Reproducibility of Maximal Exercise Ergometer Testing in Patients With Cystic Fibrosis*

E. F. McKone, MD; S. C. Barry, BPhysio; M. X. FitzGerald, MD, FCCP; and C. G. Gallagher, MD, FCCP

Objectives: Exercise testing in patients with cystic fibrosis (CF) has become an important tool in assessing disease severity and predicting overall outcome. The reproducibility of maximal exercise testing was examined in adult subjects with stable CF.

Methods: Nine subjects with CF underwent a total of three maximal exercise tests carried out under identical circumstances over a 28-day period. Oxygen uptake (\( \dot{V}O_2 \)), minute ventilation (\( \dot{V}E \)), respiratory frequency (f), heart rate (HR), and arterial oxygen saturation (\( SaO_2 \)) were measured at rest, at end exercise, and at 40% and 70% of maximum workload.

Results: There were no significant differences in these measurements among the three tests. Reproducibility of exercise performance was assessed using the coefficient of variation. The mean within-subject coefficient of variation for test variables at end exercise are as follows: \( \dot{V}O_2 \), 6.9%; \( \dot{V}E \), 6.2%; f, 5.8%; HR, 3.0%; and \( SaO_2 \), 1.1%. The mean within-subject coefficient of variation for test variables at 40% and 70% of maximal work rates are as follows: \( \dot{V}O_2 \), 5.2% and 4.6%; \( SaO_2 \), 0.3% and 0.9%; HR, 4.0% and 3%; f, 5.7% and 6.5%; and f, 5.8% and 7.2%, respectively.

Conclusions: Variables measured during clinical cycle ergometer exercise testing in adult patients with stable CF are reproducible. No learning effect was found on repeated testing.

Key words: cystic fibrosis; exercise testing; reproducibility

Abbreviations: CAL = chronic airflow limitation; CF = cystic fibrosis; f = respiratory frequency; HR = heart rate; ILD = interstitial lung disease; MVV = maximum voluntary ventilation; \( SaO_2 \) = arterial oxygen saturation; \( VCO_2 \) = carbon dioxide output; \( \dot{V}E \) = minute ventilation; \( \dot{V}O_2 \) = oxygen uptake; \( \dot{V}O_2\text{max} \) = peak oxygen uptake; \( VT \) = tidal volume; Wmax = peak work rate

Exercise testing in patients with cystic fibrosis (CF) has become an important tool in assessing disease severity and in predicting overall outcome.1,2 It is well documented that patients with CF have impaired exercise performance, as shown by reduced peak oxygen uptake (\( \dot{V}O_2\max \)), reduced peak work rate (Wmax), and abnormal ventilatory and cardiovascular responses to exercise.3–7 Despite the use of clinical exercise testing in CF patients, the reproducibility of maximal exercise testing has not, to our knowledge, been examined in these patients. In order to assess the clinical significance of repeated maximal exercise testing, we examined the reproducibility of clinical exercise testing in our patients with CF. Both group mean and individual responses to exercise were tested in adult patients with stable CF undergoing repeated incremental exercise tests under identical conditions.

Materials and Methods

Patients

Characteristics of our study population are outlined in Table 1. The subjects (six men and three women) were recruited from outpatients in the National Adult Cystic Fibrosis Unit, St. Vincent’s University Hospital. All subjects had clinical, radiologic, and physiologic evidence of CF-related lung disease, and they all had CF diagnoses based on clinical features, abnormal sweat test (sweat sodium and chloride > 60 mmol/L), and genotyping. All subjects were clinically stable for a period of 2 months prior to taking part in the study, and they all had no evidence of rheumatologic, neuromuscular, cardiac, or peripheral vascular disease, or any disease other than CF that might impair exercise tolerance.

Pulmonary function testing was carried out for each patient...
Comparisons of tests 1, 2, and 3 were performed by repeated measures analysis of variance.

Data are presented as mean (SD). VC

Leg discomfort, Borg scale 3.6 (1.3) 3.7 (1.0) 4.0 (1.1) NS

Dyspnea, Borg scale 4.0 (1.1) 4.2 (1.2) 4.1 (1.0) NS

HR, beats/min 168 (17) 164 (17) 167 (15) NS

Maximum V\textsubscript{\textsuperscript{o}2}, % predicted 62 (8)

Maximum HR, % predicted 86 (6)

VE/MVV, %† 98 (27)

Ve/V\textsubscript{\textsuperscript{o}2}† 42 (4.8)

Desaturation, % 4 (2.7)

*Data are given as mean (SD). Maximum V\textsubscript{\textsuperscript{o}2} = maximum oxygen consumption at end exercise; Maximum HR = maximum heart rate at end exercise.

