Evaluation of the Canadian Aerobic Fitness Test
With 10- to 15-Year-Old Children

Anne W. Garcia and Jennifer S. Zakrajsek

The objective of this study was to evaluate the utility of the Canadian Aerobic Fitness Test (CAFT), a field measure of cardiovascular fitness. After providing anthropometric measures, 31 subjects, ages 10 to 15, completed a maximal treadmill test and the CAFT, a 3-stage step test. Multiple regression analyses were conducted where maximal oxygen consumption from the treadmill test was estimated based on the oxygen cost of stepping, age and various combinations of body composition. For the total sample, the best model ($R = 0.79$, SEE = 6.7), obtained from the sum of 4 skinfolds, was the body composition estimate. This model was slightly more accurate for males ($R = 0.83$, SEE = 6.0) than for females ($R = 0.77$, SEE = 7.0). When the regression equation incorporated less time consuming indicators of body composition, the predictive power, albeit lower, was still satisfactory. It appears that the CAFT can be a useful option for measuring cardiovascular fitness for youth, with the decision dependent on the purpose of the test, the testing resources, the setting, and the motivation of the subjects.

The Surgeon General’s landmark report in 1996 (28) highlights the significance of physical activity to our nation’s health. One issue emphasized in that report is the importance of promoting active lifestyles early among today’s youth. As part of the development of effective exercise programs, it is critical that experts utilize fitness tests that are valid, reliable, and feasible in the field. It is especially important to assess cardiovascular fitness, as this particular fitness component is generally accepted for its strong relationship with other health indices (9). The purpose of this study is to evaluate the utility of the Canadian Aerobic Fitness Test (CAFT) as a field step-test measure of children’s cardiovascular fitness.

Historically, the field test that has been most widely used to assess cardiovascular fitness in school settings has been some variant of the run/walk (e.g., 1 mile, half mile, 600 yard, Cooper’s 12-min test). Running tests are easy to administer to groups and have provided good estimates of cardiovascular endurance in general. Moreover, they can be used specifically to predict maximum oxygen consumption reasonably well under a number of conditions (11, 14, 20). The accuracy of the prediction has been improved by including other key factors such as body size (defined variously as body mass, skinfolds, BMI, etc.), gender, and/or age (6, 11).

The authors are with the Division of Kinesiology at the University of Michigan, Ann Arbor, MI 48109-2214.
Despite these advantages and the general popularity of the run/walk test, criticisms have been raised in the field. In 1977, Cureton and his associates (10) noted that determinants of run/walk tests are complex as only 66% of the variance in a 1-mile run test was explained even when \( \dot{V}O_{2\text{max}} \), height, percent fat, and 50-yard dash time were included in the model. This sentiment was echoed by Simons-Morton and his colleagues (26): “A number of factors, however, can negatively influence children’s performance [in a run/walk test] including experience, motor efficiency, environmental conditions, and motivation” (p. 296). Several investigators have expressed concerns over the validity of the run test when the motivation of the participants is questionable as in the case of the mentally handicapped (21), young children (11), or otherwise inexperienced runners (12). Also run tests are dependent on the space available and climate.

An alternative field test that has received considerable attention is the CAFT. It was first introduced in the 1970s (as the Canadian Home Fitness test) when, based on the positive evaluation by several researchers (4, 15, 17, 23), Recreation Canada administered over 14,000 tests in Canadian homes. A double-step system (20 cm/8 in.) was employed to mirror the typical steps found in a household, making this test simple, accessible, and feasible for children and adults of varying height and fitness levels. The CAFT protocol that has been studied over the past 20 years has varied in terms of the number of stages and exact cadences, depending on the age of the subjects and the researchers’ goals. However, the basic protocol requires that the subject complete successive 3-min stages of stepping, with each stage at an increased pace. Heart rates are taken for a 10-s period five seconds after each stage is completed. Although many researchers have evaluated the validity of the CAFT with adults (2, 21, 24), only a few studies were uncovered that examined its utility with youth.

