Exercise testing to stratify risk in aortic stenosis

Paul Das*, Helen Rimington, and John Chambers

Cardiothoracic Centre, Guy’s and St Thomas’ Hospitals, London, UK

Received 5 June 2004; revised 21 February 2005; accepted 3 March 2005; online publish-ahead-of-print 8 April 2005

See page 1252 for the editorial comment on this article (doi:10.1093/eurheartj/ehi295)

KEYWORDS
Aortic stenosis; Treadmill exercise testing; Prognosis

Aims The aims of this study were to assess the accuracy of exercise testing in predicting symptom onset within 12 months in patients with asymptomatic aortic stenosis and to establish the criteria that define a positive test.

Methods and results A total of 125 patients with aortic stenosis [effective orifice area (EOA) 0.9 ± 0.2 cm²] were assessed by Specific Activity Scale (SAS) classification, transthoracic echocardiography, and treadmill exercise testing using the modified Bruce protocol. During follow-up, 36 patients (29%) developed spontaneous symptoms within 12 months. Of these, 26 (72%) had had symptoms revealed by exercise testing and 24 (67%) had severe stenosis (EOA ≤ 0.8 cm²). Exercise-limiting symptoms were the only independent predictors of outcome at 12 months, and an abnormal blood pressure response or ST segment depression did not improve the accuracy of the exercise test. The positive predictive accuracy for exercise-induced symptoms was 57% in the whole population and 79% for patients aged <70 in SAS Class I. The negative predictive accuracy was 87% in the whole population and 86% in the subgroup.

Conclusion A significant proportion of patients with apparently asymptomatic aortic stenosis experience limiting symptoms on treadmill exercise testing. The subsequent development of spontaneous symptoms is strongly related to the severity of stenosis and to limiting symptoms on exercise testing, but less so to an abnormal blood pressure response or ST segment depression.

Introduction
The management of patients with asymptomatic aortic stenosis remains controversial. Aortic valve replacement is not normally recommended before the onset of symptoms, because the associated risks are believed to outweigh those of asymptomatic aortic stenosis. However, the overall risk of death before the aortic valve can be replaced remains significant. Sudden death is reported in up to 6% of overall risk of death before the aortic valve can be replaced in patients with asymptomatic severe aortic stenosis, although it is usually <1% per year. In addition, 3% of patients may die within 3–6 months after the onset of symptoms and a mortality rate of 6.5% has been reported in symptomatic patients awaiting valve replacement. Ideally, therefore, surgical referral should be made just before the onset of symptoms.

Recent prospective studies have shown that the severity of stenosis on echocardiography and the rate of progression of stenosis predict the need for valve replacement. However, these factors may themselves influence referral decisions: for example, Otto et al. found that 44% of patients underwent surgery in the absence of symptoms.

Exercise testing might be a better predictor of clinical outcome although there is uncertainty about which specific parameters are most useful. Therefore, we tested the ability of exercise time, electrocardiographic changes, abnormal blood pressure response, and symptoms invoked during treadmill exercise testing to predict the onset of spontaneous symptoms within 12 months. Spontaneous symptoms are the best-validated criteria for surgery, and we chose a period of 12 months as this provides a reasonable period for discussion and invasive investigation and allows for waiting lists that still exist in many countries.

Methods
Patients
A total of 125 patients (85 males and 40 females) were recruited from the echocardiography department between August 1996 and December 2001. The mean age was 65 years (56–74 years). All met the inclusion criteria of aortic valve thickening and effective orifice area (EOA) < 1.4 cm², normal left ventricular systolic function (defined by a fractional shortening >28% and no regional wall motion abnormality), no more than mild aortic regurgitation, and no other significant valve disease. Patients with known pulmonary disease were excluded. Aortic stenosis was graded according to continuity EOA at rest by commonly used arbitrary criteria as mild (area > 1.2 cm²), moderate (area 0.8–1.2 cm²), or severe (area <0.8 cm²).

All were assessed for exertional symptoms by both clinical history and a structured written questionnaire. Each patient also completed a Specific Activity Scale (SAS) questionnaire of daily activity. This grades exercise ability into four classes and has been validated...
against treadmill exercise capacity. All of the patients denied symptoms and none had been referred for aortic valve replacement. Exercise testing may be an unfamiliar and ‘supra-physiologic’ stressor for some patients. To test whether the exercise test might better predict the onset of spontaneous symptoms in patients who were younger and more active, subgroup analysis was performed for 58 patients aged <70 and SAS Class I. The study was approved by the local Ethical Committee and all patients gave written, informed consent.

