5.1) was used for data management.

The OM, the criterion measure for determination of EE, is a portable gas analyzer that measures breath-by-breath respiratory gas exchange under free-living conditions. Volume and gas calibration were performed before each trial following the manufacturer's recommendations. The OM has been shown to be valid and reliable.<sup>18, 19</sup>

The Actigraph data were processed using 2 different 2RM approaches,<sup>14</sup> one developed with vertical axis data (VA2RM) and one developed using vector magnitude data (VM2RM). The Actigraph data were also processed for the 1RM equations developed by Freedson/Trost et al. (FT1RM),<sup>3</sup> Trost et al. (TR1RM),<sup>6</sup> Puyau et al. (PU1RM),<sup>4</sup> and Treuth et al. (TH1RM)<sup>5</sup> to provide comparisons with previous results (see Supplementary Table 2). To ensure effective comparisons, all data were presented using equivalent outcome units of  $\text{MET}_{\text{RMR}}$ .  $\text{MET}_{\text{RMR}}$  is a MET value that incorporates one's resting metabolic rate (RMR). In this study, the Schofield equation<sup>20</sup> was used to predict RMR of the children, which is the same methodology used in previous validation studies.<sup>21, 22, 8, 23</sup> The specific  $\text{MET}_{\text{RMR}}$  value was calculated by dividing the measured activity  $\text{VO}_2$  value (i.e.  $\text{ml}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$ ) by the predicted (Schofield<sup>20</sup>) RMR value (i.e.  $\text{ml}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$ ).

The 2RM method by Crouter et al.<sup>14</sup> predicts EE in  $MET_{RMR}$ . However, FT1RM and TH1RM yield  $MET_{3,5}$ , which is a MET value based on a resting metabolic rate of  $3.5\text{ml}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$ . Therefore,  $MET_{RMR}$  values for the FT1RM and TH1RM were obtained by first multiplying MET values predicted from the equations by  $3.5\text{ml}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$ , then dividing by a RMR estimated using the Schofield equation.<sup>20</sup> The TR1RM and PU1RM provide estimates of EE in  $\text{Kcal}\cdot\text{min}^{-1}$  and  $\text{Kcal}\cdot\text{kg}\cdot\text{min}^{-1}$ , respectively. Therefore, different conversions were needed to compute  $MET_{RMR}$ . An established conversion factor ( $1\text{ L O}_2 = 4.825\text{ kcal}$ ) was used to convert EE estimates into estimates of oxygen consumption in  $\text{ml}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$ . This was then divided by the weight (kg) and the estimated RMR value for TR1RM, and by the estimated RMR value for PU1RM to compute  $MET_{RMR}$ .<sup>23</sup> For the OM data,  $MET_{RMR}$  values were calculated by dividing  $VO_2$  values ( $\text{ml}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$ ) of performed activities by the estimated RMR value.<sup>20</sup>

A statistical “equivalence test”<sup>24, 25</sup> was used to examine measurement agreement between the Actigraph equations (i.e. 2RMs and 1RMs) and the OM. In traditional analytic approaches, the null hypothesis is that the two methods are equal. Failure to reject the null suggests “no evidence of a difference”; however, this does not necessarily imply that there is “evidence of equivalence”.<sup>26</sup> With an equivalence test, the null hypothesis is flipped to specify a difference between two means, thereby allowing a direct test of equivalence. In the context of the present study, a 95% equivalence test (i.e.  $\alpha=5\%$ ) would conclude that Actigraph EE models (i.e. 2RMs and 1RMs) and the OM are considered to be significantly equivalent to each other if a 90% of confidence interval (CI) for a mean of the Actigraph EE model fell into a proposed equivalence zone (i.e.  $\pm 10\%$  of the mean) of the OM.

To evaluate overall agreement between the Actigraph EE models and OM across all the activities, the 90% CIs for means of each Actigraph EE model were estimated from a mixed model ANOVA in which participants are included as a random effect to account for potential correlation of measurements from the same participant. Whether or not the Actigraph measurement for each of the 24 activities was significantly equivalent to the corresponding OM measurement was tested using an equivalence test with an alpha-level of 5%. The null hypothesis of non-equivalence was rejected when the 90% confidence interval for the mean Actigraph EE measurement fell inside the equivalence zone, which we defined as  $\pm 10\%$  of the mean for OM. The equivalence testing and Pearson correlation coefficient were used to examine group-level agreement.

