

The reproducibility of symptoms during a submaximal exercise test in chronic heart failure

S. Grant,¹ J. McMurray,² T. Aitchison,³ T. McDonagh,² J. Christie,² A. Davie² & H. J. Dargie²

¹Institute of Biomedical and Life Sciences and ³Statistics Department, University of Glasgow, and ²Department of Cardiology, Western Infirmary, Glasgow

Aims The aim of this study was to evaluate the use of a submaximal test with a symptom limited endpoint and to measure the reproducibility of symptoms in patients with CHF.

Methods Ten patients with chronic heart failure were studied. Based on two maximal treadmill tests an individual protocol using a constant work rate at a submaximal intensity was derived. The projected maximum treadmill time for the constant workrate was between 8 and 17 min. Tests were carried out 1, 2, 4 and 6 weeks after the maximum tests. Every 2.5 min during the submaximal test patients recorded their symptoms of breathlessness and fatigue using computer automated visual analogue (VAS) and Borg CR10 scales. The measure of reproducibility used was the proportion of total variability explained by the between subject variability.

Results Using the VAS scale, general fatigue was reasonably reproducible ranging from 77–86%. For VAS breathlessness reproducibility ranged from 66% to 83%. Reproducibility for breathlessness and fatigue for the Borg CR10 scale was much lower than the VAS scale. Reproducibility for the treadmill times was 51% but increased to 76% if one test of one subject was excluded.

Conclusions The use of the VAS during submaximal exercise offers a useful means of evaluating symptoms in CHF and potentially their response to treatment. These findings show that individual submaximal protocols can be easily prescribed for CHF patients. Using such an approach, clinically desirable tests lasting around 12 min can be developed. These tests are reasonably reproducible and may provide a useful means of assessing patient disability and the impact of treatment.

Keywords: chronic heart failure, exercise testing, and symptoms

Introduction

Patients with chronic heart failure (CHF) experience dyspnoea and general fatigue during their everyday activities. Quantification of these symptoms during submaximal exercise may provide a useful means of measuring patient disability and the impact of drug treatment. It is important that the clinician is able to assess the patient's symptoms. For such an approach to be of value, the measures used need to be reproducible.

Many studies rely on subjective measures to evaluate outcomes including quality of life, dyspnoea and fatigue. Major debate has centred around the selection of the optimal response options. A number of subjective scales (Visual Analogue Scale (VAS)) and various versions of Borg scales have been widely used to assess breathlessness, rate of perceived exertion and fatigue in a normals and a range of patient groups [1–6]. No study has examined the reproducibility of symptoms (breathlessness and general fatigue) in CHF patients.

Poole-Wilson [7] concluded in his review of exercise testing in CHF that 'there is a need for a simple exercise test which can be used to evaluate symptoms of patients

with CHF and to test the efficacy of new drugs as they become available'. He advocated that the test must be simple, cheap, reproducible and sensitive. In addition, he stressed the need for tests which relate to everyday life.

A treadmill test which simulates daily activities, can measure symptoms and has a defined endpoint, and is reproducible would be of great benefit in the evaluation and monitoring of CHF patients.

The aim of this study was to evaluate the use of a submaximal test with a symptom limited endpoint and to measure the reproducibility of symptoms in patients with CHF.

Methods

Subjects

Ten male subjects aged 59.9 ± 7.9 years (mean \pm s.d.) with chronic heart failure (NYHA 1 and 2) were studied. All subjects gave informed consent for the study which had been approved by the West Ethical Committee.

Research design

Subjects underwent two maximal incremental treadmill tests at least 4 days apart. From the longer of the two tests, a

Correspondence: Dr S. Grant, Institute of Biomedical and Life Sciences, University of Glasgow, 64 Oakfield Avenue, Glasgow, G12 8LT, U.K.

submaximal workrate was calculated (see below). The subjects performed four submaximal tests at weeks 1, 2, 4 and 6 after the second maximal test.

