Development and Validation of a One-Mile Treadmill Walk Test to Predict Peak Oxygen Uptake in Healthy Adults Ages 40 to 79 Years

David M. Pober, Patty S. Freedson, Gregory M. Kline, Kyle J. McInnis, and James M. Rippe

---

Catalog Data

Pober, D.M., Freedson, P.S., Kline, G.M., McInnis, K.J., and Rippe, J.M. (2002). Development and validation of a one-mile treadmill walk test to predict peak oxygen uptake in healthy adults ages 40 to 79 years. **Can. J. Appl. Physiol.** 27(6): 575-588. ©2002 Canadian Society for Exercise Physiology.

Key words: *$\dot{V}O_{2max}$ estimation, sub-maximal exercise testing, fitness, aerobic power, Treadwalk test*

Mots clés: *estimation du $\dot{V}O_{2max}$, épreuve sous-maximale, condition physique, puissance aérobie maximale, test de marche (*treadwalk*)*

Abstract/Résumé

*The purpose of this investigation was to determine whether the Rockport one-mile walk test equation to predict maximal oxygen uptake was valid for application to treadmill walking. When the Rockport model was found to be inappropriate, a new regression model was developed for predicting peak oxygen uptake ($\dot{V}O_{2peak}$) from a one-mile treadmill walk. 304 healthy volunteers ages 40 to 79 years (mean age = 57.6 years, 154 men and 150 women) completed a $\dot{V}O_{2peak}$ test and a one-mile treadmill walk. Stepwise regression was used to build a model for the relationship between $\dot{V}O_{2peak}$ and a variety of predictor variables in a sub-sample development group (*n *= 154). This new model was then applied to a sub-sample validation group (*n *= 150). The new equation produced a correlation of 0.87, SEE = 4.7 $ml \cdot kg^{-1} \cdot min^{-1}$ with a mean residual of 0.96 $ml \cdot kg^{-1} \cdot min^{-1}$. The equation for predicting $\dot{V}O_{2peak}$ developed in this investigation provides a means of assessing $\dot{V}O_{2peak}$ that is easy to*

D.M. Pober and P.S. Freedson are with the Department of Exercise Science, Exercise Physiology Laboratory, University of Massachusetts, Amherst, MA; G.M. Kline, K.J. McInnis, and J.M. Rippe are with the Rippe Lifestyle Institute, Worcester, MA.

administer, allows for careful supervision of subjects, and can be completed at a low financial and temporal cost.

Cette étude cherche à établir si le test de marche de Rockport (1,6 km), conçu pour l'estimation de la consommation maximale d'oxygène, est valide lorsqu'adapté au tapis roulant. L'équation de régression du test initial ne convenant plus, une nouvelle équation fut mise au point pour l'épreuve sur tapis roulant. Trois cent quatre volontaires, en bonne santé, âgés de 40 à 79 ans (150 femmes et 154 hommes dont la moyenne d'âge est de 57,6 ans) participent à une épreuve d'évaluation du $\dot{V}O_2$ *de crête et à un test de marche sur une distance de 1,6 km (1* mile*). Une analyse de régression par degrés est utilisée pour estimer le* $\dot{V}O_2$ *de crête à partir de quelques variables indépendantes d'un sous-échantillon (*n = *154). La nouvelle équation appliquée au sous-échantillon de validation (*n *= 150) révèle un coefficient de corrélation de 0,87 avec une erreur type d'estimation égale à 4,7 mL ·* kg^{-1} *·* min^{-1} *et une moyenne des résidus égale à 0,96 mL ·* kg^{-1} *·* min^{-1}*. Cette nouvelle équation permet donc la prédiction du* $\dot{V}O_2$ *de crête au cours d'un test de marche sur tapis roulant; le test est facile à administrer, permet un encadrement minutieux, prend peu de temps et est peu coûteux.*

Introduction

Peak oxygen uptake ($\dot{V}O_{2peak}$), or peak aerobic exercise capacity, is widely considered to be an excellent indicator of cardiorespiratory fitness. The direct assessment of $\dot{V}O_{2peak}$ is common in clinical and research settings to assess aerobic fitness, for diagnostic purposes, and to examine the effectiveness of endurance training programs for healthy individuals and individuals involved in clinical exercise programs. However, this direct assessment has several drawbacks that limit its practical application. Direct determination of $\dot{V}O_{2peak}$ is relatively expensive and time consuming to conduct, and therefore not well suited for testing large numbers of people. Additionally, administering maximal exercise tests, with or without measurement of gas exchange, is often not feasible within the context of limited third-party reimbursement for such services. Moreover, direct assessment of $\dot{V}O_{2peak}$ requires a maximal or near-maximal effort from the subject, which is often difficult to elicit from individuals with low cardiorespiratory fitness. The need for direct supervision by a physician when evaluating older individuals or those with known cardiovascular disease or associated risk factors adds to the challenge of administering maximal exercise tests to assess $\dot{V}O_{2peak}$.

