Accuracy of $\dot{V}O_{2max}$ Prediction Equations in Older Adults

MATTHEW J. PETERSON[1,2], CARL F. PIEPER[1–4], and MIRIAM C. MOREY[1–3]

[1]*Geriatric Research, Education and Clinical Center, Department of Veterans Affairs Medical Center, Durham, NC;* [2]*Claude D. Pepper Older Americans Independence Center/Center on Aging and Human Development,* [3]*Department of Medicine, and* [4]*Department of Biometry and Bioinformatics, Duke University Medical Center, Durham, NC*

ABSTRACT

PETERSON, M. J., C. F. PIEPER, and M. C. MOREY. Accuracy of $\dot{V}O_{2max}$ Prediction Equations in Older Adults. *Med. Sci. Sports Exerc.*, Vol. 35, No. 1, pp. 145–149, 2003. **Purpose:** We explored the accuracy and bias of prediction equations (ACSM and Foster) in older, deconditioned men and women. We also examined the predictors of $\dot{V}O_{2max}$ to further understand which variables affect respiratory fitness in the elderly. **Methods:** One hundred seventy-one community dwelling, men (72.6 ± 4.8 yr) and women (71.0 ± 5.1 yr) screened in a clinical trial were retrospectively examined. $\dot{V}O_{2max}$ was measured using a standardized protocol with gas exchange measured. Measured $\dot{V}O_{2max}$ values were compared with prediction equations via mean difference analyses, and bias was explored using Bland-Altman analyses. Regression analysis determined significant predictors of measured $\dot{V}O_{2max}$. Alpha was $P \leq 0.05$. **Results:** In men and women, Foster, 21.6 ± 5.7 and 17.9 ± 5.1 $mL \cdot kg^{-1} \cdot min^{-1}$, respectively, was not significantly different from measured $\dot{V}O_{2max}$, 21.7 ± 4.8 and 17.3 ± 4.0, respectively. ACSM overestimated $\dot{V}O_{2max}$ in men and women, 26.3 ± 8.2 and 20.9 ± 7.3, respectively. By using Bland-Altman plots, ACSM showed significant overestimation bias in more fit women (r = 0.29), whereas Foster showed no estimation bias in either gender. Significant predictors of $\dot{V}O_{2max}$ were gender, BMI, age, treadmill grade, and speed, with an equation R^2 of 0.70. A measure of current activity levels did not make it into the final model ($P = 0.0505$) but is worthy of future consideration using more sensitive measures than ours. **Conclusion:** ACSM is not appropriate for use when treadmill testing older adults. We believe the Foster equation's $\dot{V}O_{2max}$ prediction accuracy is acceptable, showing no bias along a continuum of aerobic capacity. **Key Words:** EXERCISE TESTING, AGING, AEROBIC CAPACITY, OXYGEN CONSUMPTION, AGING FACTORS

Exercise testing via treadmill ergometry is the most widely used type of noninvasive diagnostic, prognostic, and functional cardiovascular testing (1,13). A recent National Ambulatory Medical Care Survey study showed an average of 3.1 million exercise stress tests ordered per year, with patients 65 yr and older accounting for approximately 1.25 million tests (6). In clinical settings, estimation of maximal or peak oxygen uptake levels is the norm (10,18). Maximal achieved workload is most often converted via estimation to metabolic equivalents (METs) or $\dot{V}O_{2max}$. To truly measure $\dot{V}O_{2max}$, indirect spirometry with measurement of expired gases is necessary. However, equipment costs and staff training limit direct measurement mainly to research and few clinical settings.

