

# Reproducibility of incremental maximal cycle ergometer testing in patients with restrictive lung disease

Darcy D Marciniuk, Rodney E Watts, Charles G Gallagher

## Abstract

**Background**—Exercise testing has become an important tool in the diagnosis and treatment of restrictive lung disease. The reproducibility of variables measured during exercise testing was examined in subjects with stable restrictive lung disease.

**Methods**—Six subjects, who had never previously undergone exercise testing, each underwent three maximal incremental exercise studies on a bicycle ergometer conducted during a 28 day period.

**Results**—Data collected at rest, before exercise, were not significantly different during the three study days. Comparison of results at the end of the exercise tests from the three studies also revealed no evidence of a significant learning effect. Reproducibility of exercise performance by subjects was assessed by the coefficient of variation. The mean within subject coefficient of variation at the end of the exercise tests was 5.6% for work rate, 7.9% for exercise duration, and 9.5% for dyspnoea. The mean within subject coefficient of variation obtained at the end of the exercise tests was 5.3% for oxygen uptake ( $\dot{V}O_2$ ), 2.5% for oxygen saturation ( $SAO_2$ ), 4.0% for heart rate (HR), 5.5% for minute ventilation ( $\dot{V}E$ ), 5.8% for respiratory frequency ( $f$ ), and 4.6% for tidal volume ( $V_T$ ). The mean within subject coefficient of variation at 40% and 70% of maximal work rates for  $\dot{V}O_2$  was 5.7% and 5.6% respectively, for  $SAO_2$  1.3% and 1.5%, for HR 4.8% and 4.0%, for  $\dot{V}E$  6.3% and 6.6%, for  $f$  10.1% and 7.8%, and for  $V_T$  6.0% and 4.5%.

**Conclusions**—Variables measured during clinical exercise testing in subjects with restrictive lung disease are highly reproducible. No significant learning effect was found on repeated testing in subjects who had never previously undergone exercise testing.

(Thorax 1993;48:894-898)

Exercise testing in patients with restrictive lung disease has become an important tool in early diagnosis for assessing disability and evaluating responses to treatment.<sup>1,2</sup> It is well established that patients with restrictive lung disease have impaired exercise tolerance, as

shown by reduced peak oxygen uptake (peak  $\dot{V}O_2$ ) and reduced maximal work rate ( $W_{max}$ ),<sup>3</sup> abnormal ventilatory responses,<sup>4-6</sup> and a significant drop in arterial oxygen saturation ( $SAO_2$ ) during exercise.<sup>3,7,8</sup> Despite the diagnostic and prognostic implications derived from these various responses, however, the reproducibility of these variables measured during clinical exercise testing has not previously been examined in these patients. In order to assess its clinical significance we therefore studied the reproducibility of incremental exercise tests in patients with restrictive lung disease. Both group mean and individual patient reproducibility of variables was assessed in patients with stable restrictive lung disease undergoing repeated incremental exercise studies conducted under identical circumstances.

## Methods

### PATIENTS

Characteristics of the study population are outlined in table 1. Subjects (five men and one woman) were recruited from the outpatient clinics of the Division of Respiratory Medicine. All subjects had clinical, radiographic, and pulmonary function testing evidence of restrictive lung disease, but no evidence of pulmonary restriction secondary to chest wall disease, pleural disease, or respiratory muscle weakness. Restrictive lung disease was the result of idiopathic pulmonary fibrosis in three patients, sarcoidosis in two patients, and scleroderma in one patient; two were confirmed by open lung biopsy while one was confirmed by transbronchial lung biopsy. All subjects were clinically stable for at least two months before entering the study and had no rheumatological, neuromuscular, cardiac, peripheral vascular, or any disease apart from restrictive lung disease that might impair exercise tolerance. None of the sub-

Table 1 Mean (SD) characteristics of the study population (n = 6)

Age (y)	44(15)
TLC (l)	4.67(1.01)(68% predicted)
VC (l)	2.83(0.84)(57% predicted)
TLCO (% predicted)	57(22)
Kco (% predicted)	80(31)
FEV <sub>1</sub>	2.31(0.52)(63% predicted)
FEV <sub>1</sub> /FVC	83(8)

TLC—total lung capacity; VC—vital capacity; TLCO—transfer factor; Kco—transfer coefficient; FEV<sub>1</sub>—forced expired volume in one second; FVC—forced vital capacity.

