SPINE Volume 25, Number 13, pp 1704–1710

Aerobic Fitness Testing in Patients With Chronic Low Back Pain

Which Test Is Best?

Harriet Wittink PT, PhD, OCS,* Theresa Hoskins Michel, MS, PT, CCS,† Ronald Kulich, PhD,* Anita Wagner, PhD,‡ Andrew Sukiennik, MD,* Raymond Maciewicz, MD, PhD,§ and William Rogers, PhD‡

Study Design. This is a randomized comparison of three exercise tests in a sample of 30 patients with chronic low back pain.

Objectives. To determine, by comparing three exercise tests, which test yields the highest peak and predicted oxygen consumption in a sample of patients with chronic low back pain.

Summary of Background Data. Little is known about the level of aerobic fitness in patients with chronic low back pain, although many rehabilitation programs emphasize aerobic exercise as an important part of their therapy. Measurement of aerobic fitness levels in these patients remains a problem. In healthy individuals, the highest oxygen consumption values come from exercise tests that use the largest muscle groups. For a number of reasons, this may not be true in patients with chronic low back pain.

Methods. In this study, 30 participants with chronic low back pain performed three symptom-limited maximal exercise tests: a treadmill, an upper extremity ergometer, and a bicycle ergometer. The tests were administered in randomized order. Heart rate was continuously monitored and oxygen consumption in terms of mL/kg/minute was measured by indirect calorimetry each 30 seconds.

Results. The statistical difference among the tests was highly significant ($P < 0.0001$). The treadmill test yielded the highest peak and predicted oxygen consumption followed by the bicycle and the upper extremity ergometer test, respectively.

Conclusions. The treadmill test is the best test for measuring aerobic fitness levels in patients with chronic-low back pain. It yielded the highest peak oxygen consumption compared with the other tests, coming closest to measuring maximal oxygen consumption. [Key words: aerobic fitness, chronic low back pain, exercise testing] **Spine 2000;54:1704–1710**

The economic impact of low back pain (LBP) has been estimated to be $16 to $50 billion per year.[31] More than 50 million Americans partially or totally lose their ability to work because of pain. Consequently, a loss of 700 million work days every year is attributed to pain-related disabilities.[40] The greater the duration of disabling LBP, the greater the probability of permanent disability.[17]

Many rehabilitation programs of patients with chronic low back pain (CLBP) focus on the reversal of the "deconditioning syndrome"[24] to reduce back pain–related disability. Aerobic conditioning has been a component of several treatment approaches that report significant reductions in back pain disability.[18,24,29,35] The specific contribution of aerobic fitness (VO_{2max}) to the results achieved through these multifaceted approaches is not known.

Many factors contribute to aerobic fitness such as physical activity levels, exercise habits, and genetics. Patients with chronic pain tend to be inactive. As a result of inactivity, cardiac and skeletal muscles become inefficient at using oxygen to turn fuels into energy. This decreased efficiency results in a loss of muscle endurance and cardiac output, and thus in a lower VO_{2max}.

To establish what the literature has reported on VO_{2max} in CLBP and the optimal exercise test for these patients, a search was performed. Medline from 1966 to 1997 and CINAHL from 1982 to 1997 were searched using the text words "aerobic fitness," "aerobic capacity," "physical fitness" and "VO_{2max}" with CLBP. The references in the articles yielded by the search were explored and hand searched. This search yielded seven studies that have measured VO_{2max} in patients with CLBP.[10,13,18,20,28,37,39] In these studies, CLBP is defined as a persistent, daily back pain of 6 months or longer duration. Aerobic fitness was reported to range from very low[28] to average.[20]

The measurement of VO_{2max} in patients with CLBP and the reported data were inconsistent across all seven studies. Furthermore, most of the studies estimated VO_{2max} from submaximal testing protocols. Calculating VO_{2max} from submaximal testing tends to under- or overestimate VO_{2max} by 15% in normal subjects.[6,12,34] Clearly, the most accurate measurement of VO_{2max} would be a maximal exercise test. If a maximal exercise test cannot be performed because of patients' inability to complete a protocol as a result of pain, fatigue, or other factors, then the best test for measuring VO_{2max} would be a protocol that yields the highest level of peak oxygen

From the *New England Medical Center, Pain Management Clinic, Boston, MA, †MGH Institute for Health Professions, ‡the Health Institute, and the §Tufts University Dental School, Boston, Massachusetts.
Funded in part by the Saltonstall Fund for Pain Research.
Acknowledgment date: January 8, 1999.
First revision date: May 4, 1999.
Second revision date: August 9, 1999.
Acceptance date: September 7, 1999.
Device status category: 11.
Conflict of interest category: 14.

consumption (VO_2) because this would come closest to VO_{2max}.

The searched literature did not divulge any information on which exercise test (treadmill, bicycle, or upper extremity ergometry) yields the highest peak VO_2 in patients with CLBP. Furthermore, there was no information concerning the influence that pain has on exercise test results or whether patients with CLBP are able to exercise to maximal heart rate. These issues must be addressed before conclusions about the specific contribution of aerobic fitness to CLBP disability can be drawn.