Because of the limitations to direct $\dot{V}O_{2peak}$ assessment, substantial efforts have been directed towards the development of less strenuous, less time consuming, and more cost effective means of measuring $\dot{V}O_{2peak}$. While many submaximal $\dot{V}O_{2peak}$ prediction tests have been developed (McNaughton, 1998), these tests have generally been developed on specific samples of the population (George, 1996; McArdle et al., 1972; Rintala et al., 1997), do not have validation studies to support them (Falls et al., 1966; Fox, 1973; Mastropaolo, 1970), or do not provide a comprehensive picture of measurement error that is necessary for analysis of the predictive accuracy of the test (Cooper, 1968; Doolittle and Bigbee, 1968; Hermiston and Faulkner, 1971). The limitations of predicting $\dot{V}O_{2peak}$ from submaximal tests led to the development of the Rockport one-mile overground walk test to estimate

$\dot{V}O_{2peak}$ in 30 to 59 year olds, using time to complete the one-mile walk, heart rate at the end of the walk, age, gender, and body mass as predictor variables (Kline et al., 1987).

The Rockport test was based on a large, heterogeneous sample, and demonstrated very good predictive accuracy on a cross-validation sample from the same population (*r* = 0.92, SEE = 4.97 ml · kg^{-1} · min^{-1}). This test has since been validated on numerous occasions using a variety of populations with generally good results (*r* = 0.68 to 0.91, SEE = 3.17 to 6.26 ml · kg^{-1} · min^{-1}) (Fenstermaker et al., 1992; George et al., 1998; Rintala et al., 1997). A modification of the Rockport test was suggested by Widrick et al. (1992) who noted that if this popular prediction test that could be performed in the laboratory as well as in the field, it would have increased utility. The ability to use this test on a treadmill would allow careful monitoring of subjects, and allow testing by facilities that do not have access to a measured mile. Widrick et al. conducted a study to determine whether the Rockport equation developed for overground walking could successfully be applied to treadmill walking (Widrick et al., 1992).

Widrick et al. (1992) reported a high correlation (*r* = 0.91) between predicted and actual $\dot{V}O_{2peak}$ using the Rockport equation, with a total error of 5.26 ml · kg^{-1} · min^{-1}. It appeared from this research that the Rockport equation was suitable for application to treadmill testing. These authors, however, provided no measure of the appropriateness of the regression model for their data, such as an examination of the behavior of the residuals generated by the model. Furthermore, this technique was only applied to a relatively young (mean age = 37.8 years) and fit (mean $\dot{V}O_{2peak}$ = 42.0 ml/kg/min) group (Widrick et al., 1992). Evaluation of the appropriateness of the Rockport model using a treadmill one-mile walk test, and validation on diverse populations are both essential if this method is to be widely applied.

Therefore, the purpose of Part 1 of this investigation was to validate the use of the Rockport one-mile over-ground walk test equation to predict $\dot{V}O_{2peak}$ on a group of older individuals (40–79 years old, mean = 57.6 years) with a broad range of $\dot{V}O_{2peak}$ values (16–56 ml · kg^{-1} · min^{-1}, mean = 31.5 ml · kg^{-1} · min^{-1}), using a one-mile treadmill walk test. If the Rockport model was determined to be inappropriate, then Part 2 of this investigation would develop a new prediction equation on a subset of this sample population and cross-validate it on a different sub-sample group. This new equation would then be examined using the same techniques that were applied to the Rockport equation. The validity of the new model would be further assessed using the predicted residual error sum of squares (PRESS) method as described by Holiday et al. (1995).

Methods

SUBJECTS

There were 350 healthy male and female volunteers, ages 40 to 79 years, recruited for this study through posters, flyers, and newspaper advertisements. Subjects were included in the study if they were between the ages of 40 and 79, were not taking any medications that would alter the heart rate response to exercise or metabolic function, and had no known coronary heart disease or musculoskeletal limitations as assessed by a staff physician during a resting ECG, and a stress test with ECG and blood pressure monitoring. All subjects read and signed an informed consent document that was approved by the New England Institutional Review Board.

TESTING PROTOCOLS

Subject weight was recorded using a calibrated balance scale. Additionally, subjects' habitual physical activity levels were assessed using the 7-point scale described by Heil et al. (1995) and Matthews et al. (2000). Subjects rated their physical activity level from 0 (avoid walking or exertion, e.g. always use elevator, drive whenever possible instead of walking) to 7 (run more than 10 miles per week or spend more than 3 hours per week in comparable physical activity).

Treadmill one mile walk testing was carried out according to the protocol described by Widrick et al. (1992). During a brief (~5 minute) speed selection period, subjects were instructed to select a brisk pace that they could maintain for fifteen to twenty minutes. The test then began with the treadmill speed remaining constant throughout the test. Heart rate was monitored continuously throughout each test using Polar Accurex II ® wireless chest-band heart rate monitors (Polar Electro, Finland). When the subject reached one mile, the time to complete the test was recorded. Walk heart rate was determined by averaging the heart rate from the second to last and last full minutes of the treadmill walking test. The results of one trial were considered sufficient based on the analyses by Kline et al. (1987) showing virtually identical findings using walk test results under strict test-retest reliability criteria as compared to results from a single test.