In lieu of direct measurement, $\dot{V}O_{2max}$ is estimated using prediction equations. The ACSM walking equation (ACSM) (1) is often used in clinical settings. However, this equation was developed for use when performing steady-state exercise. This implies use of this equation for prediction of $\dot{V}O_{2max}$ in a clinical stress test setting is inappropriate, due to the increases in workload irrespective of steady-state oxygen consumption. Subsequent equations have been developed in an attempt to more accurately predict $\dot{V}O_{2max}$ during stress testing (2,3,5,9,11,14). Berry and coworkers (3) developed and validated a $\dot{V}O_{2max}$ prediction equation for older adults with diagnosed osteoarthritis of the knee and cardiovascular disease. Foster and coworkers (9) also developed a $\dot{V}O_{2max}$ prediction equation (Foster) using younger (mean age 44.6 ± 12.5 yr) subjects with handrail support during testing. Foster's equation was developed independent of treadmill protocol to allow the test technician more freedom in selecting and continuing a test, independent of treadmill grade and/or speed (9).

Based on the large reported numbers of older adults being referred for stress tests, it is important that the medical community have validated estimates of $\dot{V}O_{2max}$ values to help properly guide medical decisions. We found only one study (3) in which $\dot{V}O_{2max}$ prediction equations were developed specifically for older adults (67.7 ± 10.2 yr), and this equation was validated on a considerably younger age group (60.4 ± 9.4). Previous studies have demonstrated age-related changes in the physiological responses to maximal exercise, including reduced heart rate and stroke volume and increased blood pressure (17). Because these changes in exercise hemodynamics have a direct impact on

Address for correspondence: Matthew J. Peterson, Durham VA Medical Center, GRECC (182), 508 Fulton St., Durham, NC 27705; E-mail: peter076@mc.duke.edu.

Submitted for publication April 2002.
Accepted for publication September 2002.

0195-9131/03/3501-0145/$3.00/0
MEDICINE & SCIENCE IN SPORTS & EXERCISE®

DOI: 10.1249/01.MSS.0000043547.22724.0B

TABLE 1. Selected equations used to predict $\dot{V}O_{2max}$.

Author	Protocol	Equation
ACSM-derived from several sources	N/A	$\dot{V}O_2 = (0.1 \cdot speed) + (1.8 \cdot speed \cdot grade) + 3.5$
Foster et al. (9)	Handrail-supported	$\dot{V}O_2 = 0.694 \cdot (\text{*ACSM Predicted}) + 3.33$

* ACSM predicted is calculated using the ACSM equation (equation 1).

O_2 kinetics, we hypothesize that current $\dot{V}O_{2max}$ prediction equations may not generalize to older adults. Therefore, the aims of this paper are to explore the accuracy and bias of the ACSM and Foster treadmill prediction equations (Table 1), which are independent of test protocol, in a sample of older, deconditioned men and women. Furthermore, we will examine the predictors of $\dot{V}O_{2max}$ in this cohort, to further understand which variables affect respiratory fitness in older adults.

TABLE 2. Pepper protocol.

Stage	Minute	MPH	Grade (%)
1	0–1	1.5	0
2	1–2	1.5	2
3	2–3	1.5	4
4	3–4	2.0	3
5	4–5	2.0	5
6	5–6	2.0	6
7	6–7	2.0	7
8	7–8	2.0	8
9	8–9	2.0	9
10	9–10	2.5	8
11	10–11	2.5	10
12	11–12	2.5	12
13	12–13	2.5	14
14	13–14	3.0	12
15	14–15	3.0	13
16	15–16	3.0	14
17	16–17	3.3	14
18	17–18	3.3	16
19	18–19	3.3	18
20	19–20	3.3	20

METHODS

Subjects. Retrospective data analyses were performed on data collected by Morey and coworkers (12) under the auspices of the Durham VA Medical Center Geriatrics Research, Education, and Clinical Center and the Duke University Medical Center, Claude D. Pepper Older Americans Independence Center. Participants ($N = 171$) included individuals, ranging in ages from 65 to 90, screened for participation in a randomized clinical trial. Participants included in the present study were not regular exercisers and free from unstable symptoms of cardiovascular disease. Specific exclusion criteria and screening processes for the clinical trial have been previously published (12). The protocol for the clinical trial was reviewed and approved annually by the Duke and VA institutional review boards, and informed written consent was obtained from each participant before data collection.