The purposes of this study were 1) to determine which testing protocol yields the highest peak values of VO_2, heart rate, and respiratory exchange ratio (RER) in a sample of patients with CLBP, 2) to determine whether the VO_2 measures between the tests are significantly different from each other, 3) to determine which factors limit performance in each of these tests, and 4) to compare values of peak VO_2 and predicted VO_{2max} with aerobic fitness values published for normal subjects.

■ Methods

Participants and Setting. After approval of the protocol by the New England Medical Center Human Studies Committee, 30 English-speaking patients with CLBP ages 18 to 60 years were recruited for the study and provided informed consent. The participants in this study were a sample of consecutive patients referred to physical therapy in an outpatient pain management program at the New England Medical Center for evaluation and treatment of chronic back pain. Their referring physicians had ruled out malignant disease and cardiac or pulmonary conditions. Chronic low back pain was defined as pain persisting for 3 months or more.[1,8] Patients were excluded from the study if they were taking medications that influence heart rate or blood pressure, or if they had coexisting major medical disease (*i.e.,* progressive neurologic or systemic disease), amputations of one or more extremities, acute upper or lower extremity musculoskeletal pain that would interfere with exercise testing, or acute psychiatric illness. None of the patients received financial remuneration for participating in the study.

Testing Protocols. Three symptom-limited maximal exercise tests were compared: a modified Bruce treadmill test,[11,26] an upper-extremity ergometer test,[36] and a modified Åstrand-Ryhming bicycle ergometer test.[6] These tests were selected because the authors thought they would be least likely to provoke pain in patients whose exercise capacity might be limited by pain, and because they are used clinically and have established normative data. Furthermore, these three tests use three distinctive muscle groups and postures and reflect dissimilar spinal loading. Hence this study could help ascertain which of these tests can be performed with the least likelihood of worsening pain in patients with CLBP while yielding the highest oxygen uptake.

The three tests were administered in randomized order to eliminate order effect. Randomization was achieved by having patients draw blinded cards. As part of their regular treatment, patients filled out a pain clinic questionnaire that contained demographic data.

On the day of the exercise testing, patients filled out a questionnaire with questions pertaining to smoking status, location of pain, surgical history, frequency of exercise, and work status. Diagnoses were divided in three categories: low back pain, radicular pain, and radiculopathy. *Low back pain* was defined as pain confined to the back only or accompanied by radiation not as far as the knee. *Radicular pain* was defined as back pain radiating beneath the knee but without neurologic findings. *Radiculopathy* was defined as the presence of sensory or motor findings. To categorize patients into diagnostic groups, physician consultation reports were reviewed for patient pain reports, clinical evidence of radiculopathy, and final diagnosis.

Procedures and Instruments. Each participant was informed about the testing procedures and asked to refrain from smoking and caffeine intake for the 2 hours before the study, and to wear comfortable, loose clothing. All the participants had used the testing devices at least once. Standard instructions were given for 1) obtaining the participant's rating of perceived exertion (RPE) using the original Borg 6- to 20-point scale,[9] 2) completing the exercise tests (to exercise "as long as you can"), and 3) obtaining verbal numerical ratings of current perceived pain on scale of 0 (no pain) to 10 (pain as bad as it could be) points before and at the end of each test. Numerical ratings are easy to administer and score, are widely used clinically, and may be treated as ratio data.[21] The validity of numerical ratings has been well documented. They demonstrate positive and significant correlations with other measures of pain intensity.[15,22,23] Patients received no clinician feedback on their performance.

The participants performed the modified Bruce protocol on a motor-driven treadmill (Landice, 8700LTD; Randolph, NJ). The Åstrand-Ryhming bicycle test was performed on a Monark bicycle, (Sweden), which was calibrated before each testing session. Patients were told to maintain 50 rpm during the entire test. The workload was increased by 25 W each time steady state was achieved, as measured by the same heart rate during two consecutive minutes. Arm ergometry was performed on an upper extremity ergometer (UBE) (Tru · Kinetics Upper Cycle; Healey Healthcare, Sugarland, TX). The workload was started at 20 W and increased by 10 W every 3 minutes. Patients were asked to maintain a cycling rate of 60 rpm throughout the UBE test.

Indirect calorimetry was used to measure VO_2 in all exercise tests, which involved sampling expired air from a mixing chamber and measuring the O_2 and CO_2 of the samples (Vista CRX Metabolic System; Vacumetrics, Inc., Ventura, CA). The O_2 and CO_2 analyzers were calibrated immediately before each testing session using calibration gases. The VO_2 in mL/kg/minute, RER, minute ventilation, and tidal volume were measured. Metabolic data were updated every 30 seconds.

Expired air samples were collected over a 2-minute rest period, over the final 10 seconds of each second minute thereafter until the participant had reached a self-determined maximum, and for 2 minutes after termination of the test. Clinical end points for termination of the tests included patient-determined fatigue, pain, or dyspnea; objective signs of hyperpnea; clinical signs of cardiovascular intolerance; or plateau of heart rate or VO_2.[2] A three-lead electrocardiogram (ECG) was monitored continuously throughout the tests.