On a separate day, $\dot{V}O_{2peak}$ was assessed using a modified Balke protocol. Subjects began the test walking on a level treadmill at a self-selected brisk pace. Treadmill grade was increased 2.5% every two minutes until the subject could no longer continue. Pulmonary gas exchange variables were monitored continuously throughout each test on a breath-by-breath basis, via open circuit spirometry (Medical Graphics, Minneapolis, MN). Gas exchange variables were averaged over 30-second intervals. The gas analysis system was calibrated before and after each test with gasses of known concentration, and with a known volume of air. Heart rate and cardiac rhythm were monitored continuously throughout each test using a 12-lead electrocardiogram. Blood pressure and rating of perceived exertion (RPE) were obtained every two minutes, during the final 45 seconds of each stage. The test was considered a peak effort, and the peak oxygen consumption value was included in the analysis, if three of the following four criteria were met (a) a respiratory exchange ratio greater than 1.10, (b) a final RPE score of 17 or greater on the Borg scale, (c) a maximal heart rate within 15 beats per minute of age predicted maximum (220 – age), and (d) an increase in oxygen consumption between the final two stages of the test that was less than the average increase between the previous workloads (Powers and Howley, 2001). If at least three criteria were not met, the data were not included in the analysis of $\dot{V}O_{2peak}$. Three hundred four subjects (155 men, mean age 57.9 years, range 40 to 79 years, and 149 women, mean age 57.2 years, range 40 to 78 years) successfully completed both tests and had their data included in the final analyses. Table 1 shows means (± SD) and ranges for subject characteristics of the overall sample.

Table 1 Means (SD) and Ranges for $\dot{V}O_{2peak}$ and All Predictor Variables (m = 155, f = 149) for the Overall Sample

	Mean (SD)	Range
$\dot{V}O_{2peak}$ ($ml \cdot kg^{-1} \cdot min^{-1}$)	31.5 (8.85)	12–58
Walk time (min)	17.2 (3.16)	11.1–35.2
Walk heart rate (bpm)	121.1 (18.35)	79–182
Age (yr)	57.5 (10.52)	40–79
Body weight (kg)	76.7 (15.85)	42.1–129.3
Activity score	4.3 (2.12)	0–7

Statistical Analyses

All statistical analyses were carried out using SAS version 8 for Windows.

PART 1

Predicted $\dot{V}O_{2peak}$ values were computed using the Rockport general equation developed by Kline, et al. (1987) as follows:

> $\dot{V}O_{2peak}$ Rockport = 132.85 – 0.077 (body mass in pounds) – 0.39 (age in years) + 6.32 (gender; male = 1, female = 0) – 3.26 (walk time in minutes) – 0.16 (walk heart rate)

Pearson correlations between observed and predicted $\dot{V}O_{2peak}$ values were computed, along with the standard error of the estimate (SEE, defined as SD * [$\sqrt{1-r^2}$]) and total error (TE, defined as $\sqrt{[\Sigma(y'-y)^2/n]}$) as described by Lohman (1981). Appropriateness of the model was assessed using Bland – Altman plots (Bland and Altman, 1995) and normal probability plots of the residuals.

PART 2

When the Rockport general equation was determined to be inappropriate for treadmill walking data in this population, a new equation for predicting $\dot{V}O_{2peak}$ was developed and validated using the one-mile treadmill walk data from the present sample.

After randomization of subject numbers, subjects were placed into either a development (N = 154) or cross-validation group (N = 150) based on an odd-even case selection, for development of the new equation (TREADWALK). T-tests with adjustment for multiple comparisons were used to confirm the similarity of the development and cross-validation groups. TREADWALK was developed using stepwise regression on the development group. TREADWALK was evaluated on the cross-validation group using Pearson correlations, SEE, and TE to compare observed and predicted $\dot{V}O_{2peak}$. Bland-Altman and normal probability plots of the residuals were used to evaluate the appropriateness of the model.

To provide confirmation of the results of the cross validation, a regression model using the predicted residual error sum of squares (PRESS) method as described by Holiday et al. (1995) was applied to the sample data. The PRESS method is a statistical jackknife procedure whereby a regression model is developed using n –1 of the sample data, and then applied to the remaining subject to obtain a residual (measured $\dot{V}O_{2peak}$ – predicted $\dot{V}O_{2peak}$). This subject is then replaced in the data pool, and the procedure is repeated until individual residuals have been generated for all of the subjects. The PRESS method provides residuals, SEE, and TE, which may then be used to evaluate the suitability of regression modeling of the sample population. This analysis provided a second means of validating the results of the present investigation.

Results

PART 1

Application of the Rockport general equation to the treadmill test results from the present sample yielded a correlation of $r = 0.80$ between observed and predicted scores, SEE = 7.7 ml · kg^{-1} · min^{-1}, and TE = 9.9 ml · kg^{-1} · min^{-1}. Figure 1 illustrates observed versus predicted $\dot{V}O_{2peak}$ for the application of the Rockport general equation to the present sample.

Examination of the Bland-Altman (Figure 2) and normal probability plots of the Rockport general equation residuals (observed – predicted) indicated that the mean residual was 6.2 ml · kg^{-1} · min^{-1}, and that the residuals were not normally distributed.

