Are Current Body Mass Index Referenced Pedometer Step-Count Recommendations Applicable to US Youth?

Michael W. Beets, Guy C. Le Masurier, Aaron Beighle, David A. Rowe, Charles F. Morgan, Jack Rutherford, Michael Wright, Paul Darst, and Robert Pangrazi

Background: The purpose of this study was to cross-validate international BMI-referenced steps/d cut points for US girls (12,000 steps/d) and boys (15,000 steps/d) 6 to 12 years of age. Methods: Secondary pedometer-determined physical activity data from US children (N = 1067; 633 girls and 434 boys, 6 to 12 years) were analyzed. Using international BMI classifications, cross-validation of the 12,000 and 15,000 steps/d cut points was examined by the classification precision, sensitivity, and specificity for each age–sex stratum. Results: For girls (boys) 6 to 12 years, the 12,000 (15,000) steps/d cut points correctly classified 42% to 60% (38% to 67%) as meeting (achieved steps/d cut point and healthy weight) and failing (did not achieve steps/d cut point and overweight). Sensitivity ranged from 55% to 85% (64% to 100%); specificity ranged from 23% to 62% (19% to 50%). Conclusion: The utility of pedometer steps/d cut points was minimal in this sample given their inability to differentiate among children who failed to achieve the recommended steps/d and exhibited an unhealthy weight. Caution, therefore, should be used in applying previous steps/d cut points to US children.

Keywords: children, physical activity, obesity

The prevalence of unhealthy weight in children and adolescents is a growing international concern. Accumulated evidence identifies health-enhancing physical activity as an important factor in the primary and secondary prevention of
unhealthy weight gain and its associated comorbidities.\textsuperscript{1,2} Specifically, studies indicate youth of unhealthy weight engage in less physical activity (regardless of the methods used to assess activity) than their healthy-weight peers.\textsuperscript{3-7}

One strategy to improve physical activity levels is the development of physical activity cut points that specify minimum levels of daily physical activity for youth that are linked to a key health benefit (eg, healthy weight). The current study focuses on pedometer cut points (steps/d) related to body mass index (BMI).\textsuperscript{3,8} When developing steps/d cut points, an issue of concern is the representativeness of the sample of youth from which they were derived, especially when attempting to apply them to other samples of youth.\textsuperscript{9-11} Previously developed steps/d cut points were composed of samples with only 36\% of the youth from the United States\textsuperscript{3} or consisted of youth residing outside the United States.\textsuperscript{8} The minimal representation of US youth in these 2 samples brings into question whether the steps/d standards are suitable for US children. In previous analyses of the data used by Tudor-Locke et al,\textsuperscript{3} Vincent and colleagues\textsuperscript{12} determined that US youth engaged in fewer steps/d than their age- and sex-matched Australian and Swedish peers. Moreover, neither Australian nor Swedish children were categorized as overweight or obese,\textsuperscript{(p. 1371)} and correspondingly, US youth had significantly higher BMI values than youth of either of the other countries. This brings into question whether these cut points can be applied to US youth. A further consideration is whether the criterion can sufficiently differentiate between groups (healthy BMI versus unhealthy BMI) on the behavior of interest (steps/d).\textsuperscript{9-11}

Given these considerations, this study was undertaken to examine whether the preliminary BMI-referenced pedometer cut points derived by Tudor-Locke and colleagues\textsuperscript{3} were applicable to a large sample of US girls and boys 6 to 12 years of age. Following methodology outlined in previous studies,\textsuperscript{3,8} cross-validation procedures were carried out to examine the classification precision of current steps/d pedometer cut points\textsuperscript{3} using international BMI classifications.\textsuperscript{13}

**Methods**

The analysis conducted and presented herein is a secondary analysis of 3 existing data sets of 1067 children (633 girls and 434 boys, 6 to 12 years) from Arizona (60\%), North Carolina (25\%), and California (15\%). The convenience samples consisted of youth who were predominately White non-Hispanic (81\%), with smaller proportions of youth of Hispanic (14\%) and other ethnic (eg, African American, Asian, Native American, 5\%) decent. The statistical analyses undertaken in this study have not been conducted previously on these data sets.