**Echocardiography**

An ATL HDI 3000 (Seattle, WA, USA) was used with a 3–2, 20 mm duplex probe and 1.9 MHz continuous wave stand-alone probe. M-mode recordings were made at a level immediately apical to the tips of the mitral valve leaflets and end-diastolic measurements were made using the American Society of Echocardiography convention. Fractional shortening was calculated as the systolic decrease in left ventricular long axis excursion measured in M-mode from the apical four-chamber view. The sub-aortic diameter was measured on parasternal long axis frames frozen in systole. The average of three estimates was taken from inner edge to inner edge at the standard level in the cylindrical part of the left ventricle. Pulsed Doppler recordings were made in the apical five-chamber view with the sample volume moved axially from the level of the aortic annulus until a clear non-aliased signal was obtained, usually 0.5–1 cm below the valve. The signal was traced to obtain peak velocity, velocity time integral, and mean pressure difference using the online software. Continuous wave Doppler recordings were made from the apex and right intercostal positions and the optimal angle was determined by dividing each variable into four regular groups by value and plotting the per cent of patients who reached the endpoint in each group. No further model building procedures were used.

**Exercise testing**

Exercise testing was performed with a Quinton Q55xt treadmill and Q5000 monitor (WA, USA) according to ACC/AHA practice guidelines using a Bruce protocol modified by two warm-up stages. A technician unaware of the echocardiographic data supervised the test with a physician in attendance. Subjects were questioned for symptoms every 2 min and the heart rate, blood pressure, and a 12-lead electrocardiogram were recorded at baseline, at the end of each stage and at peak exercise.

The test was defined as positive if it was stopped prematurely because of limiting breathlessness/chest discomfort or dizziness. Each patient was questioned and observed carefully to distinguish significant breathlessness from quickly reversible minor breathlessness. Other predetermined criteria for cessation were ST segment depression of >5 mm measured 80 ms after the J point, more than three consecutive ventricular premature beats, and hypotension (defined as a fall in systolic blood pressure of >20 mmHg from baseline), although in practice no test was terminated for any of these reasons. Otherwise, the test continued until the patient was fatigued.

Total exercise time in seconds and maximum ST depression in mm were recorded. ST depression ≥2 mm in at least one lead was considered significant. Previous studies have used differing criteria to define the blood pressure response to exercise, and there is no universally agreed cut-off point. Therefore, for the purpose of clarity we defined an abnormal blood pressure response as a systolic blood pressure at peak exercise either the same as or below the baseline level.

**Follow-up**

Patients were re-assessed for symptoms at 6 and 12 months with the same written questionnaire. Follow-up data were available for all 125 subjects. The endpoint was the development of spontaneous exertional symptoms or cardiovascular death within 12 months of being studied.

**Statistical analysis**

Results are shown as mean ± standard deviation or median (lower quartile, upper quartile) for skewed data. Data were analysed with GB-stat 8.0 software. The two-sided unpaired t-test and the Mann-Whitney U test were used to compare patients who developed symptoms with those who remained asymptomatic to determine potential predictors of outcome. Categorical variables were compared using the χ² test.

Individual echocardiographic and exercise variables were tested for the ability to predict symptom onset within 12 months by univariate analysis. Variables that predicted symptom onset on univariate analysis were then entered into a multivariate logistic regression model. Categorical variables were entered directly. Continuous variables were first tested for the linearity assumption by dividing each variable into four regular groups by value and plotting the per cent of patients who reached the endpoint in each group. No further model building procedures were used.

Symptom-free survival after exercise testing was also demonstrated using the Kaplan-Meier life table. A significance level of 0.05 was used for comparison and entry into the multivariate logistic regression model.

Using the results for blood pressure rise from Amato et al., a total population size of 43 gave 80% power for detecting a difference between the groups with and without endpoints at P < 0.01. For the ST segment data, a total population size of 83 was necessary.

**Results**

The echocardiographic measures of aortic stenosis are given in Table 1. Defined by EOA, 11 (8%) patients had mild stenosis, 62 (50%) had moderate stenosis, and 52 (42%) had severe stenosis. From the SAS questionnaire, 35 (28%) patients were in SAS Class II and the remaining 90 were in SAS Class I.

All patients completed a satisfactory treadmill exercise test. Two patients experienced self-limiting asymptomatic atrial fibrillation at peak exercise, but there were no other significant adverse events during exercise testing.