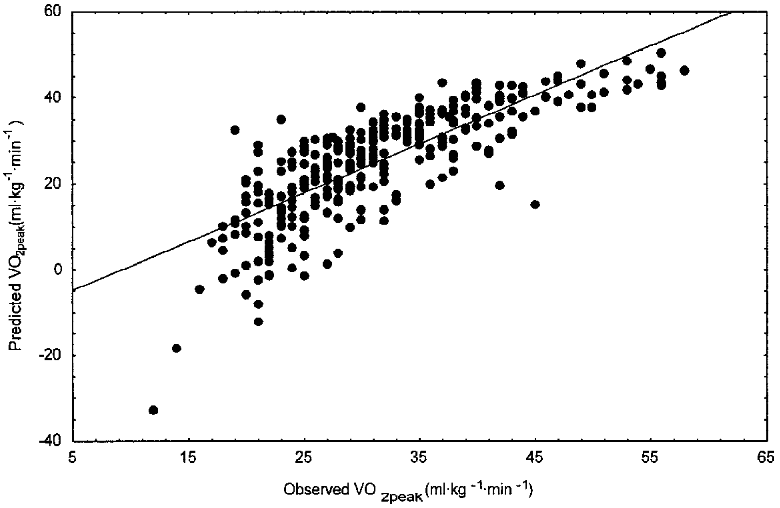

Figure 1. Observed versus predicted $\dot{V}O_{2peak}$ for the application of the Rockport general equation to the present sample.

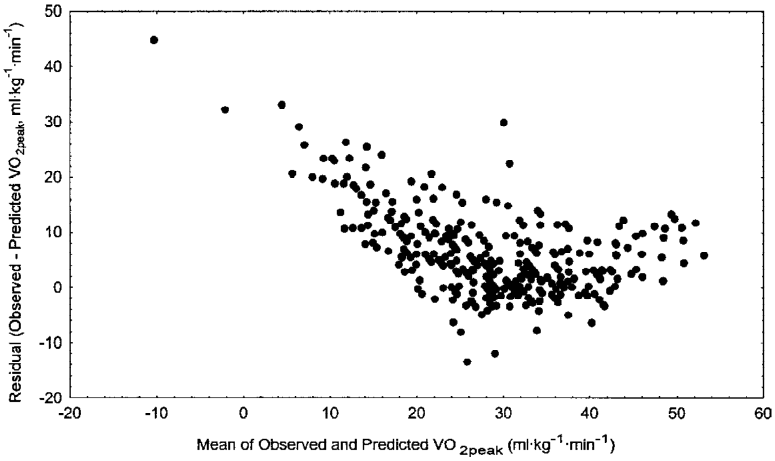

Figure 2. Bland-Altman plot of the residuals from the application of the Rockport general equation to the present sample.

PART 2

Table 2 shows means (± SD) and ranges for subject characteristics of the development and validation groups. Analysis revealed no significant differences in mean $\dot{V}O_{2peak}$, age, walk time, walk heart rate, or activity level. The development group had a significantly higher proportion of females than the validation group (0.57 vs. 0.43, $p = 0.015$). Mean body mass was significantly higher by approximately 7 kg in the cross-validation group than in the development group ($p < .0005$). The mean activity level of approximately 4 in all of the groups corresponds to a description of habitual physical activity as "participates regularly in heavy physical exercise such as running . . . swimming . . . cycling . . . tennis . . . basketball . . . for less than 30 minutes per week" (Matthews et al., 1999).

Stepwise regression on the development group from the present sample indicated that body mass, age, gender, walk time, walk heart rate, and activity level were all significant predictors of $\dot{V}O_{2peak}$, and yielded the following equation:

> $\dot{V}O_{2peak}$ TREADWALK = 92.08 – 0.10 (body weight in pounds) – 0.34 (age in years) + 9.72 (gender; male = 1, female = 0) – 1.01 (walk time in minutes and hundredths of a minute) – 0.13 (walk heart rate in bpm) + 0.86 (activity level)

Table 3 shows the multiple r, SEE, and p-value for inclusion each of the predictor variables in the regression model. Application of TREADWALK to the cross-validation group from the present sample yielded a correlation of r = 0.87 between observed and predicted $\dot{V}O_{2peak}$ values, SEE = 4.7 ml · kg^{-1} · min^{-1}, and TE = 4.8 ml · kg^{-1} · min^{-1}. Figure 3 shows observed versus predicted $\dot{V}O_{2peak}$ for the cross-validation group. Examination of the Bland- Altman (Figure 4) and normal probability plots of the TREADWALK residuals (observed – predicted) indicated

Table 2 Means (SD) and Ranges for $\dot{V}O_{2peak}$ and All Predictor Variables for the Development Sub-sample (m = 66, f = 88) and Cross-validation Sub-sample (m = 85, f = 65)