Investigations of the step test with children began with the work of Bailey and Mirwald in 1978 (3); they provided initial normative values on a pool of 212 subjects, aged 11 to 14. They reported good criterion validity based on the correspondence between the participants’ heart rates and teachers’ ratings of their involvement in sport activities. Working with a much younger population, Yoshinaga and associates (31) also noted the correspondence between a treadmill test and the step test with healthy preschool students and ones with Kawasaki disease who would be tested in a clinical setting. It should be noted, however, that no specific inferential statistic comparing the step test with the treadmill test was mentioned in their report. Other researchers (18) have successfully used the step-test as a challenge to identify children, Grades 3 to 5, with asthma that had heretofore been undiagnosed.

Only one published report was identified that used the CAFT with children or adolescents intentionally to predict \( \dot{V}O_{2\text{max}} \). Jette and his colleagues (16) administered the step test and a maximum treadmill test to 59 Canadian youth, ages 7 to 14. A regression equation was proposed in which the observed \( \dot{V}O_{2\text{max}} \) was predicted using an estimate of the oxygen cost of stepping for that individual (ml \( \cdot \) kg\(^{-1} \) \( \cdot \) min\(^{-1} \)), the heart rate achieved following final stage of stepping, and the sum of four skinfolds. The multiple Rs reported were quite high (.92 and .96 for boys and girls, respectively), supporting the criterion validity of the step test. However, it should be noted that Jette and his colleagues apparently reported the multiple R after converting the maximum oxygen consumption score from ml \( \cdot \) kg\(^{-1} \) \( \cdot \) min\(^{-1} \) to the absolute amount of oxygen utilized (L \( \cdot \) min\(^{-1} \)) instead of retaining the
more accepted measure, which adjusted for body weight. Thus, to date, direct comparisons of the predictive utility of the CAFT as compared to other field tests of cardiovascular fitness among children are not available. Additional work with the CAFT should ideally be conducted with ml·kg\(^{-1}\)·min\(^{-1}\) as the metric of interest.

Furthermore, to improve its predictive utility, many modifications of the CAFT protocol and/or regression equation for adults have been considered recently (8, 29). One such modification is identifying the best indicator of body size. In the adult literature with the CAFT, researchers have experimented with body weight, body mass index, body fat percentage, and various skinfold combinations as estimates of body fat percentage. In his lab study with children and adolescents, Jette and colleagues used the sum of four skinfolds (triceps, biceps, subscapular, and suprailiac). However, this technique could be impractical in epidemiological or applied settings as it is relatively time consuming and requires disrobing. In a field setting considerable time can be saved by eliminating or at least reducing the number of skinfolds that have to be obtained. Importantly, the four sites selected were all from the upper body, possibly mitigating the goal of being more representative than a two site protocol. Thus, the purpose of this study is to build on the work by Jette and his colleagues (16) to evaluate the validity of the CAFT as an indirect measure of maximum oxygen consumption, including a comparison of its predictive power when various indicators of body composition are included in the equation.

**Method**

**Subjects**

Human subjects approval was obtained from the Health Sciences Institutional Review Board at the University of Michigan. The 31 participants, 19 boys and 12 girls, were primarily Caucasian (74%) and ranged in age from 10 to 15, with a mean age of 12 (SD = 1.7). They were recruited from PE classes at a local middle school, a local summer sports camp, and by word of mouth. Introductory letters and consent forms were distributed to the parents of these children.

**Measures**

**Height, Weight, BMI.** Total body mass and height were measured using a Health-O-Meter (Continental Scale, Bridgeview, IL) balance scale with height rod. Measurements were made to the nearest 0.1 pound and inch, respectively. BMI was calculated as the total body mass (kg) / height (m\(^2\)).