None of the tests were stopped prematurely because of ECG changes or signs of cardiovascular abnormalities. The highest heart rate achieved was divided by the maximal heart rate (220-age) to calculate the percentage of maximal heart rate achieved. RPE was obtained during rest, every 3 minutes during the test, and at test termination using the original Borg scale. Between tests, each participant rested until pain intensity and heart rate returned to baseline levels. Each pain clinic questionnaire was scanned into an interactive database (Access, Microsoft Office 4.2).

Instruments, Reliability, and Validity.

Indirect Calorimetry. Test–retest reliability of the indirect calorimetry measurements was assessed with the modified Bruce protocol by testing and retesting a normal subject 1 week apart before the study began, then before and after the study of the patients with CLBP, and by administering a test and retest 1 week apart to another normal subject after the completion of the study. The obtained VO_2 data from the testing was subjected to intraclass correlation (ICC) analysis.

Test–retest reliability was demonstrated in an earlier study involving five subjects (ICC = 0.98).[7] The validity of measurements obtained with indirect calorimetry, such as those obtained with the Vista Metabolic System, have been demonstrated by correlating these measurements with those obtained using direct calorimetry.[7,27,30]

Aerobic Fitness or VO_{2max}. The "gold standard" for determining absolute VO_{2max} is indirect calorimetry and attainment of the following: a maximum heart rate at least 90% of the age-predicted maximum (220-age), a plateauing of VO_2, and a RER greater than 1.[4,5,38] When absolute VO_{2max} is not attained, VO_{2max} can be extrapolated from submaximal test results to age-determined maximum heart rate. This is based on the known linear increase of heart rate with increase in oxygen uptake.[6]

In the current study, predicted VO_{2max} was calculated by regression of obtained individual VO_2 peak data against individual heart rates and extrapolation to maximal heart rate (220-age). This calculation tends to under- or overestimate VO_{2max} by 15%[6,13,34] in normal subjects because the standard deviation of maximal heart rate within an age group is ±10 beats/minute.

Statistical Analysis. For statistical analysis, the STATA (StataCorp, 1997, College Station, TX) software package was used. To determine the test–retest reliability of the metabolic data, ICC correlations were used.

The sample size of 30 participants in this study was based on the number of independent variables (treadmill, bicycle, and UBE test) times ten, as recommended for multiple regression analysis by Dawson-Saunders and Trapp.[14] Order effect of exercise testing was analyzed using a generalized estimating equations model. Means and standard deviations were calculated for all parametric demographic data, peak VO_2 mL/kg/minute, predicted VO_{2max} mL/kg/minute, heart rate, RER, RPE, and pain duration. Because regressions of peak VO_2 data with heart rates that yielded r^2 less than 0.3 were considered poor predictors of VO_{2max}, they were not included in further analysis. For nonparametric data, frequency counts were used.

The *P* values for differences between the three exercise tests were calculated by analysis of variance (ANOVA). Where appropriate, 95% confidence intervals (CI) were calculated. Single and multiple regression analysis determined the significance of the influence that the independent variables (age, sex, reported pain intensity, and test duration) had on the dependent variable (peak VO_2 mL/kg/minute and predicted VO_2 mL/kg/minute).

Table 1. Patient Characteristics Mean (SD) on the Sample (n = 30): Men (n = 14) and Women (n = 16)

Variable	Sample (n = 30)	Men (n = 14)	Women (n = 16)
Age (yr)	40.4 (9.2)	39.3 (6.6)	41.3 (11.1)
Weight (kg)	80.8 (20.1)	88.5 (18.6)	74.1 (19.4)
Height (cm)	170.4 (8.1)	175.9 (8.1)	165.5 (4.0)
Pain duration (mo)	30.4 (34.1)	19.1 (15.1)	40.3 (42.7)
Ethnic background (%)			
Hispanic	5	12.5	0
White	75	62.5	83
African-American	15	25	8
Other	5	0	9
Marital status (%)			
Married	45.8	45.5	46
Separated	12.5	9	16
Divorced	20.9	27	15
Single	20.8	18.5	23
Education (%)			
High school	65	67	64
Postgraduate	0	0	0
Smoking status (%)			
Currently	40	50	31
Formerly	30	21	38
Never	30	29	31
Pack years of smoking			
Currently	23.15 (18.47)	26.68 (22.17)	18.20 (12.19)
Formerly	13.58 (8.08)	11.50 (14.60)	14.63 (3.92)
Back Surgeries (%)	20	21	19
Work (%)			
Not working	57	71	44
Working full-time	16	14	19
Modified hours	23	14	31
Modified work	3	0	6

Results

Internal Validity of Indirect Calorimetry

Test–retest reliability of peak VO_2 mL/kg/minute was assessed with the modified Bruce protocol by administering a test and retest 1 week apart to a normal subject before the study began (ICC = 0.93), then before and after the study (ICC = 0.98), and by giving a test and retest 1 week apart to another normal subject after completion of the study (ICC = 0.97). Test-retest reliability was considered good.