Division of Pulmonary Medicine,  
Department of Medicine, University of Saskatchewan,  
Saskatoon, Saskatchewan, Canada  
D D Marciniuk  
R E Watts  
C G Gallagher

Reprint requests to:  
Dr D D Marciniuk, Division of Pulmonary Medicine,  
Department of Medicine,  
Royal University Hospital,  
Saskatoon, Saskatchewan  
S7N 0X0, Canada

Received 3 December 1992  
Returned to authors  
27 April 1993  
Revised version received  
18 May 1993  
Accepted 25 May 1993

jects had previously undergone clinical exercise testing.

Pulmonary function was measured (SensorMedics 2200) in each subject before the start of the study. Forced expired volume in one second (FEV<sub>1</sub>) and forced vital capacity (FVC) were measured using recommended techniques,<sup>9</sup> lung volumes were obtained by the helium dilution method, and the carbon monoxide transfer factor (TLCO) was measured by the single breath technique. In order to calculate percentage predicted, the normal values of Morris and coworkers<sup>10</sup> were used for spirometry, those of Goldman and Becklake<sup>11</sup> for lung volumes, and those of Burrows and coworkers for transfer factor.<sup>12</sup>

The study was approved by the medical ethics committee of the University of Saskatchewan. All subjects gave informed consent to the procedures, but none was aware of the specific purpose of the study.

#### PROTOCOL

Before each exercise study FEV<sub>1</sub> and FVC were measured with a pneumotachograph. At least three well coordinated maximal efforts were obtained and the highest value obtained for each variable was recorded. The subjects underwent a total of three exercise studies which were conducted during a 28 day period, each study separated by at least seven days. Exercise was performed in an air conditioned room and at the same time of day on each occasion. The subjects were asked to avoid strenuous activity on the day of testing and food or caffeinated drinks in the preceding two hours.

Testing was performed on an electrically braked cycle ergometer (Godart) while the subject breathed room air. After mounting the cycle ergometer and being attached to the mouthpiece, each patient rested quietly for two minutes before starting to pedal. A nose clip was used for all experiments. The initial exercise work load was 15 W and was increased by 15 W every minute of exercise until exhaustion. With the use of speedometer feedback each subject chose the pedalling rate within a range of 50–70 rpm. All subjects were instructed in an identical manner by the same operator for all exercise studies. The subjects were told that they should continue to exercise until they could exercise no more. No type of encouragement was offered and no communication was made with the subjects during the testing.

Electrocardiographic leads attached to the chest enabled continuous monitoring of the heart rate (HR) and electrocardiogram. SaO<sub>2</sub> was monitored continuously by pulse oximetry (Nellcor 200). The mouthpiece was attached to a valve (Hans Rudolph Model No. 2730). Inspiratory flow and volume were measured by a standard inspiratory pneumotachograph/transducer/demodulator/integrator system using methods described previously.<sup>13</sup> The response of the system was linear over the range of inspiratory flows encountered in the study. The expired line was connected to a seven litre mixing cham-

ber with baffles. Mean expired oxygen and carbon dioxide were analysed with a mass spectrometer (Airspec 2200). All equipment was calibrated before and after each exercise test. All signals were continuously displayed in real time on an eight channel recorder (Gould), and also stored on computer for later analysis.

#### DATA ANALYSIS

Minute ventilation ( $\dot{V}_E$ ), tidal volume ( $V_T$ ), respiratory frequency ( $f$ ), HR, oxygen uptake ( $\dot{V}_{O_2}$ ), and carbon dioxide output ( $\dot{V}_{CO_2}$ ) were calculated over the last 30 seconds for each minute interval of exercise using standard formulae.<sup>14</sup>  $\dot{V}_E$  and  $V_T$  were expressed at BTPS,  $\dot{V}_{O_2}$  and  $\dot{V}_{CO_2}$  were expressed at STPD. Predicted peak  $\dot{V}_{O_2}$  during exercise was calculated as<sup>14</sup>:

$$\text{Peak } \dot{V}_{O_2} = 0.83 \text{ ht}^{2.7} \times (1 - 0.007 \text{ age}) \times (1 - 0.25S)$$

where height (ht) is in metres and S is a factor taking account of gender (S = 0 for men and 1 for women). Predicted peak HR was calculated as<sup>14</sup>:

$$\text{Peak HR} = 210 - 0.65 \text{ age (years)}$$

Sense of dyspnoea was assessed by the modified Borg scale<sup>15</sup> at the end of exercise. All subjects were shown the Borg scale before exercise and familiarised with the scale. Subjects were then asked at the end of exercise: "How difficult is your breathing?" and asked to point with their finger to a number on the scale.

