Evaluation of a treadmill test for predicting the aerobic capacity of firefighters

Constance M. Mier and Ann L. Gibson

Background
As part of a comprehensive occupational medical program for fire departments in the USA, the National Fire Protection Association and The Fire Service Joint Labor Management Wellness/Fitness Initiative endorse a standardized submaximal test that uses the Gerkin treadmill protocol for predicting the maximal oxygen uptake ($V_{O2max}$) of firefighters.

Aims
To test the validity of the Gerkin treadmill protocol in healthy men and women.

Methods
Fifty-four healthy men and women (age range 19–58 years) performed the Gerkin test and a treadmill run test to maximal exhaustion. Their heart rates were monitored continuously with an electrocardiogram during each test. During the $V_{O2max}$ test, the subjects' $V_{O2}$ was measured continuously using indirect calorimetry.

Results
Although the predicted and observed $V_{O2max}$ values correlated ($r = 0.70$, $P < 0.001$ and standard error of estimate = 5.98 ml/kg/min), the mean values differed (49.8 ± 8.3 and 41.8 ± 5.8 ml/kg/min, respectively) ($P < 0.001$). The $V_{O2max}$ value was overestimated in 50 (93%) participants. The overestimation was >25% in 18 (33%) participants. Gender, age and $V_{O2max}$ did not affect the Gerkin protocol's predictability of $V_{O2max}$.

Conclusions
The Gerkin treadmill protocol overpredicts $V_{O2max}$ in healthy men and women and, therefore, should not be used for predicting $V_{O2max}$ in individual firefighters, particularly if $V_{O2max}$ is a criterion for inclusion or exclusion from duty. At this time, a valid treadmill running test is needed for predicting the $V_{O2max}$ value of individual firefighters.

Key words
Gerkin treadmill protocol; maximal heart rate; submaximal treadmill test; $V_{O2max}$ estimation.

Introduction
Firefighting is an aerobically demanding occupation sometimes requiring a firefighter to work at or near maximal capacity for several minutes at a time [1–6]. Because of the high physical demands of firefighting, successful job performance and minimization of injury depends largely on the fitness level of firefighters [1,2]. In addition, the combination of the high physical demands of the job, abrupt changes from rest to high intensities, a number of environmental and psychological stressors and use of protective garments contribute to the increased mortality from cardiovascular disease among firefighters [7,8]. In order to address these health and safety issues, The Fire Service Joint Labor Management Wellness/Fitness Initiative was developed [9]. This initiative recommends a maximal oxygen uptake ($V_{O2max}$) of at least 42 ml/kg/min in order to meet the aerobic demands of the job adequately. This minimal value is believed to allow a sufficient reserve capacity that offers a margin of safety for performing necessary firefighting duties [1,4,10].

A number of fire departments are following the
recommendations of the initiative and have established a minimum $V\dot{O}_{2\text{max}}$ for their firefighters. Most of the time, $V\dot{O}_{2\text{max}}$ is not measured directly; rather, field or laboratory submaximal tests are used for predicting $V\dot{O}_{2\text{max}}$ based on heart rate response. The National Fire Protection Association and The Fire Service Joint Labor Management Wellness/Fitness Initiative endorse a standardized submaximal test that uses the Gerkin treadmill protocol [11]. The use of this test is extremely beneficial in a clinical or commercial setting. First, it is relatively easy to administer since it does not require sophisticated testing equipment other than a treadmill and a heart rate monitor. Secondly, because it is a submaximal test, there is a lower risk of cardiac complications in those individuals who may be at risk for heart disease. Thirdly, it is considerably less expensive than the tests that require direct measurement of oxygen uptake.