Incremental tests

The maximal incremental tests started at 1.4 mph at 0% gradient with progressions in speed or gradient until volitional exhaustion.

Constant work rate test

The predetermined speed and gradient were set on the treadmill for the constant work rate test by calculating 80% of the predicted oxygen cost of the stage attained in the longer of the two incremental tests. Calculation was based on the American College of Sports Medicine [8] equations. This procedure results in a workrate approximately two stages below that attained in the incremental protocol. The subjects were asked to walk for as long as possible.

Equipment—Treadmill, ECG and gas analysis

Respiratory variables were determined during exercise by an automated gas analysis system (Beckman metabolic measurement cart; classic exercise model system 2). Gas collection and analysis were continuous and respiratory values for every 30 s were given on a print-out. A marquette MAC2 treadmill and ECG console were used. Heart rate was taken from the electrocardiogram during the last 10 s of each minute.

Symptom scales.

During each test, symptom scales (see below) were administered every 2.5 min. Each scale was used twice at each time point i.e. to measure breathlessness and general fatigue. The symptom of breathlessness was always measured before general fatigue.

Equipment—symptom scales

Each scale displayed on a colour television screen in front of the subject while he exercised on the treadmill. The subject recorded his response by means of finger controls and the information was stored in a BBC computer. An audible prompt was given each time a new scale appeared

on the screen. On each occasion the subjects had to move the lever before the cursor appeared on the screen i.e. before any score was visible (thus the previous score was not displayed when the new scale was presented).

Symptom scales

The visual analogue scale [3] consisted of a horizontal line. At the left hand side of the scale the word 'none' was labelled and at the far right the word 'very severe' was placed. The Borg CR10 scale [3] consisted of a vertical scale labelled 0–10 with verbal descriptors at various numbers on the scale. The scales had a heading of either breathlessness or general fatigue. Subjects indicated their level of breathlessness or general fatigue by moving the lever on the treadmill which adjusted the score on the TV monitor. Once the subject had chosen the desired score, he pressed the button to record this in the computer. At this point another scale appeared. Subjects were randomly assigned to a VAS/Borg CR10 or a Borg CR10/VAS sequence of presentation.

Breathlessness was described as, 'an uncomfortable need to breathe'.

General fatigue was described as, 'overall tiredness, overall fatigue'.

Statistical analysis

Statistical analysis was limited to the first four timepoints i.e. 2.5, 5.0, 7.5 and 10 min as there were insufficient data after 10 min to allow for meaningful analysis. A generalised linear model (GLM) was applied to the data to include the possible effects of visit, subject, time and natural (or error) variability. A Bonferroni multiple comparisons procedure was used to assess where the significant differences, if any, lay.

Results

The reproducibility of the subjective scales

The estimated reproducibility coefficients for VAS and Borg CR10 for the first four timepoints are given in Table 1. For both breathlessness and general fatigue the reproducibility of the VAS appear to be higher than the Borg CR10. Ninety-five percent confidence intervals, however, suggested that there were no significant differences between the scales. For the VAS (breathlessness), the reproducibility appeared to decrease through time into the test (i.e. the coefficients

Scale	Reproducibility Time (min)			
	2.5	5.0	7.5	10.0
<i>Breathlessness</i>				
VAS	82% (58, 91)	83% (61, 92)	68% (36, 83)	66% (22, 78)
Borg CR10	51% (16, 73)	48% (13, 70)	46% (11, 69)	38% (-0, 57)
<i>General fatigue</i>				
VAS	77% (51, 89)	86% (66, 93)	83% (60, 92)	81% (48, 82)
Borg CR10	51% (17, 73)	77% (50, 89)	64% (32, 81)	46% (6, 64)

Confidence intervals are in brackets beneath each estimate of reproducibility.

Table 1 Reproducibility coefficients and 95% confidence intervals for visual analogue and borg scales for breathlessness and general fatigue.

were lowest at 7.5 and 10 min). In general, reproducibility was better for the VAS on general fatigue than on breathlessness. There were no significant visit effects.