	Development group mean (SD)	Range	Cross-validation group mean (SD)	Range	p-value for difference
$\dot{V}O_{2peak}$ (ml · kg^{-1} · min^{-1})	31.2 (8.52)	12–58	32.0 (9.27)	14–56	0.40
Walk time (min)	17.5 (3.28)	12.1–35.2	17.0 (3.10)	11.1–27.4	0.19
Walk heart rate (bpm)	119.4 (17.48)	79–181	122.7 (19.01)	80–182	0.11
Age (yr)	57.3 (10.56)	40–79	57.4 (10.59)	40–78	0.95
Body weight (kg)	73.3 (14.97)	42.3–125.6	79.8 (15.84)	42.1–129.3	<0.0005
Activity score	4.4 (2.07)	0–7	4.3 (2.18)	0–7	0.67

Table 3 Multiple r, Standard Error of the Estimate (SEE) and p-Value for Inclusion of Each Predictor Variable, From Stepwise Regression

	r	SEE (ml · kg^{-1} · min^{-1})	p-value
Walk time	0.59	6.9	<0.0001
Walk HR	0.71	6.0	<0.0001
Age	0.75	5.7	<0.0001
Gender	0.80	5.1	<0.0001
Body mass	0.85	4.6	<0.0001
Activity level	0.87	4.3	<0.0001

that the mean residual was 0.96 ml · kg^{-1} · min^{-1}, and that the residuals were normally distributed.

Application of the PRESS method generated a correlation of $r = 0.85$ between observed and predicted $\dot{V}O_{2peak}$. The mean residual was 0.05 ml · kg^{-1} · min^{-1}, SEE = 4.9 ml · kg^{-1} · min^{-1}, and TE = 4.6 ml · kg^{-1} · min^{-1}. Table 4 shows mean residuals, correlations, SEE, and TE for each of the analyses in this investigation.

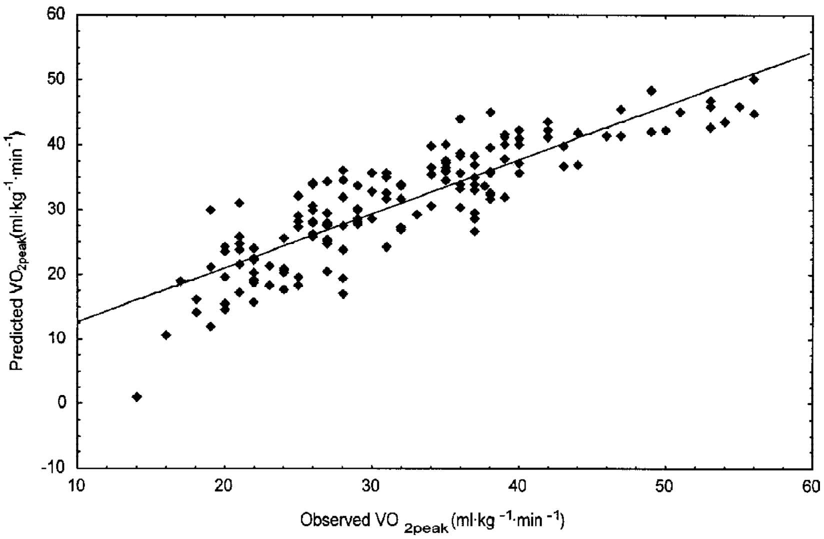

Figure 3. Observed versus predicted $\dot{V}O_{2peak}$ for the application of the Treadwalk equation to the cross-validation sub-sample.

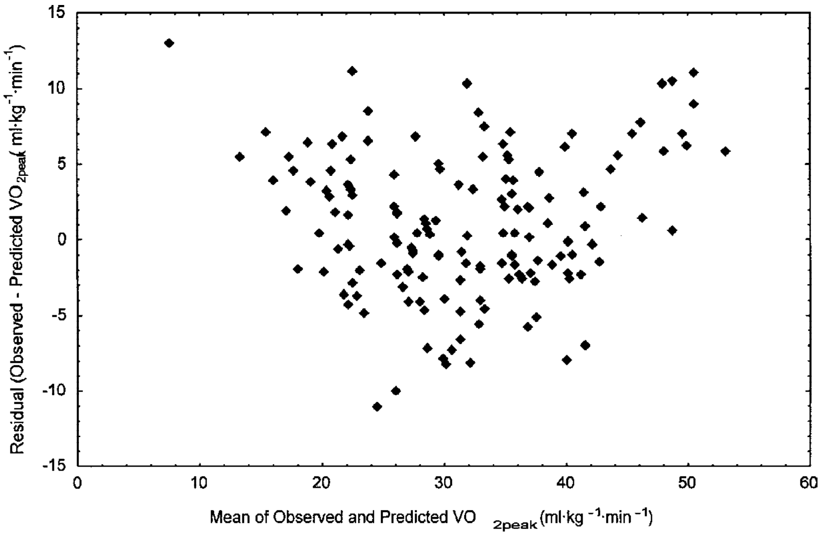

Figure 4. Bland-Altman plot of the residuals from the application of the Treadwalk equation to the cross-validation sub-sample.

