

# Improving the positive predictive value of exercise testing in women

Y K Wong, S Dawkins, R Grimes, F Smith, K D Dawkins, I A Simpson

Heart 2003;89:1416-1421

See end of article for authors' affiliations

Correspondence to:  
Dr Yuk-ki Wong, Wessex  
Cardiac Unit,  
Southampton University  
Hospital, Tremona Road,  
Southampton, Hampshire  
SO16 6YD, UK;  
yukkiwong@yahoo.com

Accepted 22 May 2003

**Objective:** To identify exercise test variables that can improve the positive predictive value of exercise testing in women.

**Design:** Cohort study.

**Setting:** Regional cardiothoracic centre.

**Subjects:** 1286 women and 1801 men referred by primary care physicians to a rapid access chest pain clinic, of whom 160 women and 406 men had ST depression of at least 1 mm during exercise testing. The results for 136 women and 124 men with positive exercise tests were analysed.

**Main outcome measures:** The proportion of women with a positive exercise test who could be identified as being at low risk for prognostic coronary heart disease and the resulting improvement in the positive predictive value.

**Results:** Independently of age, an exercise time of more than six minutes, a maximum heart rate of more than 150 beats/min, and an ST recovery time of less than one minute were the variables that best identified women at low risk. One to three of these variables identified between 11.8% and 41.2% of women as being at low risk, with a risk for prognostic disease of between 0–11.5%. The positive predictive value for the remaining women was improved from 47.8% up to 61.5%, and the number of normal angiograms was potentially reducible by between 21.1–54.9%. By the same criteria, men had higher risks for prognostic disease.

**Conclusions:** A strategy of discriminating true from false positive exercise tests is worthwhile in women but less successful in men.

The treadmill exercise test is the classic initial investigation for the diagnosis of coronary heart disease (CHD), and significant ST depression on the ECG is the commonly used indicator of a positive test. Compared with men, in women it is known that ST depression is less likely to be associated with CHD.<sup>1,2</sup> In this study, we have investigated the hypothesis that the use of other exercise test variables in addition to significant ST depression can improve the positive predictive value of exercise testing in women. We also investigated whether this approach would be useful in men for whom the positive predictive value is already high.<sup>1</sup>

Although age is not an exercise test variable, it may be related to variables that are, such as exercise test duration, and so we investigated whether exercise test variables can discriminate for CHD independently of age. Unlike age, the other conventional coronary risk factors have not been shown to directly affect exercise test results, but we also compared the prevalence of these factors in subjects with true versus those with false positive exercise tests.

## METHODS

### Subjects

We investigated a cohort of subjects referred consecutively by primary care physicians to our rapid access chest pain clinic for exercise testing between December 1997 and July 2001. Subjects were included if they were not already known to have CHD and if the Bruce protocol was used. Subjects were excluded if they were taking digoxin or a  $\beta$  blocker and if these had not been stopped within 24 hours of the exercise test. Subjects with left bundle branch block or atrial fibrillation on their resting ECG were also excluded, as were those suspected of having an acute coronary syndrome or recent myocardial infarction. New or additional horizontal or downsloping ST depression was defined as significant if it

was at least 1 mm. The decision to stop an exercise test was left to the discretion of the supervising doctor but, generally, symptom limited exercise tests were carried out. Although the achievement of 1 mm of ST depression was considered significant for the purposes of this study, in clinical practice, an attempt would be made to achieve at least 2 mm. Exercise test data were recorded on to an electronic database. Data from subsequent cardiac catheterisation and myocardial perfusion imaging (MPI) were collected from electronically stored discharge letters. The only information obtained retrospectively was the duration of ST depression during recovery. ECG traces were reviewed without knowledge of the coronary anatomy. The times of the last ECG showing ST depression greater than that on the resting ECG and the first ECG without ST depression as compared with baseline were recorded. During recovery, ECGs were recorded at intervals of one minute or less.

### Exercise test variables and cardiac catheterisation

Significant CHD was defined as the presence of at least one lesion causing at least 50% diameter stenosis. Prognostic CHD was defined as significant left main artery disease, three vessel disease, or disease in the proximal left anterior descending artery and one other vessel. The exercise test variables investigated were occurrence of usual symptoms during exercise (that is, the presenting symptoms thought to be angina), magnitude of ST depression, exercise test duration, systolic blood pressure difference between the beginning and end of exercise, time of last abnormal and

**Abbreviations:** CHD, coronary heart disease; MPI, myocardial perfusion imaging; ROC, receiver operating characteristic

time of first normal ECG in recovery, maximum heart rate and maximum heart rate as a percentage of the predicted target heart rate, and the Duke treadmill score.<sup>3</sup> The predicted target heart rate was calculated as 220 minus the subject's age.