Treadmill testing. $\dot{V}O_{2max}$ was assessed via a symptom-limited maximal exercise tolerance test. Testing was performed in a fasting or near-fasting state on a Marquette (Marquette, WI) Series 1000 treadmill, with expired gases measured via a SensorMedics (Yorba Linda, CA) 3200 breath-by-breath gas-analysis unit with continuous monitoring of ventilation, carbon-dioxide production, and oxygen consumption. The standardized protocol employed (Pepper) was developed at the Duke Aging Center laboratory specifically for older adults. The Pepper protocol is shown in Table 2. The American College of Sports Medicine indications for test termination was followed during treadmill testing of all participants (1). By using 15-s averaging, $\dot{V}O_{2max}$ was obtained from the highest nonaberrant value recorded at the peak workload.

Current activity levels determination. Current activity levels were ascertained using the Reuben questionnaire (16). This is a 4-point questionnaire, which asks duration and frequency questions related to physical activity, specifically walking. For the purposes of determining the predictive association of any recreational walking at all on $\dot{V}O_{2max}$, the patients were dichotomized into two groups based on questionnaire responses: 1) No, not currently walking at all; or 2) Yes, walking irregularly or regularly.

Data analysis. Statistical analyses included participant physical characteristics (mean ± SD), Pearson product moment correlations to determine relationships between predicted and measured values, and dependent *t*-tests to determine significant differences between predicted and measured $\dot{V}O_{2max}$ means. Mean differences, SD, and 95th percentile confidence intervals (95%CI) were also calculated. To examine prediction bias, Bland-Altman graphs (4) were used with measured $\dot{V}O_{2max}$ plotted on the y-axis and the difference between measured and predicted $\dot{V}O_{2max}$ plotted on the x-axis. Multivariate stepwise regression analysis was employed to determine significant mechanical (i.e., speed, grade, time) and physiological (i.e., gender, age, BMI, current activity levels) predictors of measured $\dot{V}O_{2max}$. Because of the relatively small size of this cohort, development and cross-validation of a $\dot{V}O_{2max}$ prediction equation was not performed. However, whether or not physiological variables, in addition to mechanical test variables, may add to an equation's predictive strength is of importance and needs investigation. Level of significance was set at $P \leq 0.05$. All analyses were performed using SAS System software version 8e. (SAS Institute, Cary, NC)

RESULTS

Participant physical characteristics and group treadmill test results are displayed in Table 3. Significant differences between men and women were seen in resting heart rate, test time, peak grade, peak speed, and measured $\dot{V}O_{2max}$; therefore, all subsequent analyses were performed separately. Correlations between predicted and measured $\dot{V}O_{2max}$ were strong in both sexes, with coefficients ranging from 0.70 to 0.75 ($P < 0.05$). Table 4 shows mean differences between measured and predicted $\dot{V}O_{2max}$, along with the SD and 95%CI. The Foster equation was not significantly different from measured $\dot{V}O_{2max}$ in men or women as evidenced by 95%CI inclusive of 0. Whereas, the ACSM equation had mean differences that were significantly different from the

TABLE 3. Physical and treadmill test characteristics of men and women used for analyses.

Variable	Men	Women
N	59	114
Age (yr)	72.5 ± 4.9	71.1 ± 5.1
Height (cm)	172.3 ± 5.8	161.3 ± 5.6*
Weight (kg)	83.1 ± 9.7	73.0 ± 12.9*
BMI	27.9 ± 2.9	28.0 ± 4.6
Resting BP_{sys} (mm Hg)	144.6 ± 24.4	143.9 ± 23.5
Resting BP_{dias}	80.5 ± 12.6	81.6 ± 10.3
HR_{rest} (BPM)	75.5 ± 11.1	81.9 ± 12.7*
Max. BP_{sys}	185.5 ± 24.6	180.6 ± 24.1
Max. BP_{dias}	83.8 ± 15.1	82.5 ± 13.2
Test time (min)	12.5 ± 4.7	10.0 ± 4.1*
Max. grade (% incline)	12.8 ± 8.5	9.8 ± 6.8*
Max. speed (mph)	2.6 ± .5	2.3 ± .5*
Measured $\dot{V}O_{2max}$ (mL·kg^{-1}·min^{-1})	21.7 ± 4.8	17.3 ± 4.0*
ACSM $\dot{V}O_{2max}$	26.3 ± 8.2†	20.9 ± 7.3*†
Foster $\dot{V}O_{2max}$	21.6 ± 5.7	17.9 ± 5.1*

Mean ± SD.
* Significantly different from men ($P \leq 0.05$).
† Significantly different from gender-specific measured $\dot{V}O_{2max}$ ($P \leq 0.05$).