**Skinfold Assessment.** Skinfold thicknesses were measured using Harpenden skinfold calipers at five sites (subscapular, triceps, biceps, suprailiac, and calf) according to standardized measurement techniques (19). Three measurements were made at each site, and the average of the three measurements was used in data analysis. The sum of subscapular, suprailiac, triceps, and biceps measurements were combined for the sum of four skinfolds (SS4; 16). An alternative measure, the sum of two skinfolds (SS2), consisted of the sum of the triceps and calf skinfold thicknesses. Percent body fat was calculated with the latter index (27).

**Treadmill Test.** Resting, exercise, and recovery heart rates were measured and recorded using a 12-lead ECG and the Marquette Electronics MAX-1 Stress
System (Marquette Electronics, Jupiter, FL). Each child was fitted with a three-way breathing valve, supported by an adjustable headgear, and noseclips. Expired air was collected and analyzed by the Sensor Medics 2900 Energy Expenditure Unit (Sensor Medics, Yorba Linda, CA). $\text{VO}_2$ (ml · kg$^{-1}$ · min$^{-1}$) and $R$ were recorded at 20-s intervals throughout rest, exercise, and recovery.

For the treadmill test, resting heart rate and expired air were measured for 3 min at baseline. The Marquette Electronics 2000 treadmill (Marquette Electronics, Jupiter, FL) was started at 3 mph at a 5% grade. Over a 1-min period, the speed and grade were gradually increased to 3.5 mph and 10%. The speed was increased by .7 mph every 2 min, while the grade remained at 10%. At the midpoint of each 2-min stage, heart rate and a Rating of Perceived Exertion (RPE) (5) were recorded. The subject was encouraged to run as long as possible, and the test was terminated at the request of the subject. The recovery period lasted a total of 3 min. Minute 1 consisted of active recovery on the treadmill and, during the final 2 min, the subject was seated on a chair. The dependent variable was defined as the highest level of oxygen consumption achieved during the treadmill test. There is a lack of consensus as to the criteria for defining a maximal test (13), thus making it difficult to ascertain whether true maximum exertion was achieved, especially in a sample of 10 to 15 year olds. We will use the term $\text{VO}_2$ for our dependent variable for the remainder of our discussion, consistent with other researchers who use this term, to indicate the highest oxygen utilization value that the subject achieved. We would note, however, that the majority of the subjects probably did attain true maximum exertion or very close to it, as 77% of them attained an RQ of 1.0 or higher.

**Step Test.** The step test was administered according to the recommendations of Jette et al. (1984) (16) and in accordance with the operations manual sponsored by the Canadian government (7). The step test protocol for youth consists of, at most, three progressive 3-min stepping bouts. Subjects coordinated their stepping to predetermined cadences provided by audio tapes from the Canadian Home Fitness Test. Tapes used for 10-year-old subjects had a maximum cadence of 132 beats · min$^{-1}$, while those for subjects 11–15 years old had a maximum cadence of 144 beats · min$^{-1}$ (16). At the end of each 3-min stepping bout, a 10-s radial pulse was measured manually and recorded. If a subject attained a heart rate greater than 180 at the completion of Stage 1 or Stage 2, the test was terminated.

Based on this test, two values were obtained: (a) the 10-s heart rate value and (b) an estimated oxygen cost of stepping. The latter value was determined in three steps. First, the rate of stepping, number of stages completed, the height of the steps, and the weight of the participant were used to calculate the work performed during stepping (kg · m · min$^{-1}$). Second, work was converted to the metabolic cost of stepping (kcal · min$^{-1}$). Finally, this value was expressed as oxygen cost of stepping (ml · kg$^{-1}$ · min$^{-1}$; see Jette et al., 1984, p. 216).

**Procedure**

Following parental consent, testing was conducted in two sessions. The sessions were completed on consecutive days. Prior to the collection of any data, each child was presented with an age appropriate written description of all testing to take place during both sessions. Subjects were instructed to read the document and sign it to indicate their assent to participate in the study. The first session consisted of the collection of the anthropometric measurements (height, weight, and the five skinfold thicknesses) and the completion of the treadmill test. The second session
consisted of only the step test. This order was maintained through the study since, given the differences in the tests and the age of the subjects, it seemed unlikely that learning or fatigue would confound our results.