To be included in the current study, data had to comply with the following criteria: (1) at least 4 days of sealed pedometer data using an acceptably accurate pedometer, (2) objectively measured height and weight, and (3) similar data collection protocols. Data for this study were collected using either the Walk4Life MLS 2525 (approximately 60\% of sample, Walk4Life, Plainfield, IL) or the Yamax SW 200 (approximately 40\% of sample, Yamax Corp., Tokyo, Japan), both of which have demonstrated acceptable accuracy in recent studies.\textsuperscript{14,15} Rationale for these inclusion criteria can be found in Le Masurier et al.\textsuperscript{16} Institutional Review Board approval was obtained, and all participants provided a written
informed consent form signed by a parent or guardian and a signed written assent form. Institutional Review Board approval was also obtained for the analyses presented herein.

**Data Collection Procedures**

Data were collected between October 2002 and May 2005. All data were collected during months when the weather was conducive to outdoor physical activity. Each participant was instructed to wear a sealed or unsealed pedometer during waking hours, excluding water activities (eg, bathing, swimming), for 4 consecutive weekdays. For each data set, pedometers were collected at the same time each day. Once collected, pedometers were opened, data were recorded, pedometers were reset, and pedometers were distributed back to the participants. This process took approximately 30 minutes each day. Each day, students were asked if they took the pedometer off for any reason during the previous day. If the pedometer was worn for less than 1 hour, the data were excluded from the analyses. In addition, students were excluded if their average daily step counts were below 1000. During the week of data collection, the research staff collected each student’s height and weight, without shoes, using objective measures (ie, digital weight scale and stadiometer). For each of the 3 data sets, pedometers and height and weight apparatus were examined for accuracy (eg, shake test, calibration with known weight) before data collection.

**Data Treatment and Statistical Analysis Procedures**

To provide comparative analysis with previously published BMI-referenced steps/d cut points for children, data treatment and analytical procedures were conducted according to those outlined by Tudor-Locke et al. For the analyses, the international BMI classification system was used to group children according to BMI status for age and sex.

**International BMI Steps/d Cut Points Classification Precision, Sensitivity, and Specificity**

Several indices were calculated that describe the accuracy of classifying US children based on meeting the international BMI-referenced steps/d cut point (12,000 and 15,000 steps/d for girls and boys, respectively) or failing to meet the BMI-referenced steps/d cut point (ie, achieves versus fails to achieve). Classification precision (probability of a correct decision) was calculated by summing the probability of true masters (TM, healthy BMI and achieve BMI-referenced steps/d cut point) and true nonmasters (TN, overweight or obese BMI and fail to achieve BMI-referenced steps/d cut point). Sensitivity of the cut points was defined as the ability of the cut points to correctly identify children that failed to meet the BMI-referenced steps/d cut point and had an unhealthy BMI. This was calculated by dividing the number of TN by the sum of TN and false masters (FM, overweight or obese and achieved BMI-referenced steps/d cut point).
Specificity of the cut points was defined as the ability of the recommendation to correctly identify children that met the BMI-referenced steps/d cut point and were classified as having a healthy BMI. This was calculated by taking the number of TM divided by the sum of TM and false nonmasters (FN, healthy BMI and fail to achieve BMI-referenced steps/d cut point). All indices were examined for each age–sex stratum and for the entire sample of girls and boys, separately.

The diagnostic odds ratio (OR) and 95% confidence interval (95% CI) of the cut points, defined as the odds of being classified as a TN for those youth who exhibited an unhealthy weight relative to the odds of being classified as a TN for those youth who exhibited a healthy weight, were calculated for each age and sex stratum.