Table 4 Correlation Coefficient (r), Mean Residual, Standard Error of the Estimate (SEE), and Total Error (TE)

	r	Mean of residuals ($ml \cdot kg^{-1} \cdot min^{-1}$)	SEE ($ml \cdot kg^{-1} \cdot min^{-1}$)	TE ($ml \cdot kg^{-1} \cdot min^{-1}$)
ROCK	0.80	6.20	7.7	9.9
TREAD	0.87	0.96	4.7	4.8
PRESS	0.85	0.05	4.9	4.6

Discussion

The correlation between observed and predicted $\dot{V}O_{2peak}$ of 0.80, and SEE of 7.7 $ml \cdot kg^{-1} \cdot min^{-1}$ produced by application of the Rockport one-mile over-ground walk test equation to treadmill walking in the first part of this investigation provide the appearance that this equation worked reasonably well to predict $\dot{V}O_{2peak}$. However, further analysis of the plots of observed vs. predicted $\dot{V}O_{2peak}$ and of the residuals indicates that the model is not appropriate. The Bland-Altman and normal probability plots of the residuals indicate that they are not normally distributed, indicating poor fit of the model to the data. Furthermore, the mean residual is 6.2 $ml \cdot kg^{-1} \cdot min^{-1}$. If the Rockport model were appropriate for the present data, it would be expected that the residuals would be normally distributed with a mean of approximately zero.

In contrast to the results found by Widrick et al. (1992), the results obtained in the present investigation indicate an overall under-prediction of $\dot{V}O_{2peak}$ when applying the Rockport general equation to treadmill walking. The high correlation between observed and predicted $\dot{V}O_{2peak}$ in Widrick et al. (1992) may be an artifact of the of the inclusion of several subjects with high $\dot{V}O_{2peak}$ values (~ 70 $ml \cdot kg^{-1} \cdot min^{-1}$) leading to a somewhat exaggerated spread in the data that would produce higher correlation. The total error generated in the Widrick et al. paper was also somewhat larger (approximately 15%, or 5.26 $ml \cdot kg^{-1} \cdot min^{-1}$ vs 4.8 $ml \cdot kg^{-1} \cdot min^{-1}$) than that found in the present investigation. This may provide additional indirect evidence that the Rockport model fit was inadequate in the investigation by Widrick et al. (1992).

The lack of acceptable results in assessing the Rockport equation in the present investigation is not particularly surprising given that the Rockport equation was developed for over-ground walking, while the present data were obtained during treadmill walking. The differences in physiological response to over-ground versus treadmill locomotion have been well documented, with over-ground locomotion generally having greater oxygen requirement at a given speed than treadmill walking (Bassett, et al., 1985; Jones and Doust, 1996; Pearce, et al., 1983). Based on these known differences, it seems reasonable that a prediction method needs to be specific to the exercise conditions (e.g., over-ground versus treadmill walking).

Using treadmill walking data to generate a $\dot{V}O_{2peak}$ prediction equation, the present study produced a correlation between observed and predicted $\dot{V}O_{2peak}$ that was higher ($r = 0.87$ vs. $r = 0.80$) and a SEE that was lower (4.7 ml · kg^{-1} · min^{-1} vs. 7.7 ml · kg^{-1} · min^{-1}) than those resulting from application of the Rockport equation. This indicates that the new equation developed in this study predicts $\dot{V}O_{2peak}$ more accurately than the Rockport equation when using treadmill walking. More importantly, the residuals produced by the new equation appeared to be normally distributed, which indicates that the new model is more appropriate for the present sample and testing methods. The results of the present investigation suggest that different prediction equations may be required when using different modes of walking to predict $\dot{V}O_{2peak}$.

The relatively high correlation between observed and predicted $\dot{V}O_{2peak}$, small SEE, and well behaved residuals indicate that the equation developed in this study is an appropriate model for predicting $\dot{V}O_{2peak}$ using treadmill walking in a sample of healthy middle-aged and older individuals. An extensive review of the submaximal exercise test literature reveals a wide range of correlation and SEE values reported for other tests. These studies generally only report one measure of the validity of their equation, and this measure is usually the result of the equation development study, rather than validation studies. The results obtained in the present investigation are similar to those reported in *validation* studies of other widely used submaximal tests such as the Åstrand-Rhyming ($r = 0.74$, SEE = 0.36 L · min^{-1}), Harvard Step ($r = 0.76$, SEE = 6.3 ml · kg^{-1} · min^{-1}) (DeVries and Klaf, 1965), and Rockport tests ($r = 0.92$, SEE = 4.97 ml · kg^{-1} · min^{-1}) (Kline et al., 1987). An advantage of the present test is that, like the Rockport test it relies on a mode of exercise that is familiar to, and well tolerated by most adults.

Another advantage of the present test over common tests used to assess functional capacity, such as the six- (Miyamoto et al., 2000) and nine-minute walk tests (Fleg et al., 2000) is that the present method obtains an estimate of $\dot{V}O_{2peak}$. In the clinical setting, $\dot{V}O_{2peak}$ remains the gold-standard for evaluating prognoses and assessing therapeutic interventions (Fleg et al., 2000). The present method relies on the familiar task of walking, reducing the possibility that lack of familiarity with the task may reduce the predictive accuracy of the test. An advantage of the present method over other walking tests is the use of the treadmill, which allows for more careful monitoring of subjects than over-ground walking tests.