Data Analysis

All data calculations and analyses were performed using the SPSS software package. To obtain the best estimate of VO\textsubscript{2peak}, regression analyses were conducted with various combinations of predictors. The estimated oxygen cost of stepping (ml \cdot kg\textsuperscript{-1} \cdot min\textsuperscript{-1}), based on the equation proposed by Jette and colleagues (16), was entered in the first step of every equation. Other determinants considered included heart rate at completion of last stage of stepping, age, and various indicators of body composition (e.g., BMI, a two-skinfold measure, or a four-skinfold measure). Following these calculations, paired t-tests were conducted comparing the directly measured VO\textsubscript{2peak} with each of three best estimates of VO\textsubscript{2peak}.

Results

Anthropometric characteristics for the boys and girls are presented in Table 1. There were no significant differences in body fat percentage between the boys and the girls. The range of fitness levels represented by the participants was relatively wide with VO\textsubscript{2peak} varying from 35.49 ml/kg/min to 71.52 ml/kg/min.

Regression analyses were performed to determine the strongest combination of predictors for estimating VO\textsubscript{2peak}. Heart rate at the conclusion of the last stage of stepping did not contribute significantly to the model and was not included. A model including the oxygen cost of stepping (ml \cdot kg\textsuperscript{-1} \cdot min\textsuperscript{-1}), age, and any one of the three measure of body composition explained the most variance in VO\textsubscript{2peak}. The resulting multiple regression equations and analyses are presented in Table 2. Three equations are presented with each of the different indices of body composition: SS4, SS2, and BMI. The equation containing SS4 resulted in the largest multiple R and lowest SEE followed by SS2 and BMI, respectively. Each equation produces stronger estimates for males than for females.

Table 1  Means and Standard Deviations of Anthropometric Characteristics of the Subjects

<table>
<thead>
<tr>
<th>Variable</th>
<th>Total Sample ((N = 31))</th>
<th>Males ((N = 19))</th>
<th>Females ((N = 12))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>12.5 ± 1.66</td>
<td>12.45 ± 1.76</td>
<td>11.41 ± 1.29</td>
</tr>
<tr>
<td>Height (cm)</td>
<td>153.76 ± 11.58</td>
<td>157.43 ± 10.85</td>
<td>147.97 ± 10.63</td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>47.07 ± 12.61</td>
<td>51.43 ± 13.10</td>
<td>40.18 ± 8.25</td>
</tr>
<tr>
<td>Body Mass Index (kg \cdot (m\textsuperscript{2})\textsuperscript{-1})</td>
<td>19.62 ± 3.2</td>
<td>20.51 ± 3.43</td>
<td>18.21 ± 2.27</td>
</tr>
<tr>
<td>Sum of two skinfolds (mm)</td>
<td>26.78 ± 9.61</td>
<td>27.13 ± 10.82</td>
<td>26.21 ± 7.73</td>
</tr>
<tr>
<td>Sum of four skinfolds (mm)</td>
<td>43.49 ± 20.55</td>
<td>43.95 ± 23.34</td>
<td>42.76 ± 16.10</td>
</tr>
<tr>
<td>Percent body fat</td>
<td>21.00 ± 6.79</td>
<td>20.94 ± 7.96</td>
<td>21.09 ± 4.72</td>
</tr>
</tbody>
</table>
Table 2  Comparison of the Observed \( \dot{V}O_{2peak} \) (From the Treadmill) to the Estimated \( \dot{V}O_{2peak} \) (From the Step Test)