To examine individual-level measurement errors, mean absolute percent errors (MAPE) and Bland Altman plots were used. MAPE was calculated by dividing an absolute difference between a regression model and the OM by the OM, multiplied by 100%. Bland-Altman plots with corresponding 95% limits of agreement and fitted lines (from regression analyses between mean and difference) with their corresponding parameters (i.e. intercept, slope) were presented. Perfect agreement between the Actigraph and OM would be identified with a fitted line that provides a slope of 0 and an intercept of 0.

## Results

Overall, none of the regression models (i.e. VM2RM, VA2RM, FT1RM, TR1RM, PU1RM, TH1RM) were significantly equivalent to the OM for the group-level comparisons as none of the 90% CIs of the 6 regression models were completely included within the equivalence zone of the OM (lower bound: 3.63, upper bound: 4.43). (See Table 2) Correlations for the VM2RM ( $r=0.53$ ) and VA2RM ( $r=0.58$ ) with the OM were all significant ( $P<0.0001$ ), and higher than for the FT1RM ( $r=0.37$ ), TR1RM ( $r=0.30$ ), PU1RM ( $r=0.37$ ) and TH1RM ( $r=0.35$ ). MAPEs for the VM2RM, VA2RM, FT1RM, TR1RM, PU1RM and TH1RM were 28.2%, 34.2%, 31.5%, 40.2%, 61.7% and 36.6%, respectively.

Regarding the agreement between the 2RMs and the OM, both VM2RM and VA2RM yielded non-equivalent underestimations for all 48 possible activity-specific comparisons (i.e., 2 methods $\times$ 24 activities) (See Figure 1.(a)). Of the 96 possible comparisons (i.e. 4 methods $\times$ 24 activities) between the 1RMs and OM, none of them were significantly equivalent with one another, except for the two comparisons between the FT1RM (i.e. 90%CI: 4.22, 4.84) and the OM (i.e. equivalence zone: 3.97, 4.85) for brisk walking, and between the PU1RM (i.e. 90%CI: 4.91, 5.64) and the OM (i.e. equivalence zone: 4.88, 5.96) for dribbling a basketball.

The MAPE values for the 2RMs ranged from 32.5% to 39.4% for sedentary activities, from 14.5% to 42.9% for light intensity, from 15.0% to 58.1% for moderate intensity, and from 19.6% to 61.5% for vigorous intensity activities (See Supplementary Table 3). The MAPE values for the four 1RMs ranged from 17.3% to 67.2% for sedentary activities, from 14.5% to 70.3% for light intensity, from 17.3% to 75.1% for moderate intensity, and from 16.3% to 78.0% for vigorous intensity activities. The 2RMs, in particular VM2RM, produced relatively smaller (and/or comparable) MAPE values in comparison with (all or most of) the 1RMs for most of the 15 LPA (light physical activity) and MPA (moderate physical activity), the exception being “Hand weight exercises” compared with TH1RM.

The Bland-Altman plots showed that VM2RM (range of 95% limits of agreement =  $1.68\text{MET}_{\text{RMR}}$ ) (i.e., -1.99 to  $-0.31\text{MET}_{\text{RMR}}$ ) and VA2RM (range:  $1.50\text{MET}_{\text{RMR}}$ ) showed relatively narrower ranges of 95% limits of agreement in comparison with FT1RM (range:  $2.11\text{MET}_{\text{RMR}}$ ), TR1RM (range:  $3.05\text{MET}_{\text{RMR}}$ ), PU1RM (range:  $1.77\text{MET}_{\text{RMR}}$ ), and TH1RM (range:  $2.12\text{MET}_{\text{RMR}}$ ). (See Supplementary Figure 1 and Supplementary Table 4) The slopes of the fitted line were not significant for VM2RM ( $P=0.286$ ), FT1RM ( $P=0.655$ ) and TH1RM ( $P=0.716$ ). This implies that patterns of systematic bias were not significant in the VM2RM, FT1RM and TH1RM. However, significant patterns of systematic errors were found for VA2RM ( $P=0.001$ ), TR1RM ( $P<0.001$ ) and PU1RM ( $P=0.002$ ).