Results

Descriptive characteristics of the total sample, separated into criterion groups based on international age- and sex-specific BMI percentiles (healthy BMI and at-risk of and overweight) for girls and boys, are presented in Table 1. Results of the classification precision, sensitivity, and specificity for the 12,000 and 15,000 steps/d are presented in Table 2. For girls, the overall probability of the 12,000 steps/d cut point resulting in a correct decision (TM and TN status) was .47 (range = .42 to .60). The sensitivity of the 12,000 steps/d cut point for correctly identifying true nonmasters (girls who failed to achieve the step-count cut point and had an unhealthy BMI) was .77 (range = .55 to .85). Conversely, the specificity of the cut point for identifying true masters (girls who achieved the cut point and were of a healthy BMI) was .34 (range = .23 to .62). For boys, the overall probability of the 15,000 steps/d cut point resulting in a correct decision was .46 (range = .38 to .67). The sensitivity of the 15,000 steps/d cut point for correctly identifying TM was .83 (range = .64 to 1.00), whereas the specificity for identifying TN was .28 (range = .19 to .50). The OR (95% CI) for the 12,000 and 15,000 steps/d cut point for girls and boys in detecting unhealthy-weight youth for each age and sex stratum was nonsignificant for all stratum except for boys age 10 (OR = 4.33; 95% CI, 1.29 to 14.59). Conversely overall, the OR (95% CI) for the 12,000 and 15,000 steps/d cut point for girls and boys in detecting unhealthy-weight youth was 1.80 (95% CI, 1.21 to 2.68) and 1.94 (95% CI, 1.17 to 3.22), respectively. This is interpreted as the following: for the 12,000 steps/d cut point, the odds for being classified as a TN (unhealthy weight and failing to achieve the steps/d cut point) among girls of an unhealthy weight is 1.80 times higher than the odds for being classified as a TN among girls with a healthy weight. To determine the impact of outliers on the results, analyses were conducted that excluded students with less than 4000 steps/d (n = 18, 1.7%) and greater than 30,000 steps/d (n = 1, 0.09%). No changes in the estimates presented earlier were observed.

Discussion

Previous criterion-referenced steps/d cut points were developed using samples that were not representative of the activity and BMI patterns observed in the United States. This is of concern because the precision of criterion-referenced cut
Table 1  Step-Count Descriptive Characteristics for Girls and Boys 6–12 y Based on International Age- and Sex-Specific Percentile Body Mass Index<sup>a</sup> Categories—United States, 2002–2005