### Discriminatory ability of exercise test variables and age for CHD

For women with significant ST depression, the ability of each variable and age to discriminate for CHD was investigated by receiver operating characteristic (ROC) curves. Binary logistic regression was used to investigate which variables were independently associated with CHD. Time of last abnormal ECG and time of first normal ECG were entered separately in regression analysis, as were maximum heart rate and maximum heart rate as a percentage of target heart rate. The Duke score is not independent, being calculated from three other variables (occurrence of symptoms, duration of exercise, and magnitude of ST depression), but in a separate analysis we entered it as a variable.

The exercise test variables have different dimensions and, to allow comparisons between variables, subjects were ranked into quintiles for each variable. The percentage of subjects with CHD in each quintile was used to identify groups at very low (< 10%) or very high (> 90%) risk. The values of the variables that identified these groups were also noted. (Although quintiles were used for this particular analysis, the actual values for each variable and not the ranking categories were used for ROC curve and logistic regression analysis.)

Variables independently associated with CHD were also used in combination with each other to identify groups of women at low risk for CHD. The number of normal angiograms and MPI studies that could be avoided if these groups were not further investigated was calculated. We also calculated revised positive predictive values by excluding these low risk groups.

The above analyses were repeated for men except that we did not review the ECGs of all men with significant ST depression. Instead, we selected a consecutive number from the beginning of the study period that was similar to the number of women with significant ST depression.

### Prevalence of coronary risk factors in subjects with true and false positive exercise tests

The risk factors considered were smoking status (current, never, or previous), hypertension, diabetes, and hypercholesterolaemia. Risk factor status was ascertained at the time of exercise testing except that the cholesterol status was not always known at this stage. Hypercholesterolaemia was defined as the presence of a fasting total cholesterol concentration of at least 5.2 mmol/l measured before or after exercise testing.

The two tailed independent samples *t* test was used to compare variables with a normal distribution (Kolmogorov-Smirnov test) between groups. Categorical variables were compared with the  $\chi^2$  test (SPSS version 10, SPSS Inc, Chicago, Illinois, USA).

## RESULTS

### Subjects

During the study period, 3087 subjects were investigated who fulfilled the inclusion criteria (table 1). Eighty three patients were excluded because of atrial fibrillation or left bundle branch block on their resting ECG, 92 were excluded because they had not stopped taking  $\beta$  blocker (some of whom were also taking diltiazem), and 88 were excluded because they were taking diltiazem. The most common primary reasons for stopping an exercise test were fatigue (47.7%), dyspnoea

(22.4%), occurrence of usual symptoms (13.2%), and limb or back pain (2.9%). Only 2.5% of tests were stopped because the target heart rate had been reached or exceeded, and this occurred in subjects with negative exercise tests. Significant ST depression was associated with age and occurrence of usual symptoms and inversely associated with exercise duration, difference in systolic BP, maximum heart rate, and Duke score (table 1).

### Discriminatory ability of exercise test variables for CHD in women

Of 160 women who had significant ST depression, 11 were not further investigated, 7 were lost to follow up, 6 declined investigation, 5 had MPI (all with normal results), and 131 had cardiac catheterisation. Sixty five (47.8%) women had significant CHD and 28 (20.6%) had prognostic disease. For occurrence of usual symptoms and the magnitude of ST depression, it was not possible to rank subjects into quintiles. Among women with significant ST depression, it was found that those with CHD were just as likely to have usual symptoms as were those without (69.2% *v* 67.6%). For magnitude of ST depression, most values were between 1–2.5 mm (table 2). Within this narrow range, discrimination for CHD was not good (table 3). All other exercise test variables had some association with CHD but only those for maximum heart rate, exercise test duration, time of last abnormal ECG, and time of first normal ECG, as well as that for age, were significant (table 3). The ROC curves show that even for these variables, discrimination for CHD was fair rather than good. In this study, we did not know the precise duration of ST depression during recovery. However, among the 53 women whose time of last abnormal ECG was 53 s or less—that is, women in the two quintiles at lowest risk—all but seven had a time of first normal ECG of 60 s or less.