Total exercise time was 10.9 ± 3.7 min. Forty-six (37%) patients stopped exercise because of limiting symptoms (28 breathlessness, 12 chest tightness, and six dizziness). The remaining 79 patients stopped because of fatigue. No test was terminated for any other reason. There was no difference in age between patients with limiting symptoms and those stopping with fatigue. Twenty-nine (23%) patients demonstrated an abnormal blood pressure response exercise, defined as a systolic blood pressure at peak exercise either the same as or below the baseline level, and ST segment depression of ≥2 mm in one or more leads occurred in 33 (26%) patients. In four cases, the electrocardiographic changes were uninterpretable because of resting bundle branch block or left ventricular hypertrophy.

**Risk stratification**

All patients were followed for 12 months and, in this time, 36 (29%) patients developed spontaneous symptoms. There were no deaths during the follow-up period. Patients who...
developed symptoms had more severe aortic stenosis on echocardiography, shorter exercise durations, and a higher proportion of positive exercise tests than those that remained symptom-free (Table 1). No patient with an EOA $\geq 1.2\, \text{cm}^2$ at baseline developed symptoms within 12 months. Overall, symptom-free survival at 12 months was 49% for patients with limiting symptoms on exercise testing and 89% for those without (Figure 1).

The predictive ability of the echocardiographic and exercise data for the onset of spontaneous symptoms is given in Table 2. Spontaneous symptoms at 12 months developed in five of six (83%) patients with exertional dizziness, six of 12 (50%) patients with chest tightness, and 15 of 28 (54%) with breathlessness. The sensitivity of exercise-limiting symptoms was 72% and the specificity was 78%. Overall, the absence of limiting symptoms had a negative predictive accuracy of 87% among all patients. An abnormal blood pressure response or ST depression gave no statistically significant benefit above limiting symptoms with respect to predictive accuracy. Only two patients with an abnormal blood pressure response, two patients with ST depression, and one patient with both but without exercise-limiting symptoms subsequently developed spontaneous symptoms.

For the 58 patients aged $<70$ years in SAS Class I, the positive predictive accuracy of limiting symptoms on exercise testing was 79% compared with 57% for the whole population. The negative predictive accuracy of limiting symptoms was maintained at 86% in this group.

For the subgroup of 52 patients with severe aortic stenosis, comparing those who reached an endpoint and those who did not, the exercise time was 7.9 vs. 11.3 min ($P = 0.0003$), the peak transaortic velocity was 4.23 vs. 4.19 m/s ($P = 0.82$), the EOA was 0.64 vs. 0.67 cm$^2$ ($P = 0.43$), and 71 vs. 32% had symptoms revealed by exercise testing ($P = 0.005$). The positive predictive accuracy of limiting symptoms on exercise testing was 65%, and the negative predictive accuracy was 73%. In comparison, there was only a 46% chance of any patient with severe stenosis developing symptoms within 12 months.

A multivariate logistic regression model was fitted with the following variables, all of which were significant on univariate analysis: total exercise time, exercise-limiting symptoms, peak transaortic velocity, EOA, abnormal blood pressure response, and ST segment depression (Table 3). Mean resistance was not included, as it did not satisfy the linearity assumption. Limiting symptoms on exercise testing remained an independent predictor of symptom onset within 12 months. Although peak velocity and EOA are calculated using the same continuous wave Doppler signal and are therefore closely related, neither was a significant predictor when included alone in the multivariate analysis.

### Discussion

In this study, symptoms during exercise testing were superior to clinical history and echocardiography in predicting the imminent onset of spontaneous symptoms. Objective haemodynamic and electrocardiographic measures did not increase predictive accuracy.

**Table 1** Comparison between patients developing symptoms (endpoint) and those remaining asymptomatic at 12 months (no endpoint)

<table>
<thead>
<tr>
<th>Demographic data</th>
<th>Endpoint ($n = 36$)</th>
<th>No endpoint ($n = 89$)</th>
<th>$P$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>70 (56, 75)</td>
<td>67 (56, 73)</td>
<td>0.43*</td>
</tr>
<tr>
<td>Gender (male/female)</td>
<td>23/13</td>
<td>62/27</td>
<td>0.53**</td>
</tr>
<tr>
<td>SAS Class (I/II)</td>
<td>24/12</td>
<td>72/17</td>
<td>0.087**</td>
</tr>
<tr>
<td>Echocardiographic data</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Peak velocity (m/s)</td>
<td>4.1 ± 0.6</td>
<td>3.7 ± 0.8</td>
<td>0.0004</td>
</tr>
<tr>
<td>Mean pressure drop (mmHg)</td>
<td>42.9 ± 14.5</td>
<td>33.2 ± 15.9</td>
<td>0.002</td>
</tr>
<tr>
<td>Effective orifice area (cm$^2$)</td>
<td>0.73 ± 0.16</td>
<td>0.94 ± 0.25</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Resistance (dyne s cm$^{-5}$)</td>
<td>240 (186, 276)</td>
<td>149 (105, 219)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Exercise test data</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exercise time (min)</td>
<td>9.1 ± 3.7</td>
<td>11.6 ± 3.5</td>
<td>0.001</td>
</tr>
<tr>
<td>Limiting symptoms</td>
<td>26 (72)</td>
<td>20 (22)</td>
<td>&lt;0.0001**</td>
</tr>
<tr>
<td>Systolic BP increase (mmHg)</td>
<td>13 ± 20</td>
<td>22 ± 19</td>
<td>0.036</td>
</tr>
<tr>
<td>Abnormal blood pressure response</td>
<td>14 (39)</td>
<td>15 (17)</td>
<td>0.026**</td>
</tr>
<tr>
<td>ST depression ≥2 mm</td>
<td>15 (42)</td>
<td>18 (20)</td>
<td>0.04**</td>
</tr>
</tbody>
</table>