A limitation of the present study is the use of a holdout cross-validation group drawn from the same sample as the equation development group was drawn from. Because the two groups were part of the same larger sample, it is not particularly surprising that the groups had similar characteristics, nor that the cross-validation produced strong results. Despite the shortcomings, it was more useful to have a cross-validation group drawn from the same sample as the development group than to have no validation at all. The large size of each group, however, provides some assurance that the results would be similar if the new equation were applied to an entirely different sample. Further assurance is provided by the results of the PRESS analysis, which are similar to those produced by the sample-splitting approach. The consistency of results between two different internal validation methods strengthens the conclusions drawn from the regression analysis.

An important consideration in the application of the present equation is that the validity of this method, as with any regression based prediction method, is based on its application to a group of subjects. The mean of the residuals produced by the TREADWALK equation is near zero, indicating accurate estimation of the sample's mean $\dot{V}O_{2peak}$. Conversely, based on the SEE generated in the present investigation, application to an individual would require a 95% confidence interval of nearly 10 ml · kg^{-1} · min^{-1} on either side of the predicted value. This is a large percentage potential error, particularly when dealing with individuals with low cardiorespiratory fitness, who are the most likely candidates for this type of test. As shown in Figure 3, the present equation may under-predict $\dot{V}O_{2peak}$ for individuals with $\dot{V}O_{2peak}$ values above 50 ml · kg^{-1} · min^{-1}, and may over-predict $\dot{V}O_{2peak}$ for individuals with values below 18 ml · kg^{-1} · min^{-1}. While the number of subjects in these categories in the present study is too small (13 of 150 subjects in the cross-validation group) to draw definite conclusions, this observation supports the caveat against using regression equations to predict individual results.

Figures 1 and 3 seem to suggest a non-linear relationship between observed and predicted $\dot{V}O_{2peak}$. However, the addition of model terms that were non-linear in the predictors, and the use of regression models that were non-linear in the coefficients produced at best only small improvements in r (0.015 to 0.02) and SEE (0.1 to 0.4 ml · kg^{-1} · min^{-1}). These modest improvements were gained at the expense of the addition of multiple exponential terms to the linear model, or the generation of a comparatively complex non-linear model. Because of this added complexity, it was determined that the linear model presented in this paper was more favorable than the non-linear options from the perspective of both statistical efficiency and ease of application of the model to practical situations.

Another issue that must be addressed with respect to the application of the present equation to individuals is the selection of predictor variables. While time to complete the one-mile walk and walk heart rate are the primary explanatory variables in the TREADWALK equation, age, gender, and body mass account for 21.8% of the variance in $\dot{V}O_{2peak}$ explained by the TREADWALK equation. The inclusion of predictor variables that do not have a strong physiological link to $\dot{V}O_{2peak}$ within individuals introduces potential bias in the equation that may further compound the prediction error associated with the application of this equation to individuals. On a population basis however, these variables are strongly associated with $\dot{V}O_{2peak}$, as exhibited by their predictive role in most submaximal exercise tests.

Further research on populations with characteristics similar to the present group, as well as on different populations, is necessary to strengthen the generalizability of these findings before the TREADWALK equation can be widely applied. Additionally, it would be useful to compare the results of treadmill and over-ground walking within one sample, to obtain a clearer picture of the differences in responses between the two modes of exercise. Though the large sample in the present investigation, and the use of much of the same staff to conduct the research reduces the likelihood that the observed differences between the Rockport and TREADWALK equations are due to the use of different subjects rather than differences in test modality, the results of the present study do not rule out that possibility. Finally, the present prediction method should be investigated to determine if it is sensitive to change that may occur with training or rehabilitation programs within groups or individuals.

In conclusion, the present investigation generated an equation that predicts $\dot{V}O_{2peak}$ with accuracy similar to other widely used submaximal exercise tests, using treadmill walking and several other quickly and easily acquired measures. This equation is applicable to healthy adults 40 to 79 years of age. Once validated in other study populations, and evaluated on its ability to detect changes in $\dot{V}O_{2peak}$ consequent to training or other intervention, the present equation has the potential to be a useful clinical and research tool to allow assessment of cardiorespiratory function with a minimum of time, expense, and risk.

References

Bassett, D.R., Giese, M.D., Nagle, F.J., Ward, A., Raab, D.M., and Balke, B. (1985). Aerobic requirements of overground versus treadmill running. **Med. Sci. Sports Exerc.** 17: 477-481.

Bland, J.M., and Altman, D.G. (1995). Comparing methods of measurement: Why plotting difference against standard method is misleading. **Lancet** 346: 1085-1087.

Cooper, K.H. (1968). A means of assessing maximal oxygen intake. Correlation between field and treadmill testing. **JAMA** 203: 201-204.

DeVries, H.A., and Klafs, C.E. (1965). Prediction of maximal O_2 intake from submaximal tests. **J. Sports Med**. 5: 207-214.

Doolittle, T.L., and Bigbee, R. (1968). The twelve-minute run-walk: A test of cardiorespiratory fitness of adolescent boys. **Res. Q. Exerc. Sport**. 39: 491-495.