Because of its simplicity and ease of use, the Gerkin treadmill protocol has become a popular mode of assessing aerobic capacity in firefighters. The Gerkin treadmill protocol incorporates stages of 1 min during which a subject’s heart rate is recorded until 85% of the age-predicted maximal heart rate ($HR_{\text{max}}$) is reached. The $V\dot{O}_{2\text{max}}$ value is then predicted from the estimated $V\dot{O}_2$ corresponding to the final stage using the American College of Sports Medicine (ACSM) metabolic equation for running [12]. If a cut-off value for $V\dot{O}_{2\text{max}}$ is to be endorsed for hiring or maintaining the active duty status of individual firefighters, it is critical that a valid test be used. Therefore, we sought to test the validity of the Gerkin treadmill protocol for estimating $V\dot{O}_{2\text{max}}$. We hypothesized that the test would overpredict the $V\dot{O}_{2\text{max}}$ value in healthy men and women owing to its reliance on the ACSM metabolic equation for $V\dot{O}_{2\text{max}}$ estimation from non-steady-state test results.

Methods

Participants

Sixty healthy individuals (men = 37 and women = 23) volunteered to participate in this study. Each participant read and signed an informed consent approved by the Barry University Institutional Review Board. The participants were recruited from Miami–Dade County and Fort Lauderdale fire departments as well as the Barry University community. Each participant completed a health questionnaire and, from the information provided, was stratified as low, medium or high risk for a cardiovascular event during exercise [12]. One male participant reported having signs or symptoms of cardiovascular disease and, therefore, he was classified as high risk and consequently excluded from the study. None of the remaining participants reported taking medications that might interfere with the physiological responses to exercise and none reported having any physical limitation that might preclude him or her from maximal exercise. Five of the 59 participants did not complete the tests for personal reasons. Therefore, data from 54 participants (men = 31 and women = 23) are presented (Table 1).

Maximal testing

Each participant performed a maximal graded treadmill test while their $V\dot{O}_{2\text{max}}$ was measured with a SensorMedics 2900 (Loma Linda, CA) metabolic cart. The flowmeter was calibrated prior to each test using a standard 3 l syringe and two-point calibrations for oxygen and carbon dioxide were performed using standard gases. The individualized running speed was constant throughout the test while the grade was increased by 2% every 2 min. The running speed was adjusted so that the first stage (0%) elicited a rating of perceived exertion of 12–13 (somewhat hard) on the Borg scale. A warm-up period of 3 min followed resting heart rate measurements. The subjects’ heart rate was continuously monitored during the test using an electrocardiogram. The test was considered maximal if $V\dot{O}_{2\text{max}}$ increased by <150 ml/min or <2.1 ml/kg/min with an increase in work rate or if a maximal respiratory exchange ratio ($RER_{\text{max}}$) of ≥1.15 and a $HR_{\text{max}}$ greater than or equal to the age-predicted maximal value were achieved.

Gerkin treadmill protocol

The test began with a warm-up period of 3 min at a walking speed of 3.5 mph. Following the warm-up, the treadmill speed was increased to a running pace of 4.5 mph. The intensity was adjusted each minute by alternating an increase of 0.5 mph with a 2% grade increase. Each 1 min stage consisted of four intervals of 15 s. The test was terminated at the end of the 15 s interval immediately following the first 15 s interval during which a heart rate greater than 85% of the age-predicted $HR_{\text{max}}$ was achieved. The treadmill time completed was expressed by identifying the four 15 s intervals as 0.1, 0.2, 0.3 and 0.4, respectively. For instance, if a test was terminated at the 30 s mark during stage 3, the treadmill time completed was expressed as 3.2.

<table>
<thead>
<tr>
<th>Table 1. Descriptive characteristics of the participants</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Age (years)</td>
</tr>
<tr>
<td>Body mass (kg)</td>
</tr>
<tr>
<td>Height (cm)</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
</tr>
<tr>
<td>Resting heart rate (beats/min)</td>
</tr>
</tbody>
</table>
Prediction of $\dot{V}O_{2\text{max}}$

Previously, a prediction equation was developed using the Gerkin treadmill protocol [11]. The $\dot{V}O_2$ corresponding to each stage was calculated using the ACSM metabolic equation for running [12]. $\dot{V}O_{2\text{max}}$ was estimated from this equation from the last stage of the Gerkin protocol performed to maximum. The following equation was derived using the estimated $\dot{V}O_{2\text{max}}$ for the stage corresponding to 85% of the age-predicted HR$_{\text{max}}$: $\dot{V}O_{2\text{max}} = 1.39(\dot{V}O_2$ at 85% of age-predicted HR$_{\text{max}}$).