TABLE 4. Mean differences in measured and predicted $\dot{V}O_{2max}$ (predicted-measured), SD, and 95th percentile confidence intervals.

Equation	Mean Difference mL·kg^{-1}·min^{-1}	SD	95% CI
Men			
ACSM	*4.5	6.0	2.9, 6.1
Foster	−0.2	4.1	−1.3, 0.9
Women			
ACSM	*3.7	5.1	2.7, 4.6
Foster	0.6	3.4	−0.04, 1.2

* Mean difference significant at $P \leq 0.05$.

measured mean value in both genders ($P < 0.05$). Foster's mean difference SD was approximately 2 mL·kg^{-1}·min^{-1} less than ACSM in both genders. The bias for each prediction equation can be seen in Bland-Altman graphs (4) in Figures 1 and 2. Gender-specific trendlines are presented, and r-values are presented when significant ($P < 0.05$). The ACSM equation (Fig. 1) displayed a trend to overestimate $\dot{V}O_{2max}$ with greater discrepancy as fitness levels increased, whereas the Foster equation (Fig. 2) showed no significant trend toward underestimating or overestimating $\dot{V}O_{2max}$ in men or women across fitness levels.

Predictors of $\dot{V}O_{2max}$. Regression analyses were performed to determine significant predictors of $\dot{V}O_{2max}$. Variables entered into the analyses included: physiological variables—age, gender, BMI, and current activity levels; and mechanical variables—test time, treadmill peak grade, and treadmill peak speed. Those variables that were significant independent predictors, or those close to significance, are listed in Table 5. All physiological variables entered in the analyses were significant predictors of $\dot{V}O_{2max}$, with the exception of the variable current activity levels, which was almost significant ($P = 0.0505$). Of the mechanical variables entered, treadmill grade and speed were significant, with test time not meeting the requirements for entry in the equation. The inclusion of physiological variables increased the explained variance from 50%, using mechanical variables only, to 70% with the full model.

DISCUSSION

As the number of older adults in this country rises with each decade (19), the number of physician-ordered stress tests for this age group is also likely to rise. As a result, the need for accurate estimation of aerobic capacity is highly desirable in the medical community. We have examined the accuracy and potential bias of two $\dot{V}O_{2max}$ prediction equations, specifically in an older adult population.

Considering accuracy and bias, the Foster equation was better than ACSM in predicting $\dot{V}O_{2max}$. We regard Foster's mean differences of −0.2 and 0.6 mL·kg^{-1}·min^{-1} in men and women, respectively, as acceptable for a prediction equation.