Demographic Data

The sample consisted of 14 men with a mean age of 39.3 ± 6.6 years and 16 women with a mean age of 41.3 ± 11.1 years. Table 1 portrays the patients' characteristics.

Unpaired Student's *t* testing showed no significant differences in ages between men and women ($t_{28} = -0.6$; $P < 0.56$) or among any of the variables except for weight ($t_{28} = 2.07$; $P < 0.05$) and height ($t_{28} = 4.56$; $P < 0.0001$), which were greater in the men. Mean body mass index for the sample was 28%: 27% for the women and 29% for the men.

The 95% CI of pain duration was 10.35 to 27.79 months for the men and 17.50 to 63 months for the women. The duration of pain was not statistically different between the men and women ($t_{28} = 1.76$; $P < 0.09$). Diagnoses for the men included LBP (36%), radicular pain (57%), and radiculopathy (7%). Diagnoses for the

Table 2. Mean (SD) and Range of the Aerobic Values on the Three Tests for the Total Sample, With 95% Confidence Interval for Peak VO_2 mL/kg/min

Variable	Treadmill	Bicycle	Upper Extremity Ergometer
HR_{max} bpm*	145.8 (19.7)	137.2 (20.7)	123.5 (20)
	108–173	100–180	82–177
Percentage of HR_{max}*	81.4 (9.8)	76.5 (10.3)	68.8 (10)
	62–96	59–96	47–91
Peak VO_2 mL/kg/min*	24.2 (8.7)	17.1 (6)	11.7 (4.1)
	9–47.2	6.7–27.9	5.2–25.7
95% CI peak VO_2 mL/kg/min	20.95–27.41	14.73–19.22	10.11–13.2
RER	0.98 (0.1)	0.90 (0.1)	0.88 (0.1)
	0.7–1.1	0.8–1.1	0.7–1.1
RPE	18 (1.9)	16.7 (2.5)	16.1 (2.6)
	15–20	9–20	11–20

* $P < 0.0001$, treadmill > bicycle > upper extremity ergometer.
VO_2 = peak oxygen consumption, HR_{max} = heart rate maximum, bpm = beats per minute, RER = respiratory exchange ratio, RPE = rating of perceived exertion.

women included LBP (25%), radicular pain (31%), and radiculopathy (44%). The genders differed in diagnostic groups, with the women having a higher proportion of radiculopathy and a lower proportion of radicular pain.

Peak VO_2 Data

Table 2 shows the means and ranges of the aerobic values obtained with peak VO_2 mL/kg/minute testing as well as the 95% CI for peak VO_2 mL/kg/minute.

Of the 30 participants, 28 achieved the highest peak VO_2 with the treadmill test. Two (female) participants, one with radicular pain and one with radiculopathy, achieved their highest peak VO_2 with the bicycle test.

Significant differences between the three tests were found for peak VO_2 mL/kg/minute ($F = 21.89$; $P < 0.0001$) and maximal heart rate achieved ($F = 29.64$; $P < 0.000$). The 95% CI of peak VO_2 for each test showed no overlap, meaning that these tests yielded significantly different peak VO_2 mL/kg/minute values.

Treadmill testing yielded the highest peak VO_2 mL/kg/minute and maximal heart rates, followed by the bicycle test, then the UBE test for both genders and all ages (Figure 1).

No differences in RER were found by ANOVA between the tests in RER ($F = 0.67$; $P < 0.52$) for the sample, for the men ($F = 1.58$; $P < 0.23$) or the women ($F = 0.18$; $P < 0.84$).

The RPE between the tests did not differ for the men ($F = 1.64$; $P < 0.22$), but it did differ for the women ($F = 6.85$; $P < 0.0044$), with the lowest RPE achieved in the UBE test.

Regression analysis of pack years of cigarette smoking against the dependent variable (peak VO_2 mL/kg/minute) showed no statistical effect for the sample or either gender in any of the tests. Equally, smoking status had no effect on peak VO_2 mL/kg/minute for the sample or either gender in any of the tests. Diagnosis had no statistical effect on peak VO_2 mL/kg/minute in any of the three tests. The order of tests used (order effect) had no statistically significant effect on the aerobic variables ($P = 0.89$).

Reason to Terminate Testing

The self-reported reason why participants stopped testing is reported in Table 3. "Other" includes inability to perform rpm as per testing protocol for the bicycle or UBE, dizziness, or stopping of the test by the tester as the protocol was completed.

On ANOVA testing for the sample, significant differences were observed between the reasons to terminate testing ($F = 4.22$, $P = 0.01$). Pain was the most reported reason why testing on the treadmill was stopped, whereas in the other two tests fatigue was the reason.