For those variables significantly associated with CHD, ranking into quintiles did not identify any groups at very low or very high risk for significant CHD, although groups at very low risk for prognostic disease were identified (table 3). Also, when each variable was considered individually, there seemed to be a limit to the lowest or highest risk identifiable. For instance, women in the quintile at lowest risk had a maximum heart rate of greater than 164 beats/min and 26.9% had significant CHD. In comparison, 30% of those with a rate greater than 170 beats/min (*n* = 10) had significant CHD, as did 25% of those with a rate greater than 180 beats/min (*n* = 4). Similarly, 75.9% of those with a rate of less than 130 beats/min had significant CHD, as did 75% of those with a rate of less than 120 beats/min (*n* = 8). However, when variables were used in combination, it was possible to identify groups at very low risk (table 4). The difference in the prevalence of CHD was relatively small between the first and second quintiles (table 3) and we therefore considered those women in the lowest two quintiles for risk for each variable. In contrast to using one variable, using variables in combination resulted in a smaller number of subjects with lower risks for significant and prognostic CHD but with lower positive predictive values and a reduced number of normal angiograms and MPI studies that could be avoided (table 4). Women at very high risk could also be identified, and those that were in the highest two quintiles for risk with respect to maximum heart rate (< 143 beats/min), time of last abnormal ECG (> 240 s), and exercise test duration (< 250 s) had a positive predictive value for significant and prognostic CHD of 92.3% and 46.2%, respectively (*n* = 13). In total, 21.3% of women with a positive exercise test could be allocated to either a very low risk or very high risk group for significant CHD.

**Table 1** Exercise test variables in subjects with and without significant ST depression during Bruce protocol treadmill exercise testing

|                                     | Women         |                 | Men           |                 |
|-------------------------------------|---------------|-----------------|---------------|-----------------|
|                                     | Positive test | Negative test   | Positive test | Negative test   |
| Number of subjects                  | 160           | 1126            | 406           | 1395            |
| Age (years)                         | 62.8 (9.8)    | 58.1 (12.2)***  | 64.0 (9.4)    | 53.8 (12.9)***  |
| Subjects with usual symptoms        | 64.3%         | 16.8%***        | 71.1%         | 19.4%***        |
| Test duration (s)                   | 330 (145)     | 402 (167)***    | 343 (148)     | 524 (198)***    |
| Difference in systolic BP (mm Hg)   | 41 (27)       | 46 (28)*        | 42 (26)       | 52 (26)***      |
| Maximum heart rate (beats/min)      | 147 (19)      | 153 (23)**      | 140 (19)      | 156 (23)***     |
| Percentage of target heart rate     | 93.4 (11.5)%  | 94.7 (12.5)%    | 89.5 (12.1)%  | 93.7 (11.8)%    |
| Duke score                          | -7.3 (5.5)    | 5.3 (4.1)***    | -8.7 (5.9)    | 7.1 (4.9)***    |
| Number with Duke score $\geq 5$     | 2 (1.3%)      | 718 (63.8)***   | 6 (1.5%)      | 1041 (74.6)***  |
| Number with Duke score -10 to 4 (%) | 110 (68.8%)   | 408 (36.2)*** † | 237 (58.4)    | 354 (25.4)*** ‡ |
| Number with Duke score $\leq -11$   | 48 (30.0%)    | 0***            | 163 (40.1%)   | 0***            |

Data are mean (SD) unless otherwise indicated. \* $p < 0.05$ ; \*\* $p < 0.01$ ; \*\*\* $p < 0.001$ .

†154 women without ST depression and a Duke score between -10 and 4 (moderate risk) had cardiac catheterisation or myocardial perfusion imaging (63.0% normal, 28.6% significant disease, 7.8% prognostic disease, 3.7% lost to follow up or declined investigation).

‡221 men without ST depression and a Duke score between -10 and 4 had cardiac catheterisation or myocardial perfusion imaging (26.7% normal, 49.3% significant disease, 19.5% prognostic disease, and 15.4% lost to follow up or declined investigation).  
BP, blood pressure.