Falls, H.B., Ismail, A.H., and MacLeod, D.F. (1966). Estimation of maximum oxygen uptake in adults from AAHPER youth fitness test items. **Res. Q. Exerc. Sport** 37: 192-201.

Fenstermaker, K.L., Plowman, S.A., and Looney, M.A. (1992). Validation of the Rockport Fitness Walking Test in females 65 years and older. **Res. Q. Exerc. Sport** 63: 322-327.

Fleg, J.L., Pina, I.L., Balady, G.J., Chaitman, B.R., Fletcher, B., Lavie, C., Limacher, M.C., Stein, R.A., Williams, M., and Bazzarre, T. (2000). Assessment of functional capacity in clinical and research applications: An advisory from the committee on exercise, rehabilitation, and prevention, council on clinical cardiology, american heart association. **Circulation** 102: 1591-1597.

Fox, E.L. (1973). A simple, accurate technique for predicting maximal aerobic power. **J. Appl. Physiol.** 35: 914-916.

George, J.D. (1996). Alternative approach to maximal exercise testing and $\dot{V}O_2$ max prediction in college students. **Res. Q. Exerc. Sport** 67: 452-457.

George, J.D., Fellingham, G.W., and Fisher, A.G. (1998). A modified version of the Rockport Fitness Walking Test for college men and women. **Res. Q. Exerc. Sport** 69: 205-209.

Heil, D.P., Freedson, P.S., Ahlquist, L.E., Price, J., and Rippe, J.M. (1995). Nonexercise regression models to estimate peak oxygen consumption. **Med. Sci. Sports Exerc.** 27: 599-606.

Hermiston, R.T., and Faulkner, J.A. (1971). Prediction of maximal oxygen uptake by a stepwise regression technique. **J. Appl. Physiol.** 30: 833-837.

Holiday, D.B., Ballard, J.E., and McKeown, B.C. (1995). PRESS-related statistics: Regression tools for cross-validation and case diagnostics. **Med. Sci. Sports Exerc.** 27: 612-620.

Jones, A.M., and Doust, J.H. (1996). A 1% treadmill grade most accurately reflects the energetic cost of outdoor running. **J. Sports Sci.** 14: 321-327.

Kline, G.M., Porcari, J.P., Hintermeister, R., Freedson, P.S., Ward, A., McCarron, R.F., Ross, J., and Rippe, J.M. (1987). Estimation of $\dot{V}O_2$max from a one-mile track walk, gender, age, and body weight. **Med. Sci. Sports Exerc.** 19: 253-259.

Lohman, T.G. (1981) Skinfolds and body density and their relationship to body fatness: A review. **Human Biology** 53: 181-225.

Mastropaolo, J.A. (1970). Prediction of maximal O_2 consumption in middle-aged men by multiple regression. **Med. Sci. Sports** 2: 124-127.

Matthews, C.E., Freedson, P.S., Hebert, J.R., Stanek, E.J., Merriam, P.A., and Ockene, I.S. (2000). Comparing physical activity assessment methods in the Seasonal Variation of Blood Cholesterol Study. **Med. Sci. Sports Exerc.** 32: 976-984.

McArdle, W.D., Katch, F.I., Pechar, G.S., Jacobson, L., and Ruck, S. (1972). Reliability and interrelationships between maximal oxygen intake, physical work capacity and step-test scores in college women. **Med. Sci. Sports** 4: 182-186.

McNaughton, L., Hall, P., and Cooley, D. (1998). Validation of several methods of estimating maximal oxygen uptake in young men. **Percept. Mot. Skills** 87: 575-584.

Miyamoto, S., Nagaya, N., Satoh, T., Kyotani, S., Sakamaki, F., Fujita, M., Nakanishi, N., and Miyatake, K. (2000). Clinical correlates and prognostic significance of six-minute walk test in patients with primary pulmonary hypertension. Comparison with cardiopulmonary exercise testing. **Am. J. Respir. Crit. Care Med.** 161: 487-492.

Pearce, M.E., Cunningham, D.A., Donner, A.P., Rechnitzer, P.A., Fullerton, G.M., and Howard, J.H. (1983). Energy cost of treadmill and floor walking at self-selected paces. **Eur. J. Appl. Physiol. Occup. Physiol.** 52: 115-119.

Powers, S.K., and Howley, E.T. (2001). **Exercise Physiology: Theory and Application to Fitness and Performance** (4th ed.). New York: McGraw Hill.

Rintala, P., McCubbin, J.A., Downs, S.B., and Fox, S.D. (1997). Cross validation of the 1-mile walking test for men with mental retardation. **Med. Sci. Sports Exerc.** 29: 133-137.

Widrick, J., Ward, A., Ebbeling, C., Clemente, E., and Rippe, J.M. (1992). Treadmill validation of an over-ground walking test to predict peak oxygen consumption. **Eur. J. Appl. Physiol. Occup. Physiol.** 64: 304-308.

Acknowledgments

The authors would like to acknowledge Christopher Palmer, Sue Damitz, Kim Demers, Keith Derusseau, Stacy Hess, Steve Jungbluth, and Janet Price, whose contributions to subject recruitment, testing, and data preparation made this research possible. This research was funded by a grant from the Whitehall-Robins Corporation.