## Discussion

The present study examined the validity of two recently developed 2RMs and four standard 1RMs for physical activity EE estimation relative to the indirect calorimetry criterion measure in children. Overall, none of the six regression models were significantly equivalent to the criterion measure. All regression models provided considerably large underestimation

of EE values measured by indirect calorimetry. Given that the 2RMs produced narrower ranges of 95% limits of agreement, less biased patterns of Bland-Altman plots and smaller MAPEs, the use of the 2RMs may result in less individual variation in measuring EE in comparison with the 1RMs.

We did not detect any particular pattern in the errors that may help to determine the underlying cause of the discrepancies. However, comparison of our results with prior related studies offers some insights into sources of error. To be specific, the findings on the validity of the 2RMs in children contrast with the original cross validation study<sup>14</sup> but are consistent with the more recent independent validation study.<sup>7</sup> In the original cross validation study of the 2RMs,<sup>14</sup> the VM2RM and VA2RM yielded  $MET_{RMR}$  values comparable with indirect calorimetry (i.e. mean differences  $< 0.8 MET_{RMR}$ ) for most of the performed activities. However, the favorable findings results may be attributable to the fact that the number, types, and duration of the activities performed were equivalent to those used to develop the VM2RM and VA2RM models. It should not be surprising to find good validity when a cross-validation study uses the exactly same general methodology that is used in a calibration study. The previous study<sup>14</sup> also used relatively crude statistical methods (i.e. one-way ANOVA) to test for agreement, and had a relatively small sample size for the comparison group ( $n=39$ ). In the current study, the 2RM method was evaluated in an independent, larger sample of children ( $n=59$ ) and employed a more rigorous methodology (i.e. equivalence testing) than previous studies. A subsequent independent validation study by Crouter et al.<sup>7</sup> evaluated the VM2RM and VA2RM using a more variable set of activities (i.e. 8 structured activities for 8 minutes each + 1 free-living measurement session for approximately 2 hours) and a larger number of participants (i.e. 32 girls and 40 boys ages 11-15yrs) in comparison to their original calibration study.<sup>14</sup> Consistent with the present study, that study<sup>7</sup> demonstrated that the VM2RM and VA2RM underestimated EE measured with indirect calorimetry for the 8 structured activities (range of mean percent errors, 1.0%–23.1%) and free-living activity measurement (range of mean percent error, 26.3% and 27.2%, respectively). The convergent findings from the two independent validation studies (i.e. the current study and the study by Crouter et al.<sup>7</sup>), the 2RM method appears to underestimate EE values in children.

From a public health standpoint, it is important to establish procedures to more accurately estimate EE in youth. Because LPA and MPA are common outcome measures in contemporary physical activity research for youth,<sup>27</sup> accurate calibration of EE equations is critical for accurately estimating EE in these two intensities. In this study, the resulting MAPE values for the 2RMs were comparable and/or relatively smaller in comparison with MAPEs for the 1RMs for the majority of the LPA and MPA. This suggests that the 2RMs may provide more accurate estimates of EE and/or activity time for LPA and MPA in comparison with the 1RMs. A similar result was reported by Crouter et al.<sup>7</sup> who compared the VM2RM, VA2RM, and 5 standard 1RMs (i.e. by Freedson/Trost et al.,<sup>3</sup> Treuth et al.,<sup>5</sup> Trost et al.,<sup>6</sup> Puyau et al.,<sup>4</sup> and Evenson et al.<sup>28</sup>) against indirect calorimetry for estimating minutes spent during 4 categories of physical activity (i.e. sedentary, LPA, MPA, VPA). In the study by Crouter et al.,<sup>7</sup> the VM2RM and VA2RM produced smaller ranges of 95% limits of agreement in comparison with the 5 1RMs in classifying LPA and MPA. This indicates that in measuring EE on LPA and MPA, the 2 RMs would have relatively smaller

individual bias in comparison with the 1RMs. Another study conducted by Trost et al.<sup>8</sup> evaluated activity classification accuracy of five 1RMs (Freedson/Trost et al.,<sup>3</sup> Puyau et al.,<sup>4</sup> Treuth et al.,<sup>5</sup> Mattocks et al.,<sup>29</sup> Evenson et al.<sup>28</sup>) using Receiver Operating Characteristic-area under the curve (ROC-AUC). In the Trost et al. study,<sup>8</sup> all the five 1RMs demonstrated relatively low classification accuracy for LPA (ROC-AUC ranges from 0.43 to 0.70) and MPA (ROC-AUC ranges from 0.56 to 0.79) time in comparison to classification accuracy for sedentary (ROC-AUC ranges from 0.80 to 0.90) and VPA (ROC-AUC ranges from 0.54 to 0.84). This suggests that LPA and MPA are more difficult to evaluate using 1RM approaches. The present results do not directly support the 2RM approach but it is clear that it provides advantages over the traditional 1RM approach.