Values are mean ± SD, or median (lower quartile, upper quartile), or n (%) of patients. Statistical comparison by unpaired $t$-test or Mann–Whitney $U$ test or $\chi^2$ independence test.

![Figure 1](https://example.com/f1.png)

Figure 1 Kaplan–Meier plot of cumulative symptom-free survival over 12 months according to the incidence of limiting symptoms on baseline exercise testing. $P < 0.0001$ (log rank test).
Previous work

A large proportion of patients were limited by symptoms on treadmill exercise testing despite claiming to be asymptomatic. This confirms previous studies\(^3,4\) and has a number of possible explanations. Patients with aortic stenosis may overestimate their exercise capacity, having gradually reduced their activity to avoid symptoms. The SAS questionnaire suggested reduced activity in 28% of patients. It is also possible that, in some cases, limiting breathlessness was defined as a symptom when it was in fact normal. While possible that, in some cases, limiting breathlessness was expected that exercise testing should be normal at baseline followed the patients for 2.5 years rather than 12 months. It is expected that exercise testing should be normal at baseline in patients with only relatively mild aortic stenosis.

The addition of physiological measures during exercise testing did not improve the predictive accuracy provided by symptom-provocation alone. The systolic blood pressure response during exercise was reduced in patients who subsequently developed symptoms. This is in agreement with the study by Otto et al.\(^6\) in which patients who reached an endpoint of death or aortic valve replacement had an average increase in systolic blood pressure of 15 mmHg on exercise, compared with 29 mmHg in those who remained asymptomatic. However, there was substantial overlap in the two groups in both Otto’s and the present study, so the blood pressure response was not prognostically significant on multivariate regression analysis. Similarly, ST segment depression occurred in 42% patients who developed symptoms but also 20% who remained asymptomatic. Although total exercise time was reduced in patients who subsequently developed symptoms, exercise time was related most closely to age in all groups and did not contribute independently to the prediction of symptom onset. The low sensitivity of these physiological measures explains the relatively low positive predictive accuracy of 55% when a positive exercise test is defined by a composite of symptoms, blunted systolic blood pressure response, sustained ventricular arrhythmias, or ST segment changes.\(^{17}\)

Exercise vs. echocardiography

This study, like those of Amato et al.\(^3\) and Alborino et al.,\(^{17}\) showed that exercise testing was superior to echocardiography in predicting clinical events. In contrast, Otto et al.\(^4\) found that haemodynamic severity at rest was a superior predictor of clinical outcome to exercise duration. This apparent discrepancy is probably caused by differences in the grade of stenosis and in the length of follow-up between individual studies. Otto et al.\(^4\) studied patients with less severe stenosis (orifice area 1.3 cm\(^2\)) than Amato et al.\(^3\) (orifice area 0.6 cm\(^2\)) or the present study (orifice area 0.9 cm\(^2\)) and followed the patients for 2.5 years rather than 12 months. It is expected that exercise testing should be normal at baseline in patients with only relatively mild aortic stenosis.

Rosenhek et al.\(^5,8\) found echocardiography useful, but did not perform exercise tests. They showed that an annual rise in peak velocity >0.3 m/s predicted surgery, although only when combined with heavy calcification of the aortic valve. However, echocardiographic signs of severe aortic stenosis or of rapid progression may lead to the clinician being more willing to interpret minor limitation of exercise as an overt symptom or to refer an asymptomatic patient for surgery. In the present study, patients with more severe stenosis were also more likely to develop symptoms during exercise testing. However, while the positive predictive accuracy of limiting symptoms in patients with severe stenosis was greater than in more moderate stenosis, the negative predictive accuracy was lower. This reflects that such patients are at high risk and require frequent assessment even when asymptomatic.