#### STATISTICAL ANALYSIS

Data collected at rest, at two matched submaximal work rates, and at the end of the exercise test were used in the analysis. The results used for the submaximal work rates were from work rates that most closely approximated 40% and 70% of Wmax from each subject's first exercise test. Comparisons were then made at matched work rates for the remaining studies.

Statistical significance of group mean data from the three experiment days was determined by repeated measures analysis of variance. The variability of subject results for the three experiments was assessed from the coefficient of variation. The coefficient of variation was derived by dividing the standard deviation by the mean.<sup>16</sup> Analysis of the Borg scale was performed using Wilcoxon's signed rank test<sup>17</sup>;  $p < 0.05$  was considered significant. The results are shown as mean (SD).

Analysis of our data showed that the sample size was sufficient to detect a 10% increase (0.14 l/min) in peak  $\dot{V}_{O_2}$  between the first and third exercise studies with 91% power.<sup>18</sup>

#### Results

All subjects completed each exercise test without any complications and no exercise test was terminated by the physician. All data collected for the study were used in the analysis and no subjects were excluded.

Table 2 Group mean (SD) values obtained at end of exercise

	Test 1	Test 2	Test 3	p*
Exercise duration (seconds)	433 (148)	434 (137)	444 (139)	NS
Work rate (W)	118 (36)	120 (30)	120 (38)	NS
$\dot{V}O_2$ (l/min)	1.39 (0.37)	1.41 (0.32)	1.39 (0.34)	NS
$\dot{V}CO_2$ (l/min)	1.71 (0.49)	1.77 (0.37)	1.72 (0.43)	NS
$\dot{V}E$ (l/min)	60.9 (18.2)	65.3 (17.8)	64.9 (17.1)	NS
$V_T$ (l)	1.7 (0.2)	1.7 (0.3)	1.7 (0.3)	NS
f (breaths/min)	36 (9)	39 (12)	38 (10)	NS
O <sub>2</sub> saturation (%)	86 (5)	85 (6)	83 (3)	NS
HR (beats/min)	148 (9)	150 (5)	146 (12)	NS
Borg scale	4.0 (0.6)	4.8 (1.5)	4.7 (1.5)	NS

\*p value for comparison of tests 1, 2, and 3 by analysis of variance;  $\dot{V}O_2$ —oxygen uptake;  $\dot{V}CO_2$ —carbon dioxide output;  $\dot{V}E$ —minute ventilation;  $V_T$ —tidal volume; f—respiratory frequency; HR—heart rate.

Figure 1 Mean (SD) values of the coefficient of variation for  $\dot{V}O_2$ ,  $SAO_2$ , and HR measured at rest, at end of exercise, and at submaximal work rates.

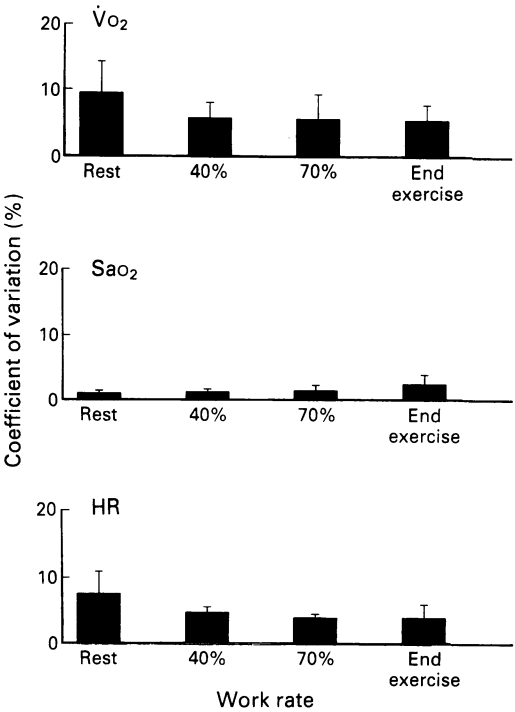
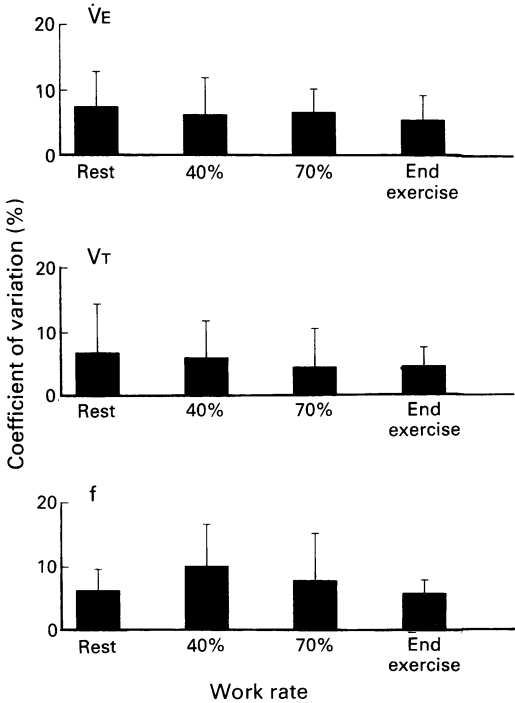