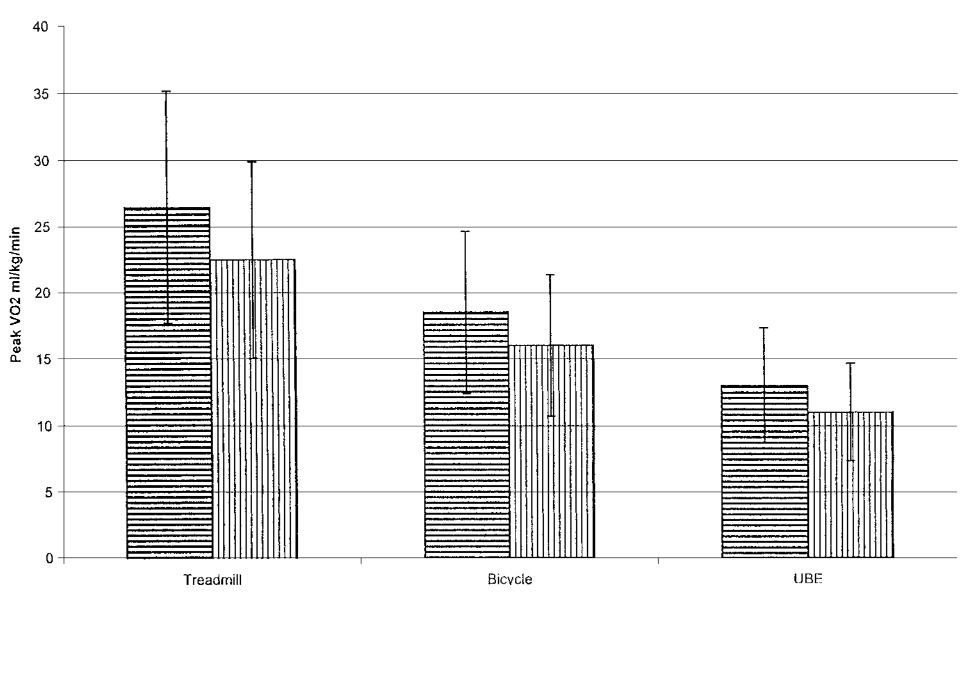

Figure 1. Peak oxygen consumption (VO_2) data for men and women.

Table 3. Reason to Stop Testing in Percentage for the Sample

Reason to Stop Testing	Treadmill	Bicycle	Upper Extremity Ergometer
Pain	56.7	40	16.7
Fatigue	36.7	56.7	70
Other	6.6	3.3	13.3

Arm fatigue was the most frequently cited reason for stopping the UBE test, and leg fatigue was the most frequently cited reason for terminating the bicycle test. Pain scores at the end of the treadmill test were significantly higher than for the other tests ($F = 3.63$; $P = 0.01$) by ANOVA.

Multiple regression equations, controlled for age, sex, fatigue, and pain, showed no statistical differences between those who stopped testing because of fatigue and those who stopped because of pain for peak VO_2 mL/kg/minute, heart rate achieved, RER, or RPE.

Predicted VO_{2max}

Regression calculations were performed for each participant's heart rates measured at each minute of the test against VO_2 measured at each minute. Predicted VO_2 was calculated by mathematically extrapolating the regression calculation to predicted heart rate (220-age). Predicted VO_{2max} values for the sample and for the men and women are shown in Table 4.

Missing predictive VO_{2max} data included one treadmill test, three bicycle tests, and seven UBE tests. When regression analysis showed a r^2 less than 0.3, predicted VO_{2max} was deemed to be invalid. A low r^2 in regression analysis could be caused by inability to maintain a cycling rate continuously as necessary in the bicycle and UBE tests. Table 5 presents the mean predicted VO_{2max} ml/kg/minute with 95% CI for the men and women by decade, along with normative values.[16]

Table 4. Mean (SD) and Range of Predicted VO_{2max} mL/kg/min for the Sample: Men and Women

Variable	Treadmill	Bicycle	Upper Extremity Ergometer
Sample			
VO_{2max} mL/kg/min*	35.3 (10.68)	30 (8.65)	22.59 (7.52)
Range	17.17–56.58	16.33–46.84	12.78–42.13
95% CI	31.24–39.36	26.59–33.43	19.34–25.84
Men			
VO_{2max} mL/kg/min*	39.77 (10.19)	32.87 (7.57)	23.82 (9.08)
Range	25.99–56.58	23.05–46.84	13.80–42.13
95% CI	33.89–45.66	28.29–37.44	17.72–29.91
Women			
VO_{2max} mL/kg/min†	31.12 (9.63)	27.36 (9.00)	21.46 (5.96)
Range	17.17–49.3	16.33–40.78	12.78–31.62
95% CI	25.79–36.46	22.16–32.55	17.68–25.25

* $P < 0.0001$, treadmill > bicycle > upper extremity ergometer.
† $P < 0.001$, treadmill > bicycle > upper extremity ergometer.
VO_{2max} = predicted maximal oxygen consumption, CI = confidence interval.

Table 5. Mean (SD) of Predicted VO_{2max} mL/kg/min (95% CI) for Men and Women by Decade in Comparison With Normative Data

Age (yr)	Men	Male Norms	Women	Female Norms
20–29			37.49 (10.7) (10.9–64.1)	36 (6.9)
30–39	41.43 (10.9) (30.1–49.8)	42 (7.0)	31.71 (7.1) (22.9–40.5)	34 (6.2)
40–49	40.23 (11.1) (12.6–67.8)	40 (7.2)	38.60 (15.1) (−97.4–147.6)	32 (6.2)
50+	31.64 (2.1) (12.8–50.5)	36 (7.1)	23.73 (5.7) (16.7–30.8)	29 (5.4)

VO_{2max} = predicted maximal oxygen consumption; CI = confidence interval.