<table>
<thead>
<tr>
<th>Sex</th>
<th>Age (y)</th>
<th>n</th>
<th>Mean</th>
<th>SD</th>
<th>CV</th>
<th>Min</th>
<th>Max</th>
<th>n</th>
<th>Mean</th>
<th>SD</th>
<th>CV</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Girls</td>
<td>6</td>
<td>23</td>
<td>11,453.2</td>
<td>±2635.3</td>
<td>23.0%</td>
<td>7124.8</td>
<td>17,218.0</td>
<td>5</td>
<td>10,254.5</td>
<td>±2745.8</td>
<td>26.8%</td>
<td>7447.0</td>
<td>13,029.3</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>27</td>
<td>12,191.2</td>
<td>±3202.0</td>
<td>26.3%</td>
<td>7492.5</td>
<td>18,714.0</td>
<td>17</td>
<td>10,393.8</td>
<td>±4194.2</td>
<td>40.4%</td>
<td>5473.5</td>
<td>20,379.3</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>37</td>
<td>11,773.7</td>
<td>±3752.8</td>
<td>31.9%</td>
<td>3522.7</td>
<td>22,274.0</td>
<td>11</td>
<td>9127.6</td>
<td>±4250.7</td>
<td>46.6%</td>
<td>2282.0</td>
<td>17,252.8</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>36</td>
<td>12,704.4</td>
<td>±4554.6</td>
<td>35.9%</td>
<td>2053.8</td>
<td>22,169.3</td>
<td>20</td>
<td>11,772.3</td>
<td>±4027.4</td>
<td>34.2%</td>
<td>3884.0</td>
<td>17,284.0</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>50</td>
<td>12,639.0</td>
<td>±3447.8</td>
<td>27.3%</td>
<td>5033.5</td>
<td>21,811.0</td>
<td>22</td>
<td>10,092.9</td>
<td>±4113.4</td>
<td>40.8%</td>
<td>4777.0</td>
<td>19,498.3</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>75</td>
<td>10,697.2</td>
<td>±3809.0</td>
<td>35.6%</td>
<td>3061.0</td>
<td>24,254.8</td>
<td>57</td>
<td>9000.6</td>
<td>±2761.5</td>
<td>30.7%</td>
<td>3954.0</td>
<td>17,525.5</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>155</td>
<td>10,240.2</td>
<td>±3129.5</td>
<td>30.6%</td>
<td>2356.0</td>
<td>24,662.0</td>
<td>94</td>
<td>9059.1</td>
<td>±4075.3</td>
<td>45.0%</td>
<td>1378.0</td>
<td>34,739.0</td>
</tr>
<tr>
<td>Boys</td>
<td>6</td>
<td>8</td>
<td>14,120.7</td>
<td>±3607.9</td>
<td>25.6%</td>
<td>7787.5</td>
<td>18,493.5</td>
<td>7</td>
<td>11,277.8</td>
<td>±3134.6</td>
<td>27.8%</td>
<td>6960.8</td>
<td>16,153.0</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>22</td>
<td>13,117.5</td>
<td>±3881.9</td>
<td>29.6%</td>
<td>5420.8</td>
<td>22,266.3</td>
<td>15</td>
<td>13,711.6</td>
<td>±2871.4</td>
<td>20.9%</td>
<td>7414.5</td>
<td>18,514.0</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>25</td>
<td>14,858.4</td>
<td>±4162.9</td>
<td>28.0%</td>
<td>6883.3</td>
<td>24,003.0</td>
<td>12</td>
<td>12,975.3</td>
<td>±1448.7</td>
<td>11.2%</td>
<td>10,446.5</td>
<td>14,961.3</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>31</td>
<td>14,219.8</td>
<td>±5241.1</td>
<td>36.9%</td>
<td>4764.8</td>
<td>26,967.3</td>
<td>15</td>
<td>11,309.0</td>
<td>±3260.7</td>
<td>28.8%</td>
<td>6150.0</td>
<td>17,585.7</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>22</td>
<td>13,948.3</td>
<td>±4812.0</td>
<td>34.5%</td>
<td>2090.0</td>
<td>21,068.3</td>
<td>36</td>
<td>12,347.6</td>
<td>±4660.6</td>
<td>37.7%</td>
<td>3195.0</td>
<td>26,466.5</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>57</td>
<td>11,833.0</td>
<td>±4530.8</td>
<td>38.3%</td>
<td>3870.0</td>
<td>26,279.7</td>
<td>38</td>
<td>10,525.7</td>
<td>±3745.6</td>
<td>35.6%</td>
<td>1881.0</td>
<td>16,963.8</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>92</td>
<td>12,258.5</td>
<td>±4091.5</td>
<td>33.4%</td>
<td>5075.0</td>
<td>23,376.0</td>
<td>54</td>
<td>10,609.0</td>
<td>±3810.0</td>
<td>35.9%</td>
<td>2013.7</td>
<td>17,609.8</td>
</tr>
</tbody>
</table>

Abbreviation: CV, coefficient of variation.