### Discriminatory ability of exercise test variables for CHD in men

The ECG traces of 124 men who had significant ST depression and cardiac catheterisation were reviewed. The majority had disease with 112 (90.3%) and 70 (56.5%) men having significant and prognostic disease, respectively. As for women, among men with significant ST depression, usual symptoms were just as likely to occur in those with CHD as in those without (73.9% v 63.6). Also, the majority of men had ST depression of less than 2.5 mm. Good discrimination for CHD was not seen, although the prevalence of significant and prognostic disease was high in the few subjects with greater ST depression (table 2). In contrast to women, age was not an independent predictor for CHD in men; however, for men the time of the last abnormal ECG in recovery was associated with CHD (table 5). When age was not included in the regression analysis, exercise test duration and maximum heart rate were also significantly associated. However, the prevalence of CHD in the lowest quintiles for coronary risk was relatively high. Even when men were categorised according to the quintile values for women, the prevalence of disease was always greater in men. For instance, in men with a maximum heart rate of greater than 150 beats/min, an exercise time of more than six minutes, and an ST recovery time of less than one minute, the positive predictive values for significant and prognostic CHD were 37.5% and 25%, respectively.

### Coronary risk factors in subjects with true and false positive exercise tests

In women, the prevalences of current, previous, and never smokers, diabetes, and hypertension were the same in subjects regardless of whether they had true or false positive exercise tests (16.9 v 18.3%, 24.6 v 19.7%, 58.5 v 62.0%, 6.2 v

2.8%, and 38.0 v 38.5%). The prevalence of hypercholesterolaemia was generally high, but higher in those with true positive exercise tests (87.7 v 66.2%,  $p = 0.007$ ).

Only 12 men had false positive results. No difference was seen in the prevalence of any of the clinical characteristics. The corresponding prevalences are 15.2 v 0%, 37.5 v 41.7%, 47.3 v 58.3%, 12.5 v 0%, 45.5 v 75%, and 65.2 v 66.7%.

## DISCUSSION

### Usefulness of exercise testing in women

It has been argued that treadmill exercise testing is of limited value in women because of false positive rates that have been reported to be as high as 67%.<sup>2</sup> We found that exercise testing was useful since 88% of women had a negative test (table 1) and the majority were reassured without further investigation. However, similar to other studies, our positive predictive value or true positive rate was low at 47.8%.

### ST recovery time

Despite the low discriminatory power of ST depression for CHD in women, the ability of other exercise test variables to differentiate between true and false positive results has been poorly investigated.<sup>4</sup> Anecdotally, the duration of ST depression in recovery is associated with the probability of CHD, but quantitative data are sparse. In a study of both sexes, a recovery time of three minutes provided the best separation of true from false positive results,<sup>5</sup> but another study of mainly men found that 80% of subjects with a recovery time of less than one minute had CHD.<sup>6</sup> In a third study of women, the recovery time was an average 6.6 minutes in those with CHD and 3.7 minutes in those without.<sup>4</sup> In our study, we confirmed that there is no cut off value that can distinguish between those with and those without CHD, but that a recovery time of less than one minute identifies a group of

**Table 2** Magnitude of ST depression and prevalence of coronary heart disease (CHD)

| ST depression (mm) | Women              |                 |                | Men                |                 |                |
|--------------------|--------------------|-----------------|----------------|--------------------|-----------------|----------------|
|                    | Number of subjects | Significant CHD | Prognostic CHD | Number of subjects | Significant CHD | Prognostic CHD |
| 1-1.5              | 80 (58.8%)         | 40.0%           | 21.2%          | 44 (35.5%)         | 86.4%           | 54.5%          |
| 1.6-2              | 40 (29.4%)         | 57.5%           | 29.0%          | 62 (50.0%)         | 90.3%           | 51.6%          |
| 2.1-2.5            | 10 (7.4%)          | 50.0%           | 42.9%          | 4 (3.2%)           | 100%            | 25%            |
| >2.5               | 5 (3.6%)           | 80.0%           | 66.6%          | 14 (11.3%)         | 100%            | 92.9%          |