The large 95% CIs are a reflection of a small sample size. Although the sample size was too small in each decade for drawing firm conclusions, the predicted VO_{2max} mL/kg/minute of these patients with CLBP appears to fall within the normative values for a healthy population.

The ANOVA testing showed significant differences between predicted VO_{2max} mL/kg/minute values determined by the three different testing methods for the sample, for the men ($F = 54.15$; $P < 0.0001$) and for the women ($F = 8.26$; $P < 0.002$). The treadmill test yielded the highest values, followed by the bicycle test. The UBE test yielded the lowest mean results for both the men and the women. This finding matches the peak VO_2 findings.

Peak VO_2 mL/kg/minute on treadmill, bicycle, and UBE testing was 74%, 63%, and 56%, respectively, of predicted VO_{2max} mL/kg/min. The ANOVA testing between the three tests showed that peak VO_2 as a percentage of predicted VO_{2max} was highest for the treadmill test ($F = 16.15$; $P < 0.0001$), with the lowest percentage VO_2 achieved in the UBE test (Figure 2).

The bicycle-predicted VO_{2max} was 85% of the treadmill-predicted VO_{2max} mL/kg/minute. The UBE-predicted VO_{2max} was 56% of the predicted bicycle-predicted VO_{2max} mL/kg/minute. The predicted VO_{2max} mL/kg/minute for the women as a fraction of that for the men was 78% for the treadmill, 83% for the bicycle, and 69% for the UBE.

■ Discussion

Significantly higher heart rates, peak VO_2, and predicted VO_{2max} mL/kg/minute were achieved by the modified Bruce treadmill test in this sample of patients with CLBP than by the bicycle or UBE tests. This result was independent of age, gender, smoking status, fatigue, or diagnosis. In normal subjects, the highest VO_{2max} was obtained with treadmill testing because of the muscle mass quantity involved, followed by bicycle testing. Unexpectedly, pain did not seem to alter this pattern of testing response. In fact, the testing response of patients with CLBP was remarkably similar to that of normal subjects.

Peak and predicted VO_{2max} showed gender differences consistent with published results for normal sub-

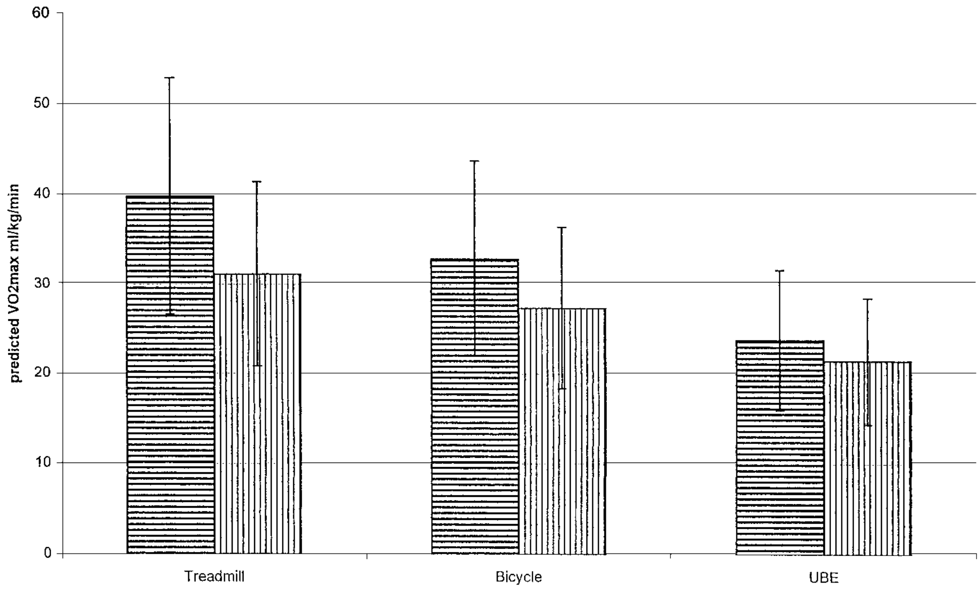

Figure 2. Predicted aerobic fitness (VO_{2max} mL/kg/minute) for men and women.

jects.[16] Normal women reach 75% to 85% of VO_{2max} values for normal men.[3] This was true for the women in the current study, with the exception of UBE testing, in which the VO_{2max} for the women was 69% of that for the men. This may be the result of arm fatigue, as supported by the lower UBE testing times for the women.

When results for the participants in this study were grouped by decade and compared with known normative values,[16] the patients with CLBP fell within the range considered normal for their age and gender, except for the patients with CLBP older 50 years who were less fit than their normal peers. Efforts must be made to substantiate this finding further, however, because the sample size for each decade was small.