†Ve/MVV and Ve/V\textsubscript{\textsuperscript{o}2} are measured at end exercise.

using a spirometer (Pneumocheck; Welch Allyn; Skaneateles Falls, NY). FEV\textsubscript{1} and FVC were measured using recommended techniques,\textsuperscript{a} and predicted normal values\textsuperscript{b} were used to calculate percentage predicted values.

The study was approved by the Ethics Committee of St. Vincent’s University Hospital. All subjects gave informed consent for the procedures.

Protocol

FEV\textsubscript{1} and FVC were measured before and after each exercise test. At least three well-coordinated maximal efforts were obtained, and the highest value obtained for each variable was recorded. The subjects underwent three exercise tests over a 28-day period, with each two tests separated by at least 7 days. Exercise was performed at the same time of day on each occasion. The subjects were asked to avoid strenuous activity for at least 24 h and food or caffeinated drinks for 2 h prior to exercise testing. All patients were instructed to take all of their medications during the testing to ensure consistency of the protocol. Management was offered, and no communication was made with the subjects during the testing to ensure consistency of the protocol.

ECG leads attached to the chest enabled continuous monitoring of the heart rate (HR). Arterial oxygen saturation (Sa\textsubscript{O2}) was monitored by pulse oximetry (SAT-TRAK Pulse Oximeter; SensorMedics; Yorba Linda, CA). Each patient’s mouthpiece was connected to a heated wire flowmeter (Mass Flow Sensor; SensorMedics). The flow signal was digitally integrated to give tidal volume (VT), and expired gases were continually analyzed by rapidly responding oxygen (paramagnetic) and carbon dioxide (infrared) analyzers. All equipment was calibrated before each exercise study using calibration syringes and precision oxygen and carbon dioxide gas mixtures. All signals were continuously displayed breath by breath on a computer screen in real time during the exercise test. Data were also stored on computer hard disk for later analysis.

Data Analysis

Minute ventilation (Ve), VT, respiratory frequency (f), HR, oxygen uptake (VO\textsubscript{2}), and carbon dioxide output (VCO\textsubscript{2}) were measured breath by breath using standard formulas.\textsuperscript{10,11} Ve and VT were expressed at body temperature and pressure, saturated with water vapor; VO\textsubscript{2} and VCO\textsubscript{2} were expressed at standard temperature pressure, dry. Predicted VO\textsubscript{2max} during exercise was calculated as follows\textsuperscript{9b}: 

\[
\text{VO}_{2\text{max}} = 0.83 \text{ht}^2 \times (1 - 0.007 \times \text{age}) \times (1 - 0.25S)
\]

where height (ht) is in meters, age is in years, and S represents gender (S = 0 for men and S = 1 for women). Predicted peak HR was calculated as follows\textsuperscript{9b}:

\[
\text{Peak HR} = 210 - (0.66 \times \text{age [years]})
\]

Table 1—Characteristics of the Study Population (n = 9)*

<table>
<thead>
<tr>
<th>Variables</th>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, yr</td>
<td>26.3 (8.3)</td>
</tr>
<tr>
<td>Gender, No.</td>
<td>Male 6, Female 3</td>
</tr>
<tr>
<td>FEV\textsubscript{1}, L [% predicted]</td>
<td>2.00 (0.66) [56]</td>
</tr>
<tr>
<td>FVC, L [% predicted]</td>
<td>3.23 (0.67) [76]</td>
</tr>
<tr>
<td>FEV\textsubscript{1}/FVC, %</td>
<td>62 (12)</td>
</tr>
<tr>
<td>Maximum VO\textsubscript{2}, % predicted</td>
<td>62 (8)</td>
</tr>
<tr>
<td>Maximum HR, % predicted</td>
<td>86 (6)</td>
</tr>
<tr>
<td>Ve/MVV, %†</td>
<td>98 (27)</td>
</tr>
<tr>
<td>Ve/V\textsubscript{\textsuperscript{o}2}†</td>
<td>42 (4.8)</td>
</tr>
<tr>
<td>Desaturation, %</td>
<td>4 (2.7)</td>
</tr>
</tbody>
</table>

*Data are given as mean (SD). Maximum VO\textsubscript{2} = maximum oxygen consumption at end exercise; Maximum HR = maximum heart rate at end exercise.