There are some key strengths of the present study. To our knowledge, this study is the first study to employ the novel method of equivalence testing in validation studies on physical activity measurement tools in children. Given that equivalence testing is intended to determine how closely one method equates to another method, it is a more appropriate analytic approach for this type of validation research. Another strength is that we included a wide variety of activities (i.e. a total of 24) designed to mimic the range of children's free-living activities.

However, the results of the current study still need to be interpreted with some caution. The current study used slightly different methodologies than that employed in the original 2RM development study.<sup>14</sup> Most importantly, the activities of this study were performed in a random fashion whereas in the validation study,<sup>14</sup> the 2RMs were developed and validated with activities being performed from the lowest to the highest intensity category. We view the random order of activities as an important design feature since it more directly captures free-living behavior. However, it is also possible that this fully randomized order could lead to overestimation of the “criterion” EE values for lower intensity activities if performed after more vigorous activities. Moreover, children's RMR was not measured, but estimated using the Schofield's equation.<sup>20</sup> However, the Schofield's equation accurately estimated RMR values for pre-school children,<sup>30</sup> and has been widely used in previous validation studies<sup>21, 22, 8, 23</sup> to evaluate different sets of Actigraph EE equations for youth.

## Conclusion

None of the two 2RMs and four 1RMs were significantly equivalent to the indirect calorimetry for both overall group comparisons and activity-specific comparisons (only except for 1 activity with the TH1RM). Relatively smaller individual variations were identified for the 2RMs in comparison with the 1RMs. In measuring estimates of EE for LPA and MPA, the 2RMs showed relatively smaller individual errors in comparison with the 1RMs. Future studies using the Actigraph to measure EE should be informed of the smaller potential biases inherent in the use of the 2RMs than the 1RMs.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

We would like to thank all the participants and their parents for their participation in the current study. We also would like to thank Yang Bai and Cassandra Weickert for their remarkable help with the data collection. All the authors declare no conflicts of interest. The present study was supported by the National Institute of Health grant (R01 HL091006).