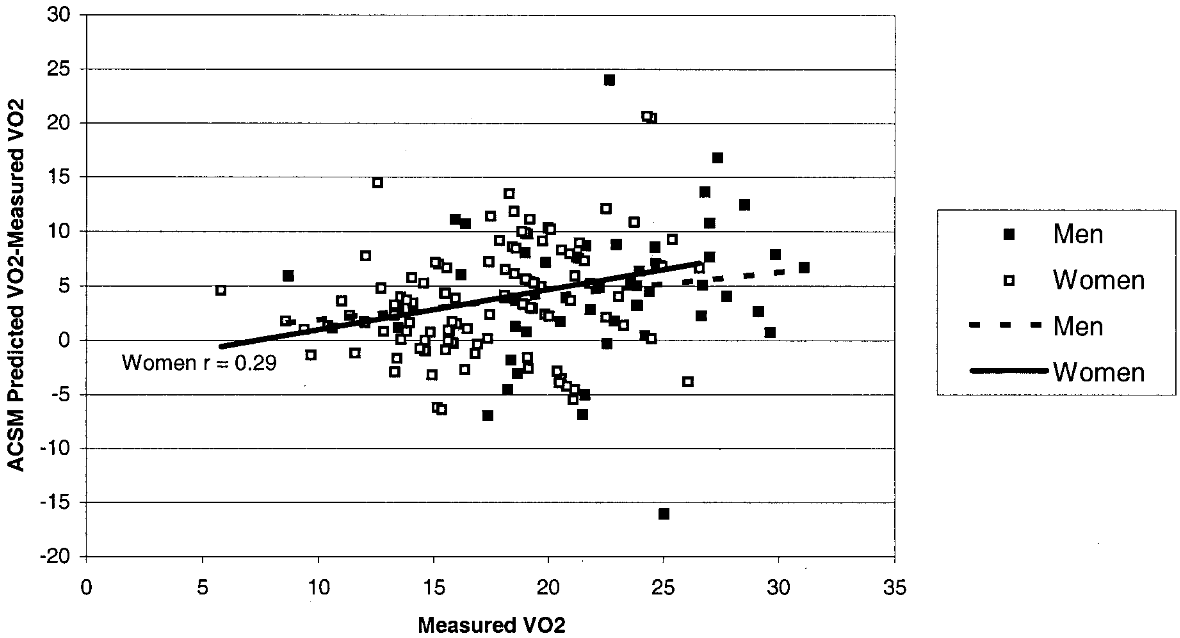

FIGURE 1—Bland-Altman plot of ACSM $\dot{V}O_2$ and measured $\dot{V}O_2$.

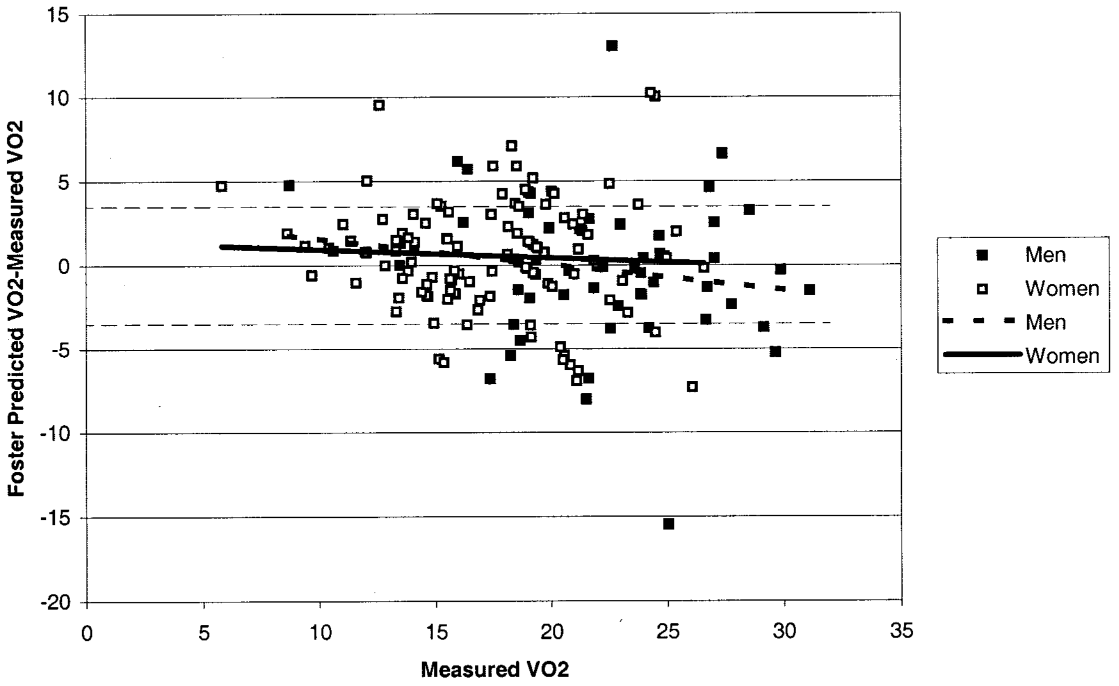

FIGURE 2—Bland-Altman plot of Foster $\dot{V}O_2$ and measured $\dot{V}O_2$, displaying upper and lower limits of agreement with 1 MET (3.5 mL·kg^{-1}·min^{-1}).