†Ve/MVV and Ve/V\textsubscript{\textsuperscript{o}2} are measured at end exercise.

Table 2—Group Values Obtained at End of Exercise*

<table>
<thead>
<tr>
<th>Variables</th>
<th>Test 1</th>
<th>Test 2</th>
<th>Test 3</th>
<th>p Value†</th>
</tr>
</thead>
<tbody>
<tr>
<td>VO\textsubscript{2}, L/min</td>
<td>1.48 (0.53)</td>
<td>1.52 (0.37)</td>
<td>1.57 (0.47)</td>
<td>NS</td>
</tr>
<tr>
<td>Ve, L/min</td>
<td>62.1 (20.49)</td>
<td>65.5 (17.15)</td>
<td>64.7 (16.2)</td>
<td>NS</td>
</tr>
<tr>
<td>Exercise duration, s</td>
<td>559 (177)</td>
<td>593 (177)</td>
<td>591 (173)</td>
<td>NS</td>
</tr>
<tr>
<td>Work rate, W</td>
<td>129 (33)</td>
<td>139 (30)</td>
<td>140 (30)</td>
<td>NS</td>
</tr>
<tr>
<td>VT, L</td>
<td>1.45 (0.37)</td>
<td>1.45 (0.40)</td>
<td>1.46 (0.44)</td>
<td>NS</td>
</tr>
<tr>
<td>VT/VC</td>
<td>0.46 (0.05)</td>
<td>0.45 (0.06)</td>
<td>0.46 (0.06)</td>
<td>NS</td>
</tr>
<tr>
<td>f, breaths/min</td>
<td>43.9 (9.4)</td>
<td>43.9 (9.0)</td>
<td>45.6 (8.3)</td>
<td>NS</td>
</tr>
<tr>
<td>O\textsubscript{2} saturation, %</td>
<td>90.5 (3.5)</td>
<td>91.4 (2.7)</td>
<td>91.4 (2.7)</td>
<td>NS</td>
</tr>
<tr>
<td>HR, beats/min</td>
<td>168 (17)</td>
<td>164 (17)</td>
<td>167 (15)</td>
<td>NS</td>
</tr>
<tr>
<td>Dyspnea, Borg scale</td>
<td>4.0 (1.1)</td>
<td>4.2 (1.2)</td>
<td>4.1 (1.0)</td>
<td>NS</td>
</tr>
<tr>
<td>Leg discomfort, Borg scale</td>
<td>3.6 (1.3)</td>
<td>3.7 (1.0)</td>
<td>4.0 (1.1)</td>
<td>NS</td>
</tr>
</tbody>
</table>

*Data are presented as mean (SD). VC = vital capacity; NS = not significant.

†Comparisons of tests 1, 2, and 3 were performed by repeated measures analysis of variance.
Maximum voluntary ventilation (MVV) was estimated as follows:

\[ \text{FEV}_1 \times 35. \]

Sense of dyspnea was assessed using the Borg scale in response to the question, “How breathless do you feel?” with the subject pointing to the appropriate number on the scale. Leg discomfort was assessed using the Borg scale, in response to the question, “How much leg discomfort do you feel?” Each patient was also asked the reason for stopping exercise immediately after each exercise test.

**Statistical Analysis**

Data collected at rest, at maximal exercise, and at two matched submaximal work rates were used in the analysis. Submaximal work rates of 40% and 70% of the highest Wmax achieved during the three tests were chosen for each patient. Comparisons were then made at matched work rates for the three studies.

Statistical significance of group mean data from the three experiment days were determined by repeated-measures analysis of variance. The variability of subject results for the three experiments was assessed using the coefficient of variation. The coefficient of variation was derived by dividing the SD by the mean. Analysis of the Borg scale was performed using Wilcoxon.

| Table 3—Mean Group Coefficient of Variation Measured at Rest, End Exercise, and at Matched Submaximal Levels of Exercise |
|---|---|---|---|---|---|---|
| | \( \text{VO}_2 \) \%, | \( \text{Ve} \) \%, | \( \text{VE} \) \%, | \( f \) \%, | \( \text{HR} \) \%, | \( \text{SaO}_2 \) \% |
| End exercise | 6.9 | 6.2 | 3.8 | 5.8 | 3.0 | 1.1 |
| 70% Wmax | 4.6 | 6.5 | 4.9 | 7.2 | 3.0 | 0.9 |
| 40% Wmax | 5.2 | 5.7 | 5.7 | 5.8 | 4.0 | 0.3 |
| Baseline | 7.1 | 6.5 | 6.9 | 8.6 | 4.6 | 1.3 |

**Figure 1.** Comparison of each patient’s \( \text{VO}_2 \) and exercise duration at end exercise for all three tests.

**Figure 2.** Graphic representation of the relationship of \( \text{Ve} \) to \( \text{VO}_2 \) throughout each exercise test in subjects 2, 4, and 9. Their \( \text{FEV}_1 \) as percent predicted is shown.
RESULTS

All subjects completed each exercise test without any complications, and no exercise test was terminated by the physician.