## References

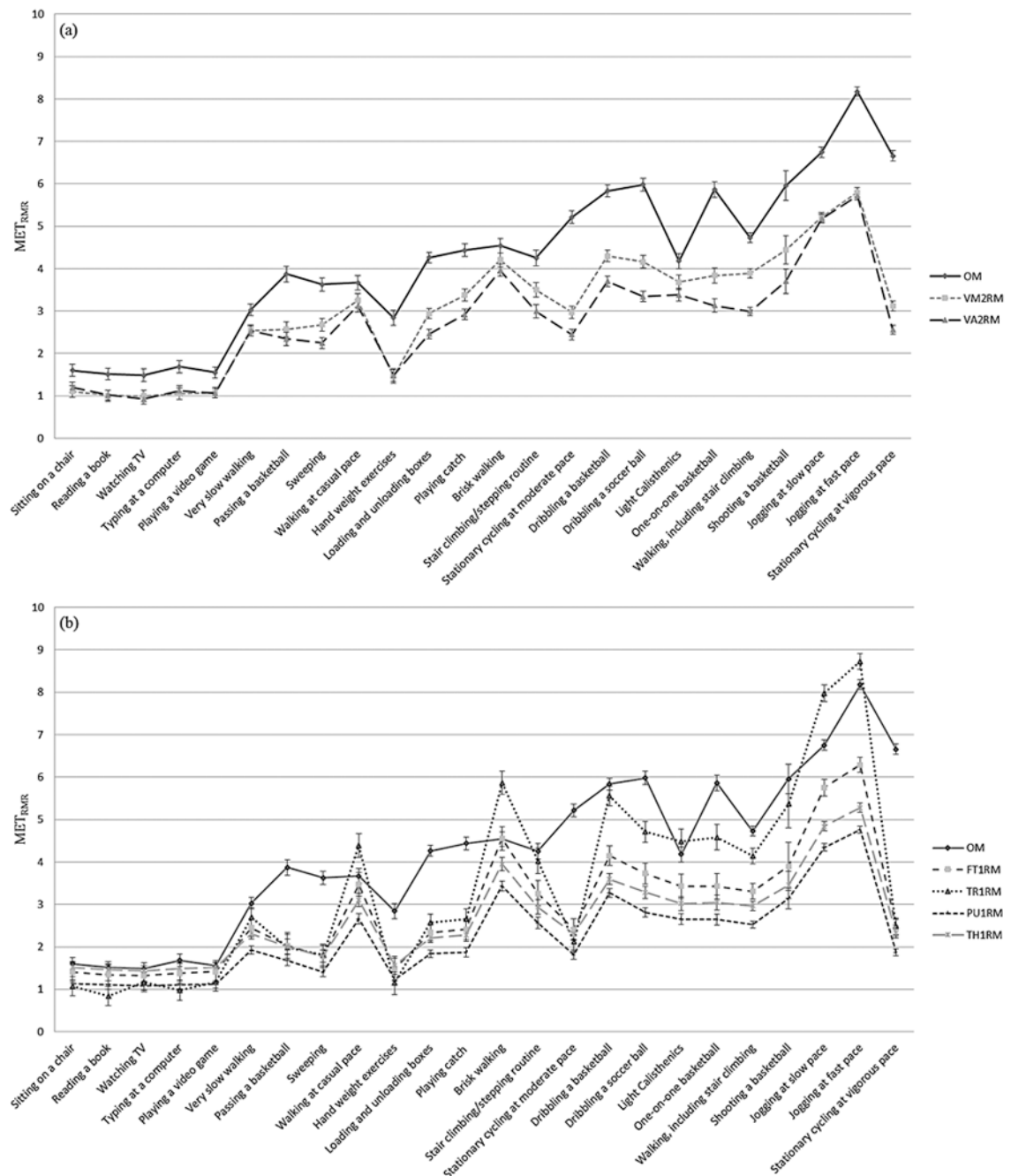
1. Matthews CE, Hagstromer M, Pober DM, et al. Best Practices for Using Physical Activity Monitors in Population-Based Research. *Med Sci Sport Exer.* 2012; 44:S68–S76.
2. Welk GJ. Principles of design and analyses for the calibration of accelerometry-based activity monitors. *Med Sci Sports Exerc.* 2005; 37(11 Suppl):S501–511. [PubMed: 16294113]
3. Freedson P, Pober D, Janz KF. Calibration of accelerometer output for children. *Med Sci Sports Exerc.* 2005; 37(11 Suppl):S523–530. [PubMed: 16294115]
4. Puyau MR, Adolph AL, Vohra FA, et al. Validation and calibration of physical activity monitors in children. *Obes Res.* 2002 Mar; 10(3):150–157. [PubMed: 11886937]
5. Treuth MS, Schmitz K, Catellier DJ, et al. Defining accelerometer thresholds for activity intensities in adolescent girls. *Med Sci Sports Exerc.* 2004 Jul; 36(7):1259–1266. [PubMed: 15235335]
6. Trost SG, Ward DS, Moorehead SM, et al. Validity of the computer science and applications (CSA) activity monitor in children. *Med Sci Sports Exerc.* 1998 Apr; 30(4):629–633. [PubMed: 9565947]
7. Crouter SE, Horton M, Bassett DR Jr. Validity of ActiGraph child-specific equations during various physical activities. *Med Sci Sports Exerc.* 2013 Jul; 45(7):1403–1409. [PubMed: 23439413]
8. Trost SG, Loprinzi PD, Moore R, et al. Comparison of accelerometer cut points for predicting activity intensity in youth. *Med Sci Sports Exerc.* 2011 Jul; 43(7):1360–1368. [PubMed: 21131873]
9. Kim Y, Beets MW, Welk GJ. Everything you wanted to know about selecting the “right” Actigraph accelerometer cut-points for youth, but...: a systematic review. *J Sci Med Sport.* 2012 Jul; 15(4): 311–321. [PubMed: 22306372]
10. Crouter SE, Clowers KG, Bassett DR Jr. A novel method for using accelerometer data to predict energy expenditure. *J Appl Physiol.* 2006 Apr; 100(4):1324–1331. [PubMed: 16322367]
11. Crouter SE, Kuffel E, Haas JD, et al. Refined two-regression model for the ActiGraph accelerometer. *Med Sci Sports Exerc.* 2010 May; 42(5):1029–1037. [PubMed: 20400882]
12. Rothney MP, Brychta RJ, Meade NN, et al. Validation of the ActiGraph two-regression model for predicting energy expenditure. *Med Sci Sports Exerc.* 2010 Sep; 42(9):1785–1792. [PubMed: 20142778]
13. Kuffel EE, Crouter SE, Haas JD, et al. Validity of estimating minute-by-minute energy expenditure of continuous walking bouts by accelerometry. *Int J Behav Nutr Phys Act.* 2011; 8:92. [PubMed: 21864359]
14. Crouter SE, Horton M, Bassett DR Jr. Use of a two-regression model of estimating energy expenditure in children. *Med Sci Sports Exerc.* 2012; 44(6):1177–1185. [PubMed: 22143114]
15. Bassett DR, Rowlands A, Trost SG. Calibration and Validation of Wearable Monitors. *Med Sci Sport Exer.* 2012; 44:S32–S38.
16. Strath SJ, Pfeiffer KA, Whitt-Glover MC. Accelerometer Use with Children, Older Adults, and Adults with Functional Limitations. *Med Sci Sport Exer.* 2012; 44:S77–S85.
17. Harrell JS, McMurray RG, Baggett CD, et al. Energy costs of physical activities in children and adolescents. *Med Sci Sports Exerc.* 2005 Feb; 37(2):329–336. [PubMed: 15692331]
18. Hodges LD, Brodie DA, Bromley PD. Validity and reliability of selected commercially available metabolic analyzer systems. *Scand J Med Sci Sports.* 2005 Oct; 15(5):271–279. [PubMed: 16181250]
19. Rosdahl H, Gullstrand L, Salier-Eriksson J, et al. Evaluation of the Oxycon Mobile metabolic system against the Douglas bag method. *Eur J Appl Physiol.* 2010 May; 109(2):159–171. [PubMed: 20043228]
20. Schofield WN. Predicting basal metabolic rate, new standards and review of previous work. *Hum Nutr Clin Nutr.* 1985; 39(Suppl 1):5–41. [PubMed: 4044297]