Figure 2 Mean (SD) values of the coefficient of variation for  $\dot{V}E$ ,  $V_T$ , and f measured at rest, at end of exercise, and at submaximal work rates.



**RESTING RESULTS**  
Analysis of group mean data collected at rest, immediately before exercise, revealed that spirometric values were not significantly different on any of the study days. The coefficient of variation for FEV<sub>1</sub> was 3.5% (range 0.4–10.8%) and for FVC was 3.4% (range 0.6–9.0%). The mean resting SaO<sub>2</sub> was 95% (range 90–99%). In addition, values obtained at rest for  $\dot{V}O_2$ ,  $\dot{V}CO_2$ ,  $\dot{V}E$ ,  $V_T$ , f, SaO<sub>2</sub>, and HR were also not significantly different among the three separate studies.

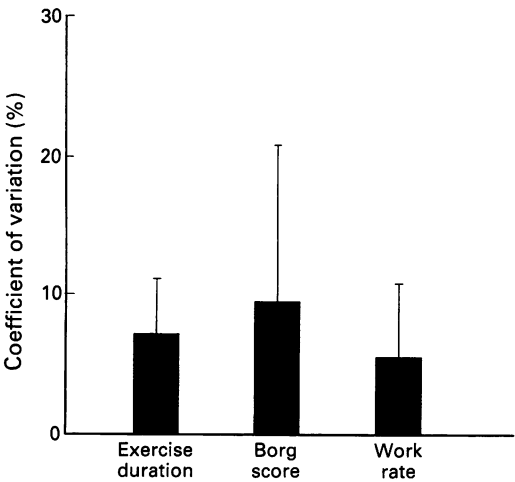
**EXERCISE RESULTS**  
Exercise performance was impaired in all subjects. The mean (SD) peak  $\dot{V}O_2$  was 55% (10%) (range 41–68%) of the age predicted normal value. The mean (SD) maximal HR was 81% (5%) of predicted. All subjects showed significant oxygen desaturation (mean (SD) 11% (5%), range 6–21%). The ratio of peak  $\dot{V}E$  (mean of each subject's peak  $\dot{V}E$  obtained from the three exercise studies) to FEV<sub>1</sub> was 28 (7) for the group (range 17–35). Four subjects discontinued exercise because of leg fatigue, while two subjects stopped because of dyspnoea. Table 2 lists group mean data collected at the end of the exercise test. There was no significant difference in exercise duration, work rate, or the peak  $\dot{V}O_2$  at end of exercise among the three studies.  $\dot{V}CO_2$ ,  $\dot{V}E$ ,  $V_T$ , f, SaO<sub>2</sub>, HR, and Borg scores also did not differ among the three studies at end of exercise. In addition, there was no significant difference in these variables at 40% Wmax and 70% Wmax. The mean work rate at 40% Wmax was 43 W and at 70% Wmax was 80 W.