**Table 3** Ability of exercise test variables and age to discriminate for CHD in women with significant ST depression on exercise testing

| Exercise test variable                        | Percentage of subjects with CHD |                   |                  |                   |                  | p Value | Area under ROC curve† |
|---|---------------------------------|-------------------|------------------|-------------------|------------------|---------|-----------------------|
|   | First quintile                  | Second quintile   | Third quintile   | Fourth quintile   | Fifth quintile   |         |                       |
| Significant CHD                               |                                 |                   |                  |                   |                  |         |                       |
| Usual symptoms during test                    |                                 |                   |                  |                   |                  | NS      |                       |
| Magnitude of ST depression                    |                                 |                   |                  |                   |                  | NS      | 0.623                 |
| Time of first occurrence of ST depression (s) | 33.3 (320–602)                  | 25.9 (213–319)    | 44.4 (145–212)   | 64.3 (86–144)     | 69.2 (<86)       | NS      | 0.682                 |
| Time of last abnormal ECG in recovery (s)     | 20.0 (<24.0)                    | 29.6 (24–53)      | 48.1 (54–240)    | 65.2 (241–411)    | 71.4 (412–900)   | 0.006   | 0.732                 |
| Time of first normal ECG in recovery (s)*     | 25.0 (<53)                      | 11.5 (53–60)      | 60.0 (61–157)    | 47.8 (158–390)    | 71.4 (391–765)   | 0.013   | 0.728                 |
| Duration of exercise test (s)                 | 25.9 (436–720)                  | 29.6 (359–435)    | 39.3 (248–358)   | 66.7 (172–247)    | 77.8 (<171)      | 0.028   | 0.719                 |
| Maximum heart rate (beats/min)                | 26.9 (164–197)                  | 33.3 (152–163)    | 46.2 (143–151)   | 56.0 (130–142)    | 75.9 (<130)      | 0.006   | 0.711                 |
| Target heart rate (%)                         | 37.0 (103.2–128)                | 44.4 (95.8–103.1) | 32.1 (88.8–95.7) | 55.6 (83.4–88.7)  | 70.4 (<83.4%)    | NS      | 0.620                 |
| Increase in systolic blood pressure (mm Hg)   | 23.1 (68–131)                   | 48.3 (47–67)      | 51.9 (35–46)     | 64.3 (17–34)      | 50.0 (<17)       | NS      | 0.621                 |
| Duke treadmill score                          | 37.0 (–2 to 7)                  | 37.9 (–6 to –3)   | 34.8 (–9 to –7)  | 64.5 (–12 to –10) | 61.5 (<–12)      | NS      | 0.631                 |
| Age (years)                                   | 29.6 (37.6–51.6)                | 25.9 (51.9–60.0)  | 35.7 (60.1–65.4) | 81.5 (66–72.2)    | 66.7 (72.3–80.6) | 0.028   | 0.715                 |
| Prognostic CHD                                |                                 |                   |                  |                   |                  |         |                       |
| Maximum heart rate (mmHg)                     | 0                               | 20.0              | 11.5             | 24.0              | 44.8             | 0.007   | 0.742                 |
| Time of last abnormal ECG in recovery (s)     | 8.0                             | 14.8              | 22.2             | 26.1              | 32.1             | NS      | 0.664                 |
| Time of first normal ECG in recovery (s)      | 10.0                            | 3.8               | 26.7             | 26.1              | 38.1             | NS      | 0.707                 |
| Duration of exercise test (s)                 | 7.4                             | 7.4               | 14.3             | 33.3              | 40.7             | NS      | 0.706                 |
| Age (years)                                   | 10.7                            | 3.6               | 17.9             | 39.3              | 28.6             | NS      | 0.687                 |
| Duke treadmill score                          | 14.8                            | 10.3              | 8.7              | 29.0              | 38.5             | NS      | 0.677                 |

Data in parentheses indicate the range for each quintile.

\*The time of the first normal ECG was available for only 105 subjects. For subjects for whom this variable was unknown, the time of the last abnormal ECG was at least 240 s. The mean (SD) interval between the time of the last abnormal and the first normal ECG was 53 (33) s.

†Areas under the receiver operating characteristic (ROC) curve of 0.6–0.69, 0.7–0.79, and 0.8–0.89 are considered to indicate poor, fair, and good discriminatory ability, respectively.

women at low risk for having prognostic CHD. For men, recovery time was a good discriminator for CHD but, because the prevalence of CHD was high, even short recovery times could not be used to exclude prognostic disease (table 5).