However, Foster's SD values approximated 1 MET (~3.5 mL·kg^{-1}·min^{-1}) in men and women, which is a variance of about 19% of the mean aerobic capacity of both genders. Additionally, the upper and lower 1-MET agreement lines in Figure 2 show that approximately 30% of men's and women's predicted $\dot{V}O_{2max}$ values using the Foster equation were above or below a 1 MET difference from measured $\dot{V}O_{2max}$ values. One explanation for this variance could be the lack of consistency in our participants' use of handrail-supported treadmill walking. Whereas almost all participants eventually used handrail support to some degree, we were not vigilant in ensuring equal or very similar support from test start to finish, as this was not a focus of this study at the time of data collection. The Foster equation showed no bias in prediction considering aerobic capacity (Fig. 2). This is an important strength in this context, as hospitals often see a wide variance in the fitness levels of patients referred for stress testing. Additionally, many older adults are not able to complete a treadmill stress test without at least minimal use of handrails for support and balance, especially during the latter stages of the test. The Foster equation was developed and validated with subjects who used handrail support during testing. We feel this is a more realistic expectation of this population. Using an equation, such as Foster, that considers this limitation in older adults is probably necessary. Overall, we feel the accuracy of the Foster equation was acceptable.

Our data support previous studies (3,8,9) that recommended not using the ACSM equation to estimate $\dot{V}O_{2max}$ in older adults. The ACSM equation's tendency to overestimate $\dot{V}O_{2max}$ is most likely due, in part, to its intended use for estimation during steady state exercise, which makes it inappropriate for use with stress testing of any age group. Second, the derivation of the ACSM equation is questionable, with its vertical and horizontal components developed using highly fit male subjects or based on estimates (2,7).

Based on our regression analyses, it may be beneficial to develop a prediction equation that includes important physiological variables such as gender, age, BMI, and activity levels. We observed a 0.20 increase in R^2 when adding these variables to a model that originally included "traditional" test variables of treadmill grade and speed. The lack of significance with test time, typically a strong predictor of $\dot{V}O_{2max}$, in these analyses may be a product of the Pepper protocol. Pepper was developed utilizing small incremental changes in metabolic requirements as the test progresses in an attempt to reach the optimal 8- to 12-min window (15). Conversely, protocols utilizing large metabolic increments (e.g., Bruce) would most likely result in a more heterogeneous scattering of test time, thus making test time a significant predictor of $\dot{V}O_{2max}$.

Bruce showed physical activity levels to be an independent predictor of $\dot{V}O_{2max}$ in healthy adults, below only sex and age in order of significance (5). We believe some measure of activity levels may also be a significant contributor when predicting an older adult's aerobic capacity; however, due to the poor sensitivity of our method of activity measure, signif-

TABLE 5. Predictors of $\dot{V}O_{2max}$ in older men and women.

Variable	Beta Coefficient
Gender	−2.82795942**
Age	−0.20078216**
BMI	−0.22979629**
Treadmill grade	0.45148385**
Treadmill speed	2.10815407**
*Current activity levels	0.83109770

* 1 = not walking; 2 = walking.
** Significant at $P < 0.05$;

icance was not found in our model. Also, due to our sample size, we were not able to provide a new prediction equation utilizing the combined strengths of physiological and mechanical variables. Determining and validating a method of activity measure and development of a model including these measures needs further study in this population.

CONCLUSIONS

The ACSM equation, although the most widely used prediction equation in clinical settings, is not appropriate for use when treadmill stress testing older adults. Between the two equations studied, the Foster equation is the more accurate predictor of $\dot{V}O_{2max}$ when treadmill testing older men and women. This equation also showed no bias along a continuum of aerobic capacity.