**EXERCISE VARIABILITY**  
The mean within subject coefficients of variation for  $\dot{V}O_2$ ,  $SAO_2$ , and HR as a function of work rate (% Wmax) are presented in fig 1. Values obtained at rest and at end of exercise, as well as those measured at 40% Wmax and 70% Wmax, are shown. The mean within subject coefficients of variation at end of exercise for  $\dot{V}O_2$  was 5.3% (5.3–9.5% throughout exercise), for  $SAO_2$  was 2.5% (1.1–2.5% throughout exercise), and for HR was 4.0% (4.0–7.6% throughout exercise). The mean within subject coefficient of variation for the decrease in  $SAO_2$  during exercise was 17.3% (not shown in fig).

Figure 2 shows the mean within subject coefficients of variation for  $\dot{V}E$ ,  $V_T$ , and f, also as a function of work rate. The mean within subject coefficients of variation for  $\dot{V}E$  at end of exercise was 5.5% (5.5–7.5% throughout exercise), for  $V_T$  was 4.6% (4.4–6.8% throughout exercise), and for f was 5.8% (5.8–10.1% throughout exercise).

Figure 3 shows the mean within subject coefficients of variation obtained at end of exercise for exercise duration, the Borg score, and for work rate. The mean within subject coefficient of variation for exercise duration was 7.2% (range 2.7–12.3%), for Borg scale was 9.5% (range 0–26.9%), and for work rate was 5.6% (range 0–13.6%). Individual results

Figure 3 Mean (SD) values of the coefficient of variation for exercise duration, Borg score, and work rate measured at end of exercise.



for  $\dot{V}E$  versus  $\dot{V}O_2$  in the three tests are plotted in fig 4 and show that the relationship between  $\dot{V}E$  and  $\dot{V}O_2$  is very similar in each subject's three tests.

Discussion

The reproducibility of clinical exercise testing in patients with restrictive lung disease has not previously been examined. However, numerous studies have examined the reproducibility in normal subjects,<sup>19 20</sup> in patients with congestive heart failure,<sup>21 22</sup> and in patients with chronic airflow limitation.<sup>23-30</sup> Garrard and Emmons<sup>19</sup> found that values obtained during exercise testing in normal subjects were highly reproducible, although some diurnal variation was seen. On the other hand, in patients with stable congestive heart failure Elborn *et al*<sup>21</sup> found that exercise time increased by 20% with repeated testing.

In patients with chronic airflow limitation undergoing repeated walking tests Knox and coworkers<sup>30</sup> found a significant increase in walking distances with repeated testing. Swinburn *et al*,<sup>25</sup> in patients with severe

chronic airflow limitation, found a 29% increase in performance with cycle ergometer testing between the first and fourth exercise tests. Conversely, Brown *et al*,<sup>26</sup> Cox and coworkers,<sup>29</sup> Silverman and coworkers,<sup>27</sup> and Muza *et al*<sup>28</sup> found no significant learning effect with repeated exercise testing, and variability in measured variables ranged from 3% to 12%. More recently, Belman *et al*<sup>24</sup> found that, while the Borg scale ratings decreased with repeated testing, other variables stabilised after one or two practice attempts. However, except for the Borg ratings, only heart rate showed a significant change between the first and second tests, and heart rate showed no significant change between the second, third and fourth tests. Furthermore, no significant changes in oxygen uptake, minute ventilation, tidal volume, or respiratory frequency were seen between the four exercise studies, so the data appear to suggest that the reproducibility was good and no learning effect, even after the first study, was noted in these variables.

In view of the conflicting results in the literature for patients with other diseases, an accurate assessment of the variability of exercise testing in patients with restrictive lung disease was needed. Our results, in clinically stable patients who have not undergone previous cycle ergometer exercise testing, suggest that reproducibility is high. The reproducibility we noted during exercise was similar to the reproducibility we found in spirometric tests obtained before exercise. We also found no significant learning effect between the three studies, suggesting that practice exercise studies are not necessary in patients with restrictive lung disease even if they have not previously exercised on a cycle ergometer. Our protocol approximates the environment of clinical exercise testing laboratories and therefore enables the results to be clinically applicable.