21. Janssen X, Cliff DP, Reilly JJ, et al. Predictive Validity and Classification Accuracy of ActiGraph Energy Expenditure Equations and Cut-Points in Young Children. *Plos One*. 2013 Nov 11.8(11)
22. Reilly JJ, Kelly LA, Montgomery C, et al. Validation of Actigraph accelerometer estimates of total energy expenditure in young children. *Int J Pediatr Obes*. 2006; 1(3):161–167. [PubMed: 17899634]
23. Trost SG, Way R, Okely AD. Predictive validity of three ActiGraph energy expenditure equations for children. *Med Sci Sport Exer*. 2006 Feb; 38(2):380–387.
24. Dixon PM, Pechmann JHK. A statistical test to show negligible trend. *Ecology*. 2005 Jul; 86(7): 1751–1756.
25. Wellek, S. Testing statistical hypotheses of equivalence. Boca Raton, Fla: Chapman & Hall/CRC; 2003.
26. Hauck WW, Anderson S. A new statistical procedure for testing equivalence in two-group comparative bioavailability trials. *J Pharmacokinet Biopharm*. 1984 Feb; 12(1):83–91. [PubMed: 6747820]
27. Janssen I, Leblanc AG. Systematic review of the health benefits of physical activity and fitness in school-aged children and youth. *Int J Behav Nutr Phys Act*. 2010; 7:40. [PubMed: 20459784]
28. Evenson KR, Catellier DJ, Gill K, et al. Calibration of two objective measures of physical activity for children. *J Sports Sci*. 2008 Dec; 26(14):1557–1565. [PubMed: 18949660]
29. Mattocks C, Leary S, Ness A, et al. Calibration of an accelerometer during free-living activities in children. *Int J Pediatr Obes*. 2007; 2(4):218–226. [PubMed: 17852552]
30. Firouzbakhsh S, Mathis RK, Dorchester WL, et al. Measured Resting Energy-Expenditure in Children. *J Pediatr Gastroenterol Nutr*. 1993 Feb; 16(2):136–142. [PubMed: 8450379]