<table>
<thead>
<tr>
<th>Sample</th>
<th>Multiple regression equations</th>
<th>Mult. R</th>
<th>SEE</th>
<th>%error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total sample</td>
<td>( \dot{V}O_{2peak} = 2.63(OC) - .29(SS4) + .49(age) - 29.57^a )</td>
<td>0.79</td>
<td>6.7</td>
<td>12%</td>
</tr>
<tr>
<td>Males</td>
<td></td>
<td>0.83</td>
<td>6.0</td>
<td>11%</td>
</tr>
<tr>
<td>Females</td>
<td></td>
<td>0.77</td>
<td>7.1</td>
<td>13%</td>
</tr>
<tr>
<td>Total sample</td>
<td>( \dot{V}O_{2peak} = 3.53(OC) - .52(SS2) + .55(age) - 54.62^a )</td>
<td>0.76</td>
<td>7.1</td>
<td>13%</td>
</tr>
<tr>
<td>Males</td>
<td></td>
<td>0.80</td>
<td>6.5</td>
<td>12%</td>
</tr>
<tr>
<td>Females</td>
<td></td>
<td>0.73</td>
<td>7.5</td>
<td>14%</td>
</tr>
<tr>
<td>Total sample</td>
<td>( \dot{V}O_{2peak} = 3.23(OC) - 1.31(BMI) + 1.39(age) - 49.21^a )</td>
<td>0.71</td>
<td>7.7</td>
<td>14%</td>
</tr>
<tr>
<td>Males</td>
<td></td>
<td>0.76</td>
<td>7.0</td>
<td>13%</td>
</tr>
<tr>
<td>Females</td>
<td></td>
<td>0.71</td>
<td>7.8</td>
<td>14%</td>
</tr>
</tbody>
</table>

^aWhere OC = oxygen cost of stepping; SS4 = sum of four skinfold measurements; SS2 = sum of two skinfold measurements; and BMI = body mass index.

Mean \( \dot{V}O_{2peak} \) values from the treadmill test and each of the three prediction equations are presented in Table 3 for the total sample and boys and girls separately. Paired t tests conducted to compare observed \( \dot{V}O_{2peak} \) and each of the estimated \( \dot{V}O_{2peak} \) values revealed that none of the three-step test \( \dot{V}O_{2peak} \) estimates were significantly different from the treadmill measurement \((p < .01)\) for the total sample, or for boys or girls separately. Also, the estimate including BMI did not differ significantly from either the estimate using SS4 \((p < .01)\) or the estimate with SS2 \((p < .05)\) for any group.

Table 3  Means and Standard Deviations of Observed \( \dot{V}O_{2peak} \) for the Treadmill Test and Estimated \( \dot{V}O_{2peak} \) for the Step Test

<table>
<thead>
<tr>
<th>Test</th>
<th>Total sample (( N = 31 ))</th>
<th>Males (( n = 19 ))</th>
<th>Females (( n = 12 ))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treadmill Test (ml · kg(^{-1} ) · min(^{-1} ))</td>
<td>55.7 ± 10.7</td>
<td>58.0 ± 10.5</td>
<td>52.1 ± 10.5</td>
</tr>
<tr>
<td>Step Test (ml · kg(^{-1} ) · min(^{-1} )) using ( O_2 )cost, age, and SS4</td>
<td>55.7 ± 8.5</td>
<td>56.1 ± 9.1</td>
<td>55.0 ± 7.7</td>
</tr>
<tr>
<td>Step Test (ml · kg(^{-1} ) · min(^{-1} )) using ( O_2 )cost, age, and SS2</td>
<td>55.7 ± 8.1</td>
<td>56.0 ± 8.5</td>
<td>55.2 ± 7.7</td>
</tr>
<tr>
<td>Step Test (ml · kg(^{-1} ) · min(^{-1} )) using ( O_2 )cost, age, and BMI</td>
<td>55.7 ± 7.6</td>
<td>55.5 ± 7.6</td>
<td>56.0 ± 7.8</td>
</tr>
</tbody>
</table>

Note. Within each group there were no significant differences between mean values.
Figure 1 — Scatterplots for observed versus estimated $\dot{V}O_2$peak using the four-site skinfold estimate of body composition (A), the two site skinfold estimate of body composition (B), and BMI as an estimate of body composition (C).
Scatterplots of observed $V_{O_2\text{peak}}$ (treadmill test) versus estimated $V_{O_2\text{peak}}$ (step test) are presented in Figure 1. Figure 1 depicts step test $V_{O_2\text{peak}}$ calculated using SS4, SS2, and BMI, respectively.