Resting Results

Analysis of group mean data collected before exercise showed no significant difference in baseline spirometric values. Mean FEV₁ was 56% predicted (range, 30 to 82%). Mean FVC was 76% predicted (range, 55 to 100%). The mean coefficient of variation for FEV₁ was 4.7% (range, 1.0 to 8.8%), and for FVC, 5.2% (range, 2.0 to 9.4%). The mean resting SaO₂ was 95% (range, 93 to 97%). There was no significant difference in baseline values of VO₂, V̇CO₂, V̇E, V̇t, f, SaO₂, and HR in the three separate studies.

Exercise Results

All patients had evidence of impaired exercise tolerance. Mean (± SD) values for VO₂max as percent predicted was 62 ± 8% (range, 51 to 71%). The mean maximal heart rate was 86 ± 6% (range, 78 to 98%) of predicted. Mean oxygen desaturation was 4 ± 2.7% (range, 1 to 9%). The three subjects with the lowest FEV₁ had significant desaturation of >5%. The V̇E/MVV ratio was 98 ± 27% (range, 57 to 124%). Reasons for discontinuing exercise were dyspnea (three patients), leg discomfort (four patients), or both (two patients). This did not vary from test to test.

Eight of the nine patients reached their anaerobic threshold as calculated using the modified V-slope method.

Table 2 lists the group mean data collected at end of exercise. There was no significant difference at end of exercise in VO₂, V̇CO₂, V̇E, V̇t, f, HR, SaO₂, exercise time, or work rate. In addition, there was no significant difference in any variables at 40% Wmax or 70% Wmax.

Exercise Variability

The mean within-subject coefficients of variation measured during exercise are shown in Table 3. Values obtained at rest, end exercise, and 40% and 70% Wmax are shown.

The mean within-subject coefficient of variation at end of exercise for VO₂max was 6.9% (range, 1 to 13% throughout exercise); for SaO₂, 1.1% (range, 0 to 5% throughout exercise); and HR, 3.0% (range, 1.7 to 8.0% throughout exercise).

The mean within-subject coefficient of variation for V̇E at end exercise was 6.2% (range, 3.0 to 12.3% throughout exercise); for V̇t, 3.8% (range, 1 to 15% throughout exercise); and for f, 5.8% (range, 2 to 15% throughout exercise).

The mean within-subject coefficient of variation for exercise duration was 4.7% (range, 1.9 to 13.2%) and for Wmax, 6.0% (range, 0.6 to 14.0%). At end exercise, the mean coefficient of variation for Borg scale leg discomfort was 8.4% (range, 0 to 25%) and for Borg scale dyspnea, 11.3% (range, 0 to 25%).

Figure 1 shows each subject’s VO₂max and exercise duration for each of the three exercise tests. Figure 2 is a graphic representation of VO₂ against V̇E throughout the exercise tests for three patients with varying degrees of lung dysfunction. The relationship between VO₂ and V̇E is similar in each subject’s three tests.

DISCUSSION

As in previous studies of exercise in patients with CF, our patients showed reduced exercise tolerance with reduced maximal workload and VO₂max. Our patients had increased ventilatory requirements with a high ventilation (V̇E), an elevated ventilatory equivalent for oxygen (V̇E/V̇O₂) and an increased V̇E/MVV at end exercise. Two patients reached their predicted maximum HR. The three patients with the lowest FEV₁ showed significant oxygen desaturation during exercise.

We looked at the reproducibility and within-subject variability of measured parameters during maximal incremental exercise testing in our adult CF population. This variability, as measured by the coefficient of variation, is similar to that measured in studies carried out on healthy subjects and in patients with chronic airway limitation (CAL), interstitial lung disease (ILD), and cardiac failure.