According to reports, VO_{2max} achieved in bicycle testing is 5% to 15% lower than that achieved in treadmill testing in normal subjects.[19] Predicted VO_{2max} mL/kg/minute achieved in the bicycle test was 85% of the treadmill results. Astrand[6] reported a 5% to 7% difference in maximal oxygen uptake between treadmill and bicycle testing in well-trained subjects. Early leg fatigue with bicycling may explain in part why subjects fail to reach equally high values of peak and predicted VO_2 in bicycle and treadmill testing. This appeared to be the case in the current study, in which leg fatigue was the major reason (56.7%) why the patients with CLBP stopped bicycle testing.

Predicted VO_{2max} mL/kg/minute estimated from arm exercise testing is 60% to 70% that of leg exercise testing[6] in normal subjects. In this study, peak VO_2 mL/kg/minute in UBE testing was 56% of that in bicycle (leg) exercise testing. Of the patients with CLBP, 70% reported that (arm) fatigue was the reason for testing cessation. Peripheral fatigue may have played an important role in the aerobic testing of these patients. This peripheral fatigue may reflect a loss of muscle endurance as the result of prolonged inactivity.

Peak VO_2 mL/kg/minute on treadmill, bicycle, and UBE testing was 74%, 63%, and 56%, respectively, of predicted VO_{2max} mL/kg/minute. Peak VO_2 came significantly closer to predicted VO_{2max} in treadmill testing than in the other tests. The percentage of maximal heart rate achieved also was highest in the treadmill test. This means that the cardiovascular performance of the patients was better on the treadmill test than on the other tests.

Of the three tests, the treadmill test is the most functional test because it uses walking, a function of everyday life. The major disadvantage of bicycle testing is that most Americans are unaccustomed to bicycle riding. Hence their maximal values often are underestimated.[32] Many of the patients reported painful burning in the thighs and buttock pain from sitting on a hard saddle. The treadmill is the most frequently used mode of testing. It is the apparatus of choice in the laboratory because exercise intensity is determined and regulated easily.[25] Most clinics have access to a treadmill, making this a practical test.

Determining aerobic uptake by indirect calorimetric measurement is time consuming and costly. Therefore, it is not always of practical use in the clinic. A variety of tests exist that estimate VO_{2max} by submaximal testing and extrapolation to maximal heart rate. The validity of these tests in patients with CLBP is as yet unknown.[33]

In summary, aerobic testing in this sample of patients with CLBP resembled closely that reported on normal subjects, despite the subjects' pain. As compared with the other tests, the treadmill test made the highest demand on the cardiovascular system, as determined by a significantly higher percentage of maximal heart rate and by the highest peak VO_2 as a percentage of predicted VO_{2max} achieved. The treadmill test is thus the best measure of cardiovascular performance. Deconditioning of these patients expressed itself as early fatigue of peripheral muscles, limiting their performance on the bicycle and UBE tests.

■ Key Points

- Little is known about the level of aerobic fitness in patients with chronic low back pain.
- The treadmill test is the best test for measuring aerobic fitness levels in patients with chronic low back pain.
- Peak and predicted VO_{2max} showed gender differences consistent with published results for normal subjects.

Acknowledgment

The authors gratefully thank Daniel Carr, MD, Alan Jette PT, PhD, David Mostofsky, PhD, and the reviewers for their helpful comments.