**Discussion**

The Canadian Aerobic Fitness Test (CAFT) has been widely used to classify adults into relatively broad fitness categories. Moreover, some have suggested that the test is also sufficiently accurate to estimate $V_{O_2\text{max}}$. Although several groups of researchers have evaluated questions such as these with adults under various conditions, little work with youth was uncovered in the literature. The primary goal of this study was to build on the work of Jette and his colleagues (16) concerning the validity of the CAFT in assessing the cardiovascular fitness of children and teenagers. The question at hand was under what conditions and for what purposes might the CAFT prove to be the field test of choice for this age group.

We began our analyses by comparing various regression equations to maximize the variance predicted in $V_{O_2\text{peak}}$. For the total sample, the best prediction was based on the estimated cost of stepping (ml·kg$^{-1}$·min$^{-1}$) in combination with age and the sum of four skinfolds. The prediction was stronger for the males than the females, possibly because there was more variability in the skinfold measures in our sample among the males. Clearly the multiple $R$s in our study are lower than those reported by the Jette et al. study (16). However, this can be largely attributed to the use of the relative measure of oxygen utilized in the current study.
as compared to using absolute oxygen consumption. It is well accepted that
the relationship reported will not be as strong when the metric is ml · kg\(^{-1}\) · min\(^{-1}\)
because (a) the lower variance in the relative measure attenuates the relationship,
and (b) the size of the subject no longer contributes to the explained variance.

It was somewhat surprising that heart rate did not make a statistically signi-
ficant contribution to the model. This may be explained in part by the relatively
low variance in this index as (a) 87% of the subjects completed all three stages of
the step-test, and (b) with the exception of one child, all of the subjects had ending
heart rates of at least 150. Also, the accuracy of palpating fast brachial pulses has
been questioned (24). Further work is warranted to determine whether a mechani-
cal measure of heart rate, as suggested by Montgomery and his colleagues (21),
would improve the predictive power of this measure.

As mentioned earlier, the decision concerning which indicator of body com-
position should be used to accompany this field measure is probably best deter-
mined by feasibility as well as accuracy. As seen in Table 2, the equation that
incorporated the four-skinfold sum was most accurate \((R = 0.79, \text{SEE} = 6.6)\); how-
ever, the equation with the two-skinfold measure was still worthy of considera-
tion \((R = 0.76, \text{SEE} = 7.1)\) and had the added advantage of taking less administra-
tion time and not requiring disrobing. Furthermore, the equation incorporating BMI \((R
= 0.70, \text{SEE} = 7.7)\) instead of a skinfold measure could be considered, albeit at a
cost to accuracy, if there was not appropriate staff and/or time to measure skinfolds
at all. Thus the final decision as to which estimate of body composition to use
should probably depend on the number of participants under evaluation, the expe-
rience and resources of the testers, and the time allotted for administration of these
tests.

The step-test in this study, as in other studies, continues to be, at best, an
adequate measure of \(\dot{V}O_2\text{peak}\). The same criticisms that have been raised in research
with the CAFT among adults (24, 25) were likely in place in our analyses: (a) A
ceiling effect that makes predictions with highly fit individuals less accurate; (b) a
lack of variance in the measure in general; and (c) difficulty arriving at an equation
that generalizes across gender, ages, fitness levels, and body size. In response to
these comments, one group of researchers has advocated adding additional stages
as necessary so that each subject attains 85% of his or her estimated maximal heart
rate (29, 30). Although this protocol does improve the generalizability of the for-
formula, it does so at the expense of making the test more time consuming. Moreover,
this adaptation may not be needed in the youth protocol used by Jette and our-
selves, as the heart rates reported by our subjects was quite high. In fact, to the
degree that this test was intended to be a submaximal effort for most individuals,
and feasible in an applied setting (e.g., the school or health clinic), the simplicity
of the original CAFT protocol continues to have some appeal when all consider-
ations are weighed. The more accurate, but more time consuming, modified CAFT
would be an unlikely choice for a school setting. A more recent recommendation
by Shephard and Bouchard (25) was to predict adult CAFT scores differently by
taking into account age-adjusted maximum heart rates and thus producing a more
variable outcome measure. However, the new predicted value is in METs instead
of oxygen utilization, a more precise measure. Moreover, this age adjustment is
not relevant in our sample of youth.