Previous studies have examined the reproducibility of maximal exercise testing. Garrard and Emmons found some diurnal variation in their healthy subjects, with a coefficient of variation of 8.4%; V̇E, 12.0%; and HR, 3.8%. Nordrehaug et al examined the reproducibility of maximal treadmill exercise testing in healthy subjects and found a coefficient of variation for VO₂max of 5.0%; V̇E, 7.0%; and HR, 3.0%. They also found that the variability at end exercise was less than that at submaximal levels of exercise.

The reproducibility of repeated exercise testing...
has also been extensively examined in patients with CAL. Swinburn et al²⁰ examined the repeatability of walking, step, and cycle ergometer tests and found a significant learning effect on repeated exercise testing, as well as significant variability in VO₂ and VE, depending on which exercise test was performed. Noseda et al²¹ examined repeated cycle ergometer exercise tests in patients with CAL at intervals of 1 month. Their coefficient of variation for FEV₁ was 10.2%; VO₂, 9.0%; VE, 8.1%; and HR, 5.0%. Owens et al²² found similar results with maximal ergometer exercise tests also separated by 1 month, with a coefficient of variation for FEV₁ of 7.5%; VO₂, 6.6%; VE, 6.3%; and HR, 3.5%. Cox et al²³ examined reproducibility of exercise testing carried out on consecutive days and calculated a relative duplicate error of 3.5% for VO₂ and 6.6% for VE.

In patients with ILD, Marciniuk et al²⁴ examined the reproducibility of maximal ergometer exercise testing and found a coefficient of variation for FEV₁ of 3.5%; VO₂, 5.3%; VE, 5.5%; and HR, 4.0%. This is similar to findings in patients with cardiac failure where Janicki et al²⁵ have shown good reproducibility, although Elborn et al²⁶ showed significant increases in exercise duration and workload with repeated exercise testing.

Our patients with CF had a coefficient of variation for FEV₁ of 4.7%; VO₂max, 6.9%; and VE, 6.2%. This is similar to the findings of Nordrehaug et al in healthy subjects. Comparison with patients with respiratory disease shows adult patients with CF to have variability slightly greater than ILD patients but less variability than CAL patients.

All of our patients stated that they made a maximal effort and exercised until exhaustion. The Borg scores at end exercise (dyspnea, 4.1 ± 1.1; leg discomfort, 3.8 ± 1.1) are similar to those in other studies assessing breathlessness at end exercise in other diseases. Marciniuk et al²⁴ found that patients with ILD had mean Borg scores of 4.5 at end exercise. In their study of added dead space during maximal incremental exercise testing in ILD, the control group of the study (with no dead space) had mean scores of 5.0 at end exercise.²⁷ Studies looking at the reproducibility of Borg scoring after maximal exercise testing in patients with CAL have also found a mean of 5.0 after repeated testing.²⁸ Other factors favoring a maximal effort by our patients is that the majority of the patients either reached their maximal predicted HR at end exercise or showed a high (> 90%) VE/MVV ratio at end exercise. Patients with low Borg scores (≤ 3) for dyspnea all gave higher scores (> 4) for leg discomfort and vice versa. Finally, all our exercise tests were observed by a physician who felt that, at end exercise, all the patients had given maximal effort.

There are limitations to the current assessment of factors that limit exercise,²⁹ but allowing for these, one patient appeared primarily limited by cardiac factors, seven patients appeared limited by respiratory factors, and one appeared limited by a combination of both. The patients with the most severe CF were limited primarily by respiratory factors, with VE/MVV ratios > 90% and low predicted HR, while the two patients with the mildest disease were primarily limited by cardiac and/or respiratory factors, reaching > 97% predicted maximum HR.

In clinical practice, changes in individual patients are more useful than changes within groups of patients. We looked at within-subject variation and how this could be applied in a clinical setting. Our results indicate that changes of 13% in exercise duration, 19% in VO₂max, and at least 17% in peak VE are unlikely to occur by chance.¹⁴

In conclusion, we have found that parameters measured in repeated cycle ergometer exercise testing are reproducible; this variability is similar to that seen in healthy subjects and in patients with CAL and ILD. There was no obvious learning effect between the tests. We also noted that the reproducibility of spirometric measurements (FEV₁ and FVC) before and after exercise was similar to that of healthy subjects and, although slightly less than that measured in ILD, is more reproducible than that seen in CAL. In addition, we looked at exercise testing at submaximal workloads (40% and 70% Wmax) and also found this to be reproducible.

Our patients showed no learning or training effect with repeated exercise testing, indicating that practice testing is not necessary in patients with stable CF who have not previously used a cycle ergometer.

REFERENCES