Incremental tests

The peak VO_2 values (mean \pm s.d.) for the two maximal tests were 20.6 ± 4.6 and 18.8 ± 4.2 ml kg^{-1} min^{-1} for test 1 and test 2 respectively. There was no significant difference between the tests. Similarly, there was no difference for endurance time (secs) between the tests, with the test 1 time being 909 ± 158 and a test 2 time of 869 ± 148 .

Constant work rate test

The endurance times for the four tests are given in Table 2. The estimated reproducibility coefficient was 51%. If test 1 of subject 6 is removed the reproducibility coefficient rises to 76%. (The endurance time for subject 6 increased from 6 min to 17 min between tests 1 and 2. There were no significant visit/familiarisation effects.

Discussion

Symptom scales

In this study the VAS tended to be more reproducible than the Borg CR10 for breathlessness and general fatigue. However, 95% confidence intervals showed no significant difference between the scales. Reproducibility coefficients were highest with the VAS scale measuring general fatigue.

Measurement of breathlessness

In a pilot study, using three 6 min stages at relative intensities ranging from 60 to 83% of peak VO_2 , it was found that the higher the relative intensity, the higher the reproducibility coefficient. In this main study, the mean % of peak VO_2 was 89% and the mean % peak VO_2 on the third stage in the CHF pilot study was 83%. Reproducibility coefficients at the the highest pilot study intensities were 82% for the VAS and 72% for the Borg CR10 scales. While the VAS tended to be more reproducible than the Borg CR10 in the main study, there was hardly any difference between the scales in the pilot study. Using normals to compare the reproducibility of the VAS, Borg CR10 and Likert scales, Grant *et al.* [1] found the VAS was clearly more reproducible than the Borg CR10 scale—the reproducibility coefficient for the VAS was 78% for breathlessness at the third timepoint. In the early stages of the test, in this study, the reproducibility coefficients compare favourably with the

normals study, giving reproducibility coefficients of over 80%. Several groups have shown that the VAS and the Borg CR10 scale allow reproducible measurement of breathlessness, in the short term, in both normal subjects and patients. [3, 5–7, 9]. Due to different exercise protocols and statistical methods of evaluation, it is impossible to say from these studies whether one scale is more reproducible than the other.

Measurement of general fatigue

In the pilot study and this main study, the VAS tended to be more reproducible than the Borg CR10 scale. Similarly the VAS scale tended to be higher than the Borg CR10 in the VAS Borg CR10 and Likert comparison using normals [1]. In this CHF study, the reproducibility coefficients for VAS were very consistently high throughout the first 10 min whereas the reproducibility coefficients for the Borg CR10 scale showed a large fluctuation in values. No other reproducibility study has been carried out on general fatigue.

Visit effect

Analysis showed that there was no visit effect over the four tests for the VAS and the Borg CR10 scales. In this study the subjects were given two maximal incremental tests but only one maximal test was given in the comparison of the VAS, Borg CR10 and Likert scales [1]. With only one maximal test, the 'normals' study showed a significant visit effect in both VAS scales and for Likert (breathlessness) in the subsequent submaximal tests i.e. the perception of breathlessness and general fatigue decreased over four tests over time for VAS (breathlessness and general fatigue) and Borg CR10 (general fatigue). There were no significant differences between the incremental test 1 and test 2 values for any subjective scale in this CHF study. Thus, experience gained in test 1 does not appear to influence the perception of breathlessness and general fatigue in test 2. The lack of learning effect suggests that the use of the VAS offers a stable baseline from which assessment of changes in patient status may be assessed.

Test time for the submaximal test

One of the aims of this study was to establish a method of predicting a treadmill endurance time between 8–17 min. In out of 40 tests this aim was achieved on 34 occasions (85%). Three tests were over 17 min and three tests were under 8 min. The three tests under 8 min were between 6 and 8 min and the three tests above 17 min were between 18 and 22 min. Thus, this method appears to be applicable in a clinical setting.