In summary, in making decisions as to which field test of cardiovascular
endurance is best, the epidemiological researcher and the physical education teacher
should take several factors into account. On the one hand, in many situations, a run-walk test would be the best choice as large groups can be measured at one time and, under optimal testing circumstances, this test yields a highly accurate measure of VO$_{2\text{peak}}$. Indeed, Buono and his associates (6) provided some of the strongest evidence when they compared a 1-mile run with a step test and a bicycle ergometer test with 10 to 18 year olds. The multiple regression model that predicted VO$_{2\text{max}}$ the best was based on mile run time in seconds, sum of two skinfolds, gender, and weight, leading the authors to recommend the run/walk test as the best field test of cardiovascular fitness currently available. However, the comparison of the run test with the step test in their study may not have been appropriate, since for their step test they used an adult protocol and height (16.25") that would generally not be recommended for a sample of youth.

Furthermore, Buono and his colleagues (6) acknowledge that the high motivation of their subjects might limit the external validity of their findings. It is this latter point that is particularly relevant when deciding which test to use. In the authors’ experience, there is a sizable proportion of participants who will not choose to exert a maximum effort during a run/walk. Analyses of fitness data gathered on over 2,000 youth from low socioeconomic status and diverse schools (12) revealed that approximately 36% of the youth took 4-min or longer to complete a 600-yard test. Based on the original normative values (1), one would expect that 5% or fewer would fall into such a low fit category. Given that the correlations and error for the CAFT in our study ($R = 0.79$, SEE = 6.6) are not radically different than the results for a widely accepted 1-mile run equation ($R = 0.72$, SEE = 4.8; 11), the CAFT deserves further consideration. Moreover, the total accuracy of each test should be considered, including the number of students who would not provide their best effort and then receive a low estimate of their maximum score.

These findings suggest that the CAFT can be a useful field measure of cardiovascular fitness for youth. Advantages to step tests in general include (a) the low cost and ease of administration; (b) the ability to administer this test in a small indoor space, irrelevant of climate; and (c) the ability to provide a standard stimulus to children independent of their levels of motivation. However, there are still some limitations to this technique; the primary one is administration time, making this procedure less feasible for moderate to large samples. Modifications seem possible such as adapting wider, naturally occurring steps in the school (e.g., staircases, bleachers) so that more than one student can be assessed at the same time. Also, a comparable step test that employs a faster cadence, but for a shorter period of time, could be investigated. Finally as heart rate monitors have become more accessible, heart rate could be assessed more accurately and potentially contribute to the predictive power of the CAFT. Further research seems warranted to refine the step-test protocol to maximize its feasibility for a variety of settings. If the efficiency of the administration of the CAFT can be improved, we would recommend it as the cardiovascular field test of choice when the motivation of the children being tested is in question.

References


**Acknowledgements**

We would like to thank Jon Strite of Slauson Middle School and Pat Van Volkinburg and Kerry Winkleseh from the Kidsport Program at the University of Michigan for facilitating subject recruitment. We also are very grateful to many students who assisted with screening subjects and data collection, including Sean Semple, Cara Monterosso, Chumbhot Plangtrakul, Alysa Ullman, and Jennifer Chuang. Finally, we greatly appreciate the advice and counsel of James Pivarnik in our revisions of this paper.