This work was supported by the National Institutes of Health, National Institute on Aging, Claude D. Pepper Older Americans Independence Center, grant no. 5 P60 AG 11268, and by the National Institutes of Health, National Center for Research Resources, General Clinical Research Centers Program grant no. M01-RR-30.

REFERENCES

1. American College of Sports Medicine. *Guidelines for Exercise Testing, and Prescription*, 6th Ed. Philadelphia: Lippincott, Williams & Wilkins, 2000, p. 97.
2. Balke, B., and R. Ware. An experimental study of "physical fitness" of Air Force personnel. *U.S. Armed Forces Med. J.* 10:675–688, 1959.
3. Berry, M., R. Brubaker, M. O'Toole, et al. Estimation of $\dot{V}O_2$ in older individuals with osteoarthritis of the knee and cardiovascular disease. *Med. Sci. Sports Exerc.* 28:808–814, 1996.
4. Bland, J. M., and D. G. Altman. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1(8476):307–310, 1986.
5. Bruce, R. A., F. Kusumi, and D. Hosmer. Maximal oxygen intake and nomographic assessment of functional aerobic impairment in cardiovascular disease. *Am. Heart J.* 85:546–562, 1973.
6. Cohen, M. C., R. S. Stafford, and B. Misra. Stress testing: national patterns and predictors of test ordering. *Am. Heart J.* 138:1019–1024, 1999.
7. Dill, D. B. Oxygen cost of horizontal and grade walking and running on the treadmill. *J. Appl. Physiol.* 20:19–22, 1965.
8. Dominick, K. L., E. C. Gullette, M. A. Babyak, et al. Predicting peak oxygen uptake among older patients with chronic illness. *J. Cardiopulm. Rehabil.* 19:81–89, 1999.
9. Foster, C., A. Crowe, E. Daines, et al. Predicting functional capacity during treadmill testing independent of exercise protocol. *Med. Sci. Sports Exerc.* 28:752–756, 1996.
10. Foster, C., A. S. Jackson, M. L. Pollock, et al. Generalized equations for predicting functional capacity from treadmill performance. *Am. Heart J.* 107:1229–1234, 1984.
11. Kaminsky, L. A., and M. H. Whaley. Evaluation of a new standardized ramp protocol: the BSU/Bruce ramp protocol. *J. Cardiopulm. Rehabil.* 18:438–444, 1998.
12. Morey, M. C., M. Schenkman, S. A. Studenski, et al. Spinal-flexibility-plus-aerobic versus aerobic-only training: effects of a randomized clinical trial on function in at-risk older adults. *J. Gerontol. Med. Sci.* 54A:M335–M342, 1999.
13. Myers, J. A survey of exercise testing: methods, utilization, interpretation, and safety in the CAHCS. *J. Cardiopulm. Rehabil.* 20:251–258, 2000.
14. Myers, J., N. Buchanan, D. Smith, et al. Individualized ramp treadmill: observations on a new protocol. *Chest* 101:236S-241S, 1992.
15. Myers, J., N. Buchanan, D. Walsh, et al. Comparison of the ramp versus standard exercise protocols. *JACC* 17:1334–1342, 1991.
16. Reuben, D. B., L. Laliberte, J. Hiris, and V. Mor. A hierarchical exercise scale to measure function at the advanced activities of daily living (AADL) level. *JAGS* 38:855–861, 1990.
17. Stratton, J. R., W. C. Levy, M. D. Cerqueira, R. S. Schwartz, and I. B. Abrass. Cardiovascular responses to exercise: effects of aging and exercise training in healthy men. *Circulation* 89:1648–1655, 1994.
18. Tavel, M. E. Stress testing in cardiac evaluation: current concepts with emphasis on ECG. *Cardiopulm. Crit. Care J.* 119:907–925, 2001.
19. U.S. Census Bureau. *Population Division, 2000*.