References

1. ACC, the National Health Committee. New Zealand Acute Low Back Guide, Wellington, NZ: Author, 1997.
2. American College of Sports Medicine. Guidelines for exercise testing and prescription. 4th ed. Philadelphia: Lea & Febiger, 1991:1–10.
3. Astrand I. Aerobic work capacity in men and women with special reference to age. Acta Physiol Scand 1960;49:1–92.
4. Astrand I, Astrand PO, Rodahl K. Maximal heart rate during work in older men. J Appl Physiol 1959;14:562–6.
5. Astrand PO. Experimental studies of physical work capacity in relation to sex and age. Dissertation 1952:23–95.
6. Astrand PO, Rodahl K. Textbook of Work Physiology: Physiological Bases of Exercise. 3rd ed. New York: McGraw-Hill, 1986:354–91.
7. Blevins J, Broeder CE. The Vista Metabolic System, 1995. Johnson City, TN: East Tennessee State University.
8. Bigos SJ, Bowyer O, Braen G, et al. Acute Low Back Problem in Adults: Assessment and Treatment: Quick Reference Guide to Clinicians 1994. Number 14. U.S. Department of Health and Human Services, Agency for Health Care Policy and Research. AHCPR Publication No. 95-0643.
9. Borg G. Psychophysical basis of perceived exertion. Med Sci Sports Exerc 1982;14:377–81.
10. Brennan GP, Shultz BB, Hood RS, Zahniser JC, Johnson SC, Gerber AH. The effects of aerobic exercise after lumbar microdiscectomy. Spine 1994;19:735–9.
11. Bruce R, Kusumi F, Hosmer D. Maximal oxygen intake and nomographic assessment of functional aerobic impairment in cardiovascular disease. Am Heart J 1973;85:546–62.
12. Davies CTM. Limitations to the prediction of maximum oxygen intake from cardiac frequency measurements. J Appl Physiol 1968;24:700–6.
13. Davis V, Fillingim R, Doleys D, Davis M. Assessment of aerobic power in chronic pain patients before and after a multidisciplinary treatment program. Arch Phys Med Rehab 1992;73:726–9.
14. Dawson-Saunders B, Trapp RG. Basic and Clinical Biostatistics. Norwalk: Appleton & Lange, 1998:207–29.
15. Downie WW, Leatham PA, Rhind VM, Wright V, Branco JA, Anderson JA. Studies with pain rating scales. Ann Rheum Dis 1978;37:378–81.
16. Fletcher GF, Balady G, Blair SN, et al. Statement on exercise: Benefits and recommendations for physical activity programs for all Americans: A statement for health professionals by the Committee on Exercise and Cardiac Rehabilitation of the Council on Clinical Cardiology, American Heart Association. Circulation 1996;94:857–62.
17. Frymoyer JW, Cats-Baril W. Predictors of low back pain disability. Clin Orthop 1987;221:89–98.
18. Hazard RG, Fenwick JW, Kalisch SM, et al. Functional restoration with behavioral support: A one-year prospective study of patients with chronic low back pain. Spine 1989;14:157–61.
19. Hermansen L, Ekblom B, Saltin B. Cardiac output during submaximal and maximal treadmill and bicycle exercise. J Appl Physiol 1970;29:82–6.
20. Hurri H, Mellin G, Korhonen O, Harjula R, Harkapaa K, Luoma J. Aerobic capacity among chronic low back pain patients. J Spinal Disord 1991;4:34–8.
21. Jensen MP, Karoly P. Self-report scales and procedures for assessing pain in adults. In: Turk DC, Melzack R, eds. Handbook of Pain Assessment, vol 1. New York: Guilford Press, 1992:135–51.
22. Jensen MP, Karoly P, Braver S. The measurement of clinical pain intensity: A comparison of six methods. Pain 1986;27:117–26.
23. Jensen MP, Karoly P, O'Riordan EF, Bland FJ, Burns RS. The subjective experience of acute pain: An assessment of the utility of 10 indices. Clin J Pain 1989;5:153–9.
24. Mayer TG, Gatchel RJ, Kishino N, et al. A prospective short-term study of chronic low back pain patients utilizing novel objective functional measurement. Pain 1986;25:53–68.
25. McArdle WD, Katch FI, Katch VL. Exercise Physiology: Energy, Nutrition, and Human Performance. 3rd ed. Malvern: Lea & Febiger, 1991:216.
26. McInnis K, Balady G, Weiner D, Ryan T. Comparison of ischaemic and physiologic responses during exercise tests in men using the standard and modified Bruce protocols. Am J Cardiol 1992;69:84–9.
27. McLellan TM. Defence and Civil Institute of Environmental Medicine [letter]. Ontario, Canada, 1991.
28. McQuade K, Turner J, Buchner DM. Physical fitness and chronic low back pain: An analysis of the relationship among fitness, functional limitations, and depression. Clin Orthop Rel Res 1988;233:198–204.
29. Mitchell RI, Carmen GM. Results of a multicentre trial using an intensive active exercise program for the treatment of acute soft tissue and back injuries. Spine 1990;15:514–21.
30. Montoye HG, Kemper HCG, Saris WHM, Washburn RA. Measuring physical activity and energy expenditure. Champagne, IL: Human Kinetics, 1996.
31. Plowman SA. Physical activity, physical fitness, and low back pain. Exerc Sport Sci Rev 1992;20:221–51.
32. Pollock ML, Wilmore JH. Medical screening and evaluation procedures. In: Pollock ML, Wilmore JH, eds. Exercise in Health and Disease: Evaluation and Prescription for Prevention and Rehabilitation. 2nd ed. Philadelphia: WB Saunders Company, 1990:239–363.
33. Protas EJ. Aerobic exercise in the rehabilitation of individuals with chronic low back pain: A review. Crit Rev Phys Rehabil Med 1996;8:283–95.
34. Rowell LB, Taylor HL, Wang Y. Limitations to prediction of maximal oxygen uptake. J Appl Physiol 1964;19:919–27.
35. Saal JA, Saal JS. Nonoperative treatment of herniated lumbar intervertebral disc with radiculopathy: An outcome study. Spine 1989;14:431–7.
36. Sawka M. Physiology of upper body exercise. Exerc Sport Sci Rev 1986;14:175–211.
37. Schmidt A. Cognitive factors in the performance level of chronic low back pain patients. J Psychosom Res 1985;29:183–9.
38. Taylor HL, Buskirk E, Henschel A. Maximal oxygen intake as an objective measure of cardiorespiratory performance. J Appl Physiol 1955;8:73–80.
39. Thomas LK, Hislop H, Waters R. Physiological work performance in chronic low back disability: Effects of a progressive activity program. Phys Ther 1980;60:407–11.
40. Vasudevan S. Impairment, disability, functional capacity assessment. In: Turk DC, Melzack R, eds. Handbook of Pain Assessment. New York: Guildford Press, 1992.

Address reprint requests to

Harriet Wittink PT, PhD, OCS
New England Medical Center
Pain Management Clinic
750 Washington Street 298
Boston, MA 02111