<sup>a</sup> Body mass index classification based on Cole et al<sup>13</sup> international age- and sex-specific cut points for overweight and obese.
Table 2 Classification Precision and Sensitivity/Specificity of Step-Count Recommendations for Girls (12,000 steps/d) and Boys (15,000 steps/d) 6–12 y by International Body Mass Index Classification—United States, 2002–2005

<table>
<thead>
<tr>
<th>Age (y)</th>
<th>n</th>
<th>Probability of a correct decision</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>OR&lt;sup&gt;d&lt;/sup&gt; (95% CI)</th>
<th>n</th>
<th>Probability of a correct decision</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>OR&lt;sup&gt;d&lt;/sup&gt; (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>28</td>
<td>.46</td>
<td>.75</td>
<td>.42</td>
<td>2.14 (0.19–23.72)</td>
<td>15</td>
<td>.67</td>
<td>.86</td>
<td>.50</td>
<td>6.00 (0.48–75.35)</td>
</tr>
<tr>
<td>7</td>
<td>44</td>
<td>.57</td>
<td>.71</td>
<td>.50</td>
<td>2.50 (0.64–9.77)</td>
<td>37</td>
<td>.38</td>
<td>.64</td>
<td>.27</td>
<td>0.64 (0.14–2.90)</td>
</tr>
<tr>
<td>8</td>
<td>49</td>
<td>.53</td>
<td>.78</td>
<td>.48</td>
<td>3.17 (0.58–17.15)</td>
<td>37</td>
<td>.54</td>
<td>1.00</td>
<td>.39</td>
<td>NA&lt;sup&gt;e&lt;/sup&gt; (NA)</td>
</tr>
<tr>
<td>9</td>
<td>57</td>
<td>.60</td>
<td>.55</td>
<td>.62</td>
<td>2.01 (0.67–6.05)</td>
<td>46</td>
<td>.50</td>
<td>.82</td>
<td>.40</td>
<td>3.00 (0.56–16.01)</td>
</tr>
<tr>
<td>10</td>
<td>72</td>
<td>.50</td>
<td>.61</td>
<td>.46</td>
<td>1.35 (0.46–4.02)</td>
<td>58</td>
<td>.66</td>
<td>.83</td>
<td>.46</td>
<td>4.33 (1.29–14.59)</td>
</tr>
<tr>
<td>11</td>
<td>133</td>
<td>.44</td>
<td>.84</td>
<td>.24</td>
<td>1.66 (0.65–4.27)</td>
<td>95</td>
<td>.43</td>
<td>.88</td>
<td>.21</td>
<td>1.82 (0.54–6.12)</td>
</tr>
<tr>
<td>12</td>
<td>250</td>
<td>.42</td>
<td>.85</td>
<td>.23</td>
<td>1.74 (0.84–3.61)</td>
<td>146</td>
<td>.38</td>
<td>.82</td>
<td>.19</td>
<td>1.03 (0.41–2.57)</td>
</tr>
<tr>
<td>All</td>
<td>633</td>
<td>.47</td>
<td>.77</td>
<td>.34</td>
<td>1.80 (1.21–2.68)</td>
<td>434</td>
<td>.46</td>
<td>.83</td>
<td>.28</td>
<td>1.94 (1.17–3.22)</td>
</tr>
</tbody>
</table>

<sup>a</sup> Body mass index classification based on Cole et al<sup>13</sup> international age- and sex-specific cut points for overweight and obese.

<sup>b</sup> Sensitivity = True Nonmasters/(True Nonmasters + False Master) or Failed to achieve step-count recommendation and unhealthy body mass index/(Failed to achieve step-count recommendation and unhealthy weight + Unhealthy weight and achieved step-count recommendation).

<sup>c</sup> Specificity = True Masters/(True Masters + False Nonmasters) or Achieved step-count recommendation and healthy body mass index/(Achieved step-count recommendation and healthy body mass index + Healthy body mass index and failed to achieve step-count recommendation).

<sup>d</sup> The odds of being classified as a true nonmaster in those youth who exhibited an unhealthy weight relative to the odds of being classified as a true nonmaster in those youth who exhibited a healthy weight.

<sup>e</sup> Odds ratio is undefined owing to sensitivity value of 1.00.
The current findings suggest that daily pedometer step counts do not accurately differentiate US children classified as exhibiting a healthy weight or an unhealthy weight (see Table 2). Given this, it is unclear whether the existing BMI-referenced steps/d cut points are effective for identifying low-active, unhealthy-weight US children, with these findings being consistent regardless of age, gender, or sample size of the stratum.