### Table 2 Predictive ability for onset of spontaneous symptoms within 12 months

<table>
<thead>
<tr>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
<th>PPA (%)</th>
<th>NPA (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAS Class I, age ≤ 70 and limiting symptoms (n = 14)</td>
<td>65</td>
<td>93</td>
<td>79</td>
</tr>
<tr>
<td>Limiting symptoms (n = 46)</td>
<td>72</td>
<td>78</td>
<td>57</td>
</tr>
<tr>
<td>EOA ≤ 0.8 cm(^2) (n = 52)</td>
<td>67</td>
<td>68</td>
<td>46</td>
</tr>
<tr>
<td>EOA ≤ 0.8 cm(^2) and limiting symptoms (n = 37)</td>
<td>71</td>
<td>68</td>
<td>65</td>
</tr>
<tr>
<td>SAS Class II (n = 35)</td>
<td>43</td>
<td>79</td>
<td>41</td>
</tr>
<tr>
<td>Abnormal blood pressure response (n = 29)</td>
<td>39</td>
<td>82</td>
<td>48</td>
</tr>
<tr>
<td>ST depression ≥ 2 mm (n = 33)</td>
<td>42</td>
<td>79</td>
<td>45</td>
</tr>
</tbody>
</table>

PPA, positive predictive accuracy; NPA, negative predictive accuracy.

### Table 3 Univariate and multivariate logistic regression analyses to predict symptom-onset within 12 months

<table>
<thead>
<tr>
<th>Univariate analysis</th>
<th>Multivariate analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>OR</td>
<td>95% CI (OR)</td>
</tr>
<tr>
<td>------</td>
<td>-------------</td>
</tr>
<tr>
<td>Exercise time (min)</td>
<td>1.04</td>
</tr>
<tr>
<td>Limiting symptoms (yes/no)</td>
<td>8.39</td>
</tr>
<tr>
<td>Abnormal blood pressure response (yes/no)</td>
<td>3.33</td>
</tr>
<tr>
<td>Peak velocity (m/s)</td>
<td>1.04</td>
</tr>
<tr>
<td>Effective orifice area (cm(^2))</td>
<td>1.04</td>
</tr>
<tr>
<td>ST depression ≥ 2 mm (yes/no)</td>
<td>3.53</td>
</tr>
</tbody>
</table>
Safety of exercise testing

Symptomatic severe stenosis was traditionally regarded as an absolute contraindication to exercise testing and moderate stenosis without symptoms as a relative contraindication. However, in a survey of 50,000 patients with suspected coronary disease or aortic stenosis, the morbidity rate was only 0.0005% and the mortality was 0.00004%. Exercise testing should be avoided in patients with unequivocal exertional symptoms, while in those with nonspecific breathlessness or no symptoms, it should be symptom-limited and should be discontinued if the systolic blood pressure falls by more than 20 mmHg or ST depression >5 mm develops. Under these circumstances exercise testing is safe.

Limitations of the present study

Nonspecific breathlessness is common and may be normal. Although the present study excluded patients with known pulmonary disease, occult pulmonary disease or lack of fitness were potential causes of breathlessness. However, most patients stopped exercise with fatigue and the positive predictive accuracy of exercise-induced symptoms was relatively high at 79% in active patients <70 years of age. This suggests that these symptoms were genuine.

Amato et al. reported a 10% rate of sudden death among patients with a positive exercise test. There were no sudden deaths in the present study. Nearly all of our exercise tests were performed prior to the publication of the study of Amato et al., and the results of exercise testing was not available to the supervising clinician. For these reasons, patients in this study were referred for valve replacement according to established criteria, notably the presence of spontaneous symptoms.

Clinical relevance

Exercise testing has been recommended in patients with severe aortic stenosis, indicated by an EOA <1.0 cm². However, because of the progressive nature of aortic stenosis, some patients with moderate stenosis at baseline may develop symptoms within 12 months and the EOA may fall during this period. Our results, therefore, suggest that patients with asymptomatic aortic stenosis should be exercised if the EOA is ≤1.2 cm². If the exercise test is normal, the chance of developing symptoms within 12 months is low, especially when the aortic stenosis is moderate. If clear limiting symptoms are revealed, the risk of progressing to spontaneous symptoms within 12 months is 79% for active patients aged <70, so it is reasonable to recommend valve replacement. For older or less active patients, the positive predictive value of exercise testing is lower and spontaneous symptoms probably remain the best criterion for surgery, especially if the risk of surgery is relatively high.

Acknowledgements

We thank Paul Wicks and Julie Morris, of South Manchester University Hospitals NHS Trust, for statistical advice.

References