Statistical analyses

Significant differences between the predicted and observed maximal values ($\dot{V}O_{2\text{max}}$ and HR$_{\text{max}}$) were tested using a dependent $t$-test. Pearson correlation coefficients were calculated in order to show the relationship between the predicted and observed values. The standard error of the estimate (SEE) was calculated for each. The SEE was also expressed as a percentage of the mean (SEE/mean × 100). Group effects of gender, age and $\dot{V}O_{2\text{max}}$ on the mean differences between the predicted and observed $\dot{V}O_{2\text{max}}$ were determined by an independent $t$-test. For the age and $\dot{V}O_{2\text{max}}$ groups, each participant was categorized as either falling at or below the mean or falling above the mean. The data are presented as means ± SD. SPSS version 11.0 was used for analyzing the data.

Results

Fourteen participants did not achieve the criteria for maximal effort during the maximal treadmill test. In order to determine whether their data should be removed from further analyses, we compared their results to the results of those who did achieve the criteria. Except for having a greater body mass (83.8 ± 13.1 versus 74.3 ± 13.4 kg) ($P = 0.031$) and height (178.7 ± 8.2 versus 172.4 ± 10.4 cm) ($P = 0.029$) and a lower RER$_{\text{max}}$ (1.16 ± 0.06 versus 1.21 ± 0.04) ($P = 0.01$), the group not achieving the maximal effort criteria did not differ in their observed $\dot{V}O_{2\text{max}}$ (41.5 ± 5.7 versus 41.8 ± 5.9 ml/kg/min) ($P = 0.87$), predicted $\dot{V}O_{2\text{max}}$ (50.8 ± 9.3 versus 49.5 ± 8.1 ml/kg/min) ($P = 0.66$), observed HR$_{\text{max}}$ (187 ± 8 versus 193 ± 13 beats/min) ($P = 0.12$) or predicted HR$_{\text{max}}$ (183 ± 11 versus 185 ± 11 beats/min) ($P = 0.71$). Therefore, subsequent data analyses included all 54 participants.

The results from the maximal and Gerkin treadmill tests are presented in Tables 2 and 3, respectively. Despite these relationships, the predicted $\dot{V}O_{2\text{max}}$ was greater than the observed $\dot{V}O_{2\text{max}}$, with a mean difference of 8.1 ± 5.9 ml/kg/min ($t_{53} = 10.01$ and $P < 0.001$). The predicted HR$_{\text{max}}$ was less than the observed HR$_{\text{max}}$ with a mean difference of −6.5 ± 9.4 beats/min ($t_{53} = −5.1$ and $P < 0.001$). The $\dot{V}O_{2\text{max}}$ value was overpredicted in 50 of the participants. Further, the $\dot{V}O_{2\text{max}}$ value was overpredicted by 10–25% in 21 (39%) participants and greater than 25% in 18 (33%) participants. The HR$_{\text{max}}$ value was underpredicted in 41 (76%) participants. Further, the HR$_{\text{max}}$ value was underpredicted by 1–9 beats/min in 23 (43%) participants, 10–20 beats/min in 14 (26%) participants and >20 beats/min in four (7%) participants.

The group effects on the mean differences between the predicted and observed values are presented in Table 4. No differences were found between men and women. When age groups were compared, the mean differences between the predicted and observed HR$_{\text{max}}$ were greater in the older group, meaning that the observed HR$_{\text{max}}$ was underpredicted to a greater extent than for the younger group. In addition, age correlated with the mean difference between the predicted and observed HR$_{\text{max}}$ ($r = −0.33$ and $P = 0.016$). There was no age effect on the mean differences between the predicted and observed $\dot{V}O_{2\text{max}}$ values. No differences between the $\dot{V}O_{2\text{max}}$ groups were evident.