A possible reason for these findings is that pedometer-determined physical activity levels are likely to limit the development of BMI-referenced steps/d cut points because they provide only a measure of volume of physical activity. Spring-levered pedometers are not able to distinguish physical activity intensity, duration, or frequency. Previous research has demonstrated the dramatic changes in moderate-to-vigorous physical activity participation among school-age youth. Thus, spring-levered pedometers are limited in their ability to discriminate between participants who engage in similar amounts of activity as defined by duration (ie, time) but might engage in different intensity levels. Consider the example of two 12-year-old boys, one classified as overweight and the other as healthy, who both accumulated 13,000 steps/d. One engaged in light-intensity activities (walked home from school, walked with parents after dinner), whereas the other participated in after-school soccer practice for 2 hours. Although there is no difference in pedometer-determined activity, there would be a measurable difference in energy expenditure between the 2 boys. Child studies have demonstrated that body fatness is inversely related with time spent in activities of high intensity (vigorous) but not total activity. It might be that pedometer-measured physical activity (ie, steps/d) does not provide enough information about physical activity to adequately differentiate healthy-weight children and those with an unhealthy BMI.

In addition, the ability of BMI to accurately identify youth that are overweight because of excessive adiposity is debatable. Although BMI might be useful as a coarse proxy for body composition on a population level, these data demonstrate that BMI is unable to sufficiently distinguish between pedometer-determined active and inactive youth. Incorrect classification of BMI can result from excessive lean muscle mass (ie, heavy for height because of muscle) or early maturation (ie, child is compared with norms on chronological age not biological age). Maturation is associated with gains in adipose tissue for both boys and girls, and the prevalence of overweight is markedly higher in early maturing girls. A higher BMI is associated with earlier maturation (age to peak height velocity) in boys. More physically active youth have BMIs that increase to a greater degree with maturation, but these youth accumulate lower total fat mass than their nonphysically active peers. Increasing BMI is also associated with gains in stature to a greater extent than gains in adiposity, which would lead to incorrect classification of weight status. In addition, evidence suggests that increases in BMI, at least for boys, are associated with increases in lean mass rather than increases in fat mass. These scenarios might be occurring in this sample, possibly resulting in youth who are both highly active and identified as overweight according to international BMI standards.

Several limitations need to be illuminated in regard to the current findings. First, 2 pedometer brands were used in the current study. Given that these
pedometers might record step counts with more or less accuracy (actual steps taken versus steps registered by the pedometer) in comparison with the other, there is the potential for over- or underestimation of actual steps accrued. Second, the use of spring-lever pedometers to accurately register steps in individuals with an unhealthy BMI might influence the step counts reported herein. Nevertheless, other studies have demonstrated the accuracy of both models, and therefore, any possible inaccuracies are unlikely to be biased for 1 pedometer brand and/or sample pooled in this study. Third, no information was collected on dietary/caloric intake, and therefore, the authors cannot rule out the influence of this on BMI and, hence, its influence on the results presented herein.

**Future Directions**

As practitioners and researchers we need to be fully cognizant of the limitations of chosen measures, knowing that the selection of a measure of lesser accuracy (pedometer) should be offset by one of greater accuracy (magnetic resonance imaging, bioelectrical impedance). Moreover, because much of this work is taking place outside the laboratory, the feasibility of assessing large numbers of youth takes a substantial role in the selection of the measures used to develop activity cut points. Alternatively, as promoters of physically active lifestyles, the question needs to be raised as to whether blanket prescriptions (steps/d cut points) can be developed and, if so, would their use be counterproductive to increasing the activity levels of youth. These are important issues that require careful consideration when attempting to develop activity cut points for youth.

**Conclusion**

Pedometer-determined daily activity levels were unable to distinguish between youth classified as a healthy weight or unhealthy weight according to international BMI age- and sex-specific categories. The cut points evaluated herein should, therefore, be used and interpreted with caution.

**References**