### Practical Implications

- Researchers would be informed of the relative validity of different types of prediction equations for Actigraph accelerometers to estimate energy expenditure in youth.
- Epidemiological surveillance and/or intervention studies would utilize prediction equations (i.e. two-regression models) in order to obtain accurate estimates of energy expenditure levels for youth.
- Relatively more accurate estimation of energy expenditure for light and moderate activities could be achieved with the use of the two regression models.



**Figure 1.** Means and standard errors of (a) two 2-regression models (VM2RM - Vector magnitude 2 regression model; VA2RM - Vertical axis 2 regression model) and (b) four 1-regression models (FT1RM - Freedson/Trost 1 regression model; TR1RM - Trost 1 regression model; PU1RM - Puyau 1 regression model; TH1RM - Treuth 1 regression model) and Oxycon Mobile for each activity type; Note - MET<sub>RMR</sub> (metabolic equivalents taking into account resting metabolic rate).

**Table 1**  
**Physical characteristics of the participants (n=59)**

Variables	Values
Age (M ± SD, range)	9.9 ± 1.8 yrs, 7-13 yrs
Gender	
Female (n=18)	30.5%
Male (n=41)	69.5%
Height (M ± SD)	146.0 ± 12.9 cm
Weight (M ± SD)	38.9 ± 13.8 kg
Body Mass Index (BMI) <sup>a</sup>	
Normal weight (n=51)	79.7%
Overweight (n=6)	13.6%
Obese (n=3)	6.8%
Race	
Caucasian (n=56)	93.2%
Asian (n=2)	3.4%
African American (n=1)	1.7%
Hispanic (n=1)	1.7%
Resting Metabolic Rate	
All (n=59)	4.89 ml·kg <sup>-1</sup> ·min <sup>-1</sup>
Female (n=18)	4.75 ml·kg <sup>-1</sup> ·min <sup>-1</sup>
Male (n=41)	4.96 ml·kg <sup>-1</sup> ·min <sup>-1</sup>

<sup>a</sup>BMI; 5<sup>th</sup> – 85<sup>th</sup> percentile as normal Weight, 85<sup>th</sup> – 95<sup>th</sup> percentile as overweight, 95<sup>th</sup> as obese

**Table 2**

Overall measurement agreement between Actigraph regression models and Oxycon Mobile in energy expenditure measurement.

Type	Method	90% CI of Regression Models			Equivalence zone of OM			Correlations			MAPE (%)
		M	SE	Lower bound	Upper bound	Lower bound	Upper bound	r	P-value	M	
<i>Criterion measure</i>											
OM <sup>a</sup>		4.03	0.06	-	-	3.63	4.43	-	-	-	-
<i>Regression Models</i>											
VM2RM <sup>b</sup>		2.88	0.05	2.79	2.97	-	-	0.53	<0.0001	28.2	9.3
VA2RM <sup>c</sup>		2.64	0.04	2.57	2.71	-	-	0.58	<0.0001	34.2	7.2
FTIRM <sup>d</sup>		2.76	0.06	2.65	2.86	-	-	0.37	0.0040	31.5	11.0
TRIRM <sup>e</sup>		2.46	0.10	2.29	2.62	-	-	0.30	0.0221	40.2	15.5
PUIRM <sup>f</sup>		1.54	0.04	1.47	1.60	-	-	0.37	0.0036	61.7	7.2
THIRM <sup>g</sup>		2.55	0.06	2.44	2.65	-	-	0.35	0.0068	36.6	11.3

Note: Values are METRMR (metabolic equivalents taking into account resting metabolic rate) unless otherwise noted.

<sup>a</sup>Oxycon Mobile;

<sup>b</sup>Vector magnitude 2 regression model;

<sup>c</sup>Vertical axis 2 regression model;

<sup>d</sup>Freedson/Trost 1 regression model;

<sup>e</sup>Trost 1 regression model;

<sup>f</sup>Puyau 1 regression model;

<sup>g</sup>Treuth 1 regression model