Discussion

We found that the Gerkin treadmill protocol for predicting $\dot{V}O_{2\text{max}}$ in healthy men and women greatly overestimated their aerobic capacity. We see several errors with this protocol. The Gerkin treadmill protocol relies
on 85% of the age-predicted HR\textsubscript{max} value as the primary criterion for ending the test. Predicting HR\textsubscript{max} by subtracting age from 220 carries significant inter-individual variability, which has been reported to be 10–12 beats/min [12]. The traditional equation underestimates HR\textsubscript{max} in older individuals and tends to overestimate HR\textsubscript{max} in younger individuals [13,14]. In support of these observations, we found age to be related to the difference between the predicted and observed HR\textsubscript{max} values.

When the final heart rate measured during the Gerkin test was expressed as a percentage of actual HR\textsubscript{max} we observed values ranging from 75 to 91%, indicative of the large error present when using a percentage of the age-predicted HR\textsubscript{max} as a stop-test criterion. The percentage of actual HR\textsubscript{max} achieved during the Gerkin test correlated with the difference between the predicted and observed VO\textsubscript{2max} \((r = 0.46\) and \(P < 0.001\)) such that the VO\textsubscript{2max} value of those individuals reaching >85% of HR\textsubscript{max} tended to be overestimated to the greatest extent. This makes sense because those individuals had a smaller reserve capacity at the end of the Gerkin test than that which would be predicted from their heart rate responses. These data clearly demonstrate the error in using the traditional calculation for HR\textsubscript{max} for predicting VO\textsubscript{2max}.

Another problem with the Gerkin treadmill protocol is that it relies on the ACSM metabolic equation for running for predicting VO\textsubscript{2max}. It is well known that this equation is based on steady-state measures and that, if used to calculate VO\textsubscript{2max} from a maximal exercise stage, VO\textsubscript{2max} will be overestimated [12]. This is largely due to the unaccounted contribution of anaerobic energy to the maximal workload. Our results are similar to those of Foster et al. [15], who found a 17.4% overestimation of VO\textsubscript{2max} among cardiac patients who performed a graded treadmill test using increments of 1 min to the maximum and the estimation of VO\textsubscript{2} at maximal intensity from the ACSM metabolic equation. Therefore, significant error is introduced in the prediction formula devised by Gerkin et al. [11], which relies on an estimated steady-state VO\textsubscript{2} during the maximal running stage as the dependent variable.

The error is confounded further by using 85% of the age-predicted HR\textsubscript{max} during submaximal stages of 1 min as the independent variable. The steady state is generally not reached until \(\sim 2–3\) min after the onset of exercise in most individuals [16]. Chuang et al. [17] measured heart rates beat by beat, beginning with the onset of exercise through to steady state at intensities above and below the lactate threshold. Approximately 80% of the steady-state heart rate was reached within 1 min at both intensities. Since we did not measure heart rates under steady-state conditions during the Gerkin test, we cannot be certain as

![Figure 1](image1.png)

**Figure 1.** Scattergram of the predicted and observed VO\textsubscript{2max} and HR\textsubscript{max}.

<table>
<thead>
<tr>
<th>Gender</th>
<th>Age (years)</th>
<th>VO\textsubscript{2max} (ml/kg/min)</th>
<th>HR\textsubscript{max} (beats/min)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Women ((n = 23))</td>
<td>&lt;35 ((n = 28))</td>
<td>7.8 ± 6.3</td>
<td>–6.4 ± 10.0</td>
</tr>
<tr>
<td>Men ((n = 31))</td>
<td>≥36 ((n = 26))</td>
<td>8.3 ± 5.8</td>
<td>–6.7 ± 9.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6.9 ± 5.7</td>
<td>–3.9 ± 8.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>7.5 ± 6.1</td>
<td>–9.4 ± 9.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8.7 ± 5.7</td>
<td>–7.2 ± 9.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>≥41.8 ((n = 25))</td>
<td>–5.8 ± 9.0</td>
</tr>
</tbody>
</table>
to how close to the steady state each individual was at the end of 1 min. However, it is safe to surmise that most if not all of our participants had not reached a steady-state heart rate when the final heart rate was recorded. This could explain some of the overestimation of VO2max.