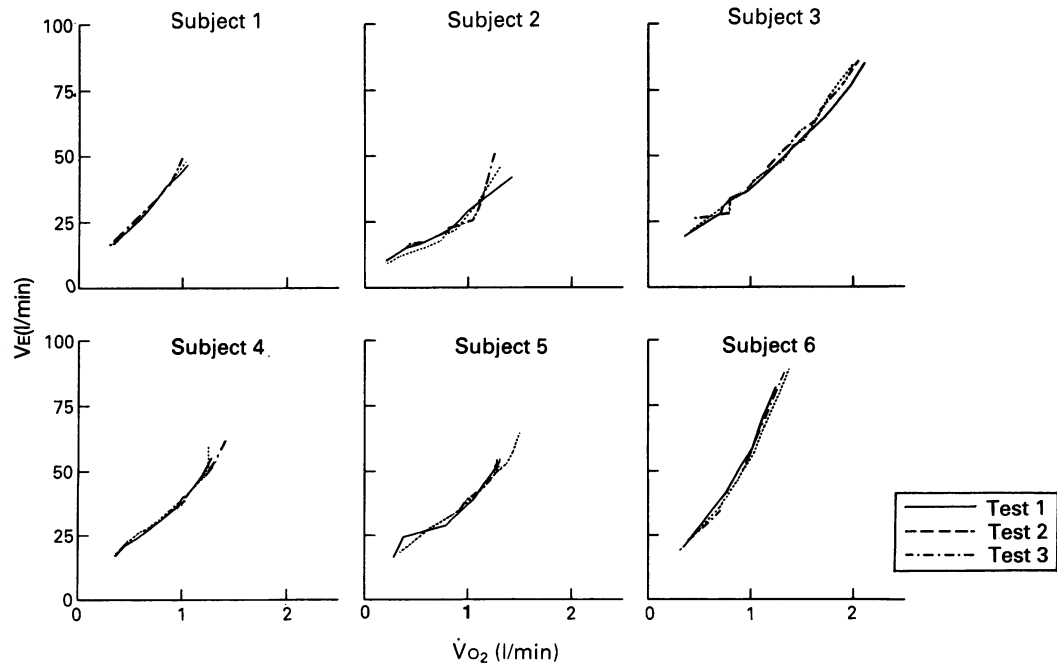


Figure 4 Individual results for  $\dot{V}E$  versus  $\dot{V}O_2$  in the three tests.



While measurements at maximal exercise are important, it is often equally important to examine variables measured during submaximal exercise, especially since many patients will stop exercise well below expected maximum values.<sup>14</sup> Comparison of our results, measured at 40% Wmax and 70% Wmax, shows that reproducibility at these submaximal levels of exercise parallels the reproducibility of measurements made at end of exercise. This suggests that, in addition to generating reproducible end of exercise measurements, measurements made at submaximal levels of exercise may also be used appropriately in evaluating the results of clinical exercise testing in these patients.

From this study we have found that measurements during clinical exercise testing in subjects with restrictive lung disease are highly reproducible and show no learning effect with successive studies. Clinical exercise testing can be reliably used to diagnose and monitor patients with restrictive lung disease.

We wish to thank Ms Candace Leuschen for her assistance in the preparation of the manuscript. The study was supported by the Saskatchewan Lung Association and by the Saskatchewan Health Research Board. Dr Marciniuk and Dr Gallagher are Saskatchewan Lung Association Scholars.