The method of selecting the constant workrate for each subject is not based on any scientific formula. This method was developed on a trial and error basis with a number of patients. Prior to the CHF pilot study the attempted use of a ventilatory threshold or an arbitrary relative intensity gave wide-ranging results. It may well be that there is no method based on a scientific formula which will allow an appropriate timescale to be selected. The results from the two incremental maximum tests show that there was no increase

Table 2 Endurance times (mean, s.d. minimum and maximum for the four constant workrate tests in seconds.

Test	Mean	s.d.	Minimum	Maximum
1	609	159	395	942
2	715	218	492	1009
3	706	295	407	1272
4	754	214	542	1295

in endurance time between the two tests. These findings indicate that two tests are sufficient to provide a stable platform from which a submaximal work rate can be selected. Selection of the work rate using 80% of the predicted energy cost has sound clinical application. Incremental tests on CHF patients are regularly carried out without gas collection. Therefore it is important to establish a method of prescribing a treadmill speed and gradient using non ventilatory parameters.

Previous studies have attempted to use sub-maximal tests in CHF to a symptom limited endpoint with a mean endurance time ranging between 6 and 24 min but some authors reported a very large standard deviation [10–12].

Conclusions

Overall the VAS tended to be more reproducible than the Borg CR10 scale for breathlessness and general fatigue. The method of selecting the workrate for the submaximal test (to elicit an endurance time of between 8–17 min) is appropriate for most patients. The use of the above exercise protocol with the VAS scale offers a useful means of evaluating symptoms in CHF.

References

- 1 Grant S, Henderson E, Abdullah I, *et al.* Is there a reproducible and sensitive method of measuring symptoms during exercise in humans? (abstract) *Eur Heart J* 1992; **13**: 118.
- 2 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 Resp Dis* 1990; **141**: 909–913.
- 3 Wilson RC, Jones PW. A comparison of the visual analogue scale and modified Borg scale for the measurement of dyspnoea during exercise. *Clin Sci* 1989; **76**: 277–282.
- 4 Carton RL, Rhodes CR. A critical review of the literature on ratings scales for perceived exertion. *Sports Med* 1985; **2**: 198–222.
- 5 Stark RD, Gambles SA, Chatterjee SS. An exercise test to assess clinical dyspnoea: estimation of reproducibility and sensitivity. *Br J Dis of Chest* 1982; **76**: 269–278.
- 6 Stark RD, Gambles SA, Lewis JA. Methods to assess breathlessness in healthy subjects: a critical evaluation and application to analyse the acute effects of diazepam and promethazine on breathlessness induced by exercise or by exposure to raised levels of carbon dioxide. *Clin Sci* 1981; **61**: 429–439.
- 7 Poole-Wilson PA. Exercise as a means of assessing heart failure and its response to treatment. *Cardiology* 1989; **76**: 347–356.
- 8 American College of Sports Medicine. *Guidelines for Exercise Testing and Prescription*. Lea and Febiger, Philadelphia. 1986; 161–162.
- 9 Wilson RC, Jones PW. Long-term reproducibility of Borg-scale estimates of breathlessness during exercise. *Clin Sci* 1991; **80**: 309–312.
- 10 Koch J, Broustet JP. The benefit of graded physical exercise in chronic heart failure. *Proceedings of the Vth World Congress on Cardiac Rehabilitation*, Ed Broustet JP. Intercept Ltd., Andover, Hampshire, 1993.
- 11 Cowley AJ, Stainer K, Rowley JM, Hampton JR. Abnormalities of the peripheral circulation and respiratory function in patients with severe heart failure. *Br Heart J* 1986; **55**: 75–80.
- 12 Sullivan MJ, Higginbotham B, Cobb FR. Increased exercise ventilation in patients with chronic heart failure: intact ventilatory control despite hemodynamic and pulmonary abnormalities. *Circulation* 1988; **77**: 552–559.

(Received 12 March 1997,
accepted 11 November 1997)