Other possible sources of error exist. First, there is some intra-individual variability in heart rates measured during submaximal exercise intensities [16]. Heart rates can be affected by several factors such as body temperature, hydration state, anxiety or stress, medications, etc. We tested the reliability of the Gerkin treadmill protocol in 18 participants and found that the predicted VO2max did not differ between trials (44.8 ± 6.8 versus 44.8 ± 6.7 ml/kg/min) (P = 0.96), and that there was a strong relationship between trials 1 and 2 (r = 0.94 and SEE = 2.42 ml/kg/min). Therefore, we can be reasonably confident that the Gerkin test results were reliable. Another potential source of error falls within the ACSM metabolic equation for running. The inter-individual variability in VO2 measured at a given speed and grade can be as high as 15% [18].

The ease by which the Gerkin treadmill protocol can be performed makes it attractive for exercise physiologists or medical personnel for predicting the VO2max value in firefighters. It should be noted here that the Gerkin treadmill protocol was originally tested with only male participants and does not account for age or body weight in its prediction of VO2max. Prediction formulas for VO2max that are derived from several other submaximal treadmill walking or running protocols exist [19–23]. When the SEE is expressed as a percentage of the mean, the reported errors among these studies ranged from 6.3 to 11.5%. For best comparison to our study, only one [19] prediction formula was developed from a wide age range (20–59 years) of men and women, which would be more representative of firefighters. This study resulted in a prediction formula that included gender, age and steady-state heart rate as predictors. The reported correlation was \( R = 0.96 \) and \( \text{SEE} = 4.85 \text{ml/kg/min} \) (11.5% error).

Most relevant to the use of a submaximal protocol for predicting aerobic capacity in firefighters specifically is a comparison of the Gerkin protocol with the Chester step test (CST), which is commonly used in the UK (K. Sykes, personal communication). Through personal correspondence with Dr Kevin Sykes, we were able to compare our results to that of a validation study recently conducted in Dr Sykes’s laboratory and discovered that the CST has a higher prediction power than the Gerkin protocol. We suspect that this is because the CST allows for steady-state measures during the submaximal stages.

In its present form, we do not recommend the Gerkin treadmill protocol for predicting VO2max values in individual firefighters, particularly if a cut-off value is incorporated. We looked at the possibility of finding a more valid prediction formula based on our results from the Gerkin test. When we entered potential prediction variables into a stepwise regression (probability of \( F \) to enter \( \leq 0.05 \)), we found a lower SEE using three predictor variables: stages performed during the Gerkin test, gender and body weight (adjusted \( R^2 = 0.66 \), \( \text{SEE} = 3.36 \text{ml/kg/min} \) and \( \% \text{SEE} = 8.0\% \)). The formula derived was \( \text{VO2max (ml/kg/min) = } 35.3 + \text{treadmill stage(1.88)} + \text{gender(7.399)} – \text{body mass (0.194)} \), where gender = 1 for female and gender = 2 for male and body mass is expressed in kilograms.

Without cross-validating this equation, we cannot be certain that this formula will offer a valid means for predicting VO2max values among the firefighter population. The next step will be to cross-validate the equation among a group of firefighters.

The better alternative to a submaximal treadmill test for predicting VO2max values is a graded exercise test performed to maximal intensity. Compared with graded submaximal tests, graded maximal tests result in higher predictive accuracy ranging from 0.93 to 0.96 with the SEE ranging from 1.95 to 3.12 ml/kg/min [24–26]. The Gerkin treadmill protocol taken to maximal intensity may offer a more accurate prediction formula, particularly when taking into account body weight and gender.

In summary, the Gerkin treadmill protocol significantly overpredicted VO2max in healthy men and women ranging in age from 19 to 58 years and, therefore, should not be used for predicting VO2max values in individual firefighters, particularly if VO2max is a criterion for inclusion or exclusion from duty. At this time, a valid treadmill running test is needed for the firefighter population.

Acknowledgements

We thank Patricia Gesch of the Firefighters Health and Wellness Center of Miami for her assistance. We would particularly like to thank all the participants and the firefighters of Fort Lauderdale Fire Rescue and Miami–Dade County.

References