- 1 Kelly MA, Daniel RP. Exercise testing in interstitial lung disease. *Clin Resp Med* 1984;5:145-56.
- 2 Epler GR, Saber FA, Gaensler EA. Determination of severe impairment (disability) in interstitial lung disease. *Am Rev Respir Dis* 1980;121:647-59.
- 3 Spiro SG, Dowdeswell IRG, Clark TJH. An analysis of submaximal exercise responses in patients with sarcoidosis and fibrosing alveolitis. *Br J Dis Chest* 1981;75:169-80.
- 4 Bye PTP, Anderson SD, Woolcock AJ, Young IH, Alison JA. Bicycle endurance performances of patients with interstitial lung disease breathing air and oxygen. *Am Rev Respir Dis* 1982;126:1005-12.
- 5 Burdon JGW, Killian KJ, Jones NJ. Pattern of breathing during exercise in patients with interstitial lung disease. *Thorax* 1983;38:778-84.
- 6 Jones NL, Rebuck AS. Tidal volume during exercise in patients with diffuse fibrosing alveolitis. *Bull Eur Physio-pathol Respir* 1979;15:321-7.
- 7 Denison D, Al-Hillawi H, Turton C. Lung function in interstitial lung disease. *Semin Respir Med* 1984;6:40-54.
- 8 Keogh BA, Lakatos E, Price D, Crystal RG. Importance of the lower respiratory tract in oxygen transfer. *Am Rev Respir Dis* 1984;129(Suppl):S76-80.
- 9 Gardner RM, Hankinson JL, Clauson JL, Crapo RO, Johnson RL Jr, Epler GR. American Thoracic Society standardization of spirometry—1987 update. *Am Rev Respir Dis* 1987;136:1285-98.
- 10 Morris JF, Koski A, Johnson LC. Spirometric standards for healthy non-smoking adults. *Am Rev Respir Dis* 1971;103:57-67.
- 11 Goldman HI, Becklake MR. Respiratory function tests: normal values at median altitudes and the prediction of normal results. *Am Rev Tuberc Pulm Dis* 1959;79:457-67.
- 12 Burrows B, Kasik JE, Niden AH, Barclay WR. Clinical usefulness of the single-breath pulmonary diffusing capacity test. *Am Rev Respir Dis* 1961;84:789-806.
- 13 Gowda K, Zintel T, McParland C, Orchard R, Gallagher CG. Diagnostic value of maximal exercise tidal volume. *Chest* 1990;98:1351-4.
- 14 Jones NL. *Clinical exercise testing*. Philadelphia: Saunders, 1988.
- 15 Borg G. Psychophysical bases of perceived exertion. *Med Sci Sports Exerc* 1982;14:377-81.
- 16 Snedecor GW, Cochran WG. *Statistical methods*. 8th edn. Ames, Iowa: Iowa State University Press, 1989:26-7.
- 17 Dixon WJ, Massey FJ. *Introduction to statistical analysis*. 4th edn. New York: McGraw-Hill, 1983:391-2.
- 18 Snedecor GW, Cochran WG. *Statistical methods*. 8th edn. Ames, Iowa: Iowa State University Press, 1989:38-53.
- 19 Garrard CS, Emmons C. The reproducibility of the respiratory responses to maximum exercise. *Respiration* 1986;49:94-100.
- 20 Wilson RC, Jones PW. Long-term reproducibility of Borg scale estimates of breathlessness during exercise. *Clin Sci* 1991;80:309-12.
- 21 Elborn JS, Stanford CF, Nicholls DP. Reproducibility of cardiopulmonary parameters during exercise in patients with chronic cardiac failure. The need for a preliminary test. *Eur Heart J* 1990;11:75-81.
- 22 Janicki JS, Gupta S, Ferris ST, McElroy PA. Long-term reproducibility of respiratory gas exchange measurements during exercise in patients with stable cardiac failure. *Chest* 1990;97:12-7.
- 23 Owens MW, Kinasewity GT, Strain DS. Evaluating the effects of chronic therapy in patients with irreversible air-flow obstruction. *Am Rev Respir Dis* 1986;134:935-7.
- 24 Belman MJ, Brooks LR, Ross DJ, Mohsenifar Z. Variability of breathlessness measurement in patients with chronic obstructive pulmonary disease. *Chest* 1991;99:566-71.
- 25 Swinburn CR, Wakefield JM, Jones PW. Performance, ventilation, and oxygen consumption in three different types of exercise test in patients with chronic obstructive lung disease. *Thorax* 1985;40:581-6.
- 26 Brown SE, Fisdor CE, Steinsburg DW, Light RW. Reproducibility of  $\dot{V}_{O_2}$ max in patients with chronic air-flow obstruction. *Am Rev Respir Dis* 1985;131:435-8.
- 27 Silverman M, Barry J, Hellerstein H, Janos J, Kelsen S. Variability of the perceived sense of effort in breathing during exercise in patients with chronic obstructive pulmonary disease. *Am Rev Respir Dis* 1988;137:206-9.
- 28 Muza SR, Silverman MT, Gilmore GC, Hellerstein HK, Kelsen SG. Comparison of scales used to quantitate the sense of effort to breathe in patients with chronic obstructive pulmonary disease. *Am Rev Respir Dis* 1990;141:909-13.
- 29 Cox NJM, Hendriks JCM, Binkhorst RA, Folgering HT, van Herwaarden CLA. Reproducibility of incremental maximal cycle ergometer tests in patients with mild to moderate obstructive lung diseases. *Lung* 1989;167:129-33.
- 30 Knox AJ, Morrison JFJ, Muers MF. Reproducibility of walking test results in chronic obstructive airways disease. *Thorax* 1988;43:388-92.