### Exercise test workload

An absence of both symptoms and adverse ECG changes, together with an adequate workload during exercise testing, indicates a good coronary prognosis. As heart rate is related to cardiac oxygen consumption<sup>7</sup> and exercise duration corresponds to the maximum oxygen consumed by the whole body,<sup>8</sup> then our results indicate that these measures of workload can also discriminate for CHD even in women with significant ST depression. The maximum heart rate attainable varies with age but there can be great individual variation. This may explain why the actual maximum heart rate as a percentage of the target heart rate was not a good discriminator for CHD. This finding supports those who argue that exercise tests should be terminated on the basis of symptoms or maximum effort rather than an arbitrary heart rate. Also, it emphasises the importance of stopping drugs

such as  $\beta$  blockers, which can attenuate the heart rate response to exercise.

### Exercise test scores

Exercise test scores have been devised to improve the predictive accuracy<sup>3,9</sup> and to give an estimate of cardiac prognosis,<sup>10</sup> and they have been promoted for general use. Only one recent score has been designed specifically for women<sup>11</sup> despite the existence of significant sex differences.<sup>1</sup> We found that one of the better known scores, the Duke treadmill score,<sup>3</sup> was not a good discriminator for CHD in either men or women with significant ST depression. One common criticism of the way in which these scores have been derived is that they are based on patients who have had both exercise tests and cardiac catheterisation but not those presumably lower risk patients who had exercise tests only. In our cohort of 1286 women, it can be seen that those assessed as being at high risk by the Duke score (score of  $\geq 11$  or less) all had ST depression while nearly all those assessed as being at low risk (score  $\geq 5$ ) did not (table 1). Of 518 women assessed as being at moderate risk, 110 had ST

**Table 4** Criteria for identifying groups at low risk for CHD in women with significant ST depression

| Criteria for low risk  | Number of women (%) | Number with CHD (%) |            | Potential reduction in normal coronary angiograms (%) | Positive predictive value for significant CHD in women not at low risk (%) |
|--|---------------------|---------------------|------------|---|--|
|  |                     | Significant         | Prognostic |   |  |
| Time of last abnormal ECG <1 min   | 52 (38.2)           | 13 (25.0)           | 6 (11.5)   | 54.9  | 61.5   |
| Exercise test duration >6 min ( $\geq$ stage III)  | 54 (39.7)           | 15 (27.7)           | 4 (7.4)    | 54.9  | 61.0   |
| Maximum heart rate >150 beats/min  | 56 (41.2)           | 17 (30.3)           | 6 (10.7)   | 54.9  | 60.0   |
| Time of last abnormal ECG <1 min and exercise test duration >6 min*                                      | 24 (17.6)           | 4 (16.7)            | 1 (4.1)    | 28.2  | 54.5   |
| Time of last abnormal ECG <1 min and exercise test duration >6 min and maximum heart rate >150 beats/min | 16 (11.8)           | 1 (6.3)             | 0          | 21.1  | 53.3   |

\*Similar results were obtained when maximum heart rate was used instead of time of last abnormal ECG.

**Table 5** Ability of exercise test variables and age to discriminate for CHD in men with significant ST depression on exercise testing

| Exercise test variable                        | Percentage of subjects with CHD |                   |                  |                   |                   | p Value | Area under ROC curve |
|---|---------------------------------|-------------------|------------------|-------------------|-------------------|---------|----------------------|
|   | First quintile                  | Second quintile   | Third quintile   | Fourth quintile   | Fifth quintile    |         |                      |
| Significant CHD                               |                                 |                   |                  |                   |                   |         |                      |
| Usual symptoms during test                    |                                 |                   |                  |                   |                   | NS      |                      |
| Magnitude of ST depression                    |                                 |                   |                  |                   |                   | NS      | 0.621                |
| Time of first occurrence of ST depression (s) | 68.0 (318–710)                  | 96.0 (213–317)    | 93.5 (161–212)   | 100.0 (91–160)    | 96.0 (<90)        | NS      | 0.795                |
| Time of last abnormal ECG in recovery (s)     | 70.8 (<39.0)                    | 87.5 (39–240)     | 96.0 (241–361)   | 100.0 (362–500)   | 100.0 (501–2603)  | 0.045   | 0.837                |
| Time of first normal ECG in recovery (s)*     | 70.6 (<57)                      | 78.9 (57–180)     | 93.8 (181–360)   | 100.0 (361–517)   | 100.0 (518–1040)  | NS      | 0.814                |
| Duration of exercise test (s)                 | 72.0 (470–769)                  | 92.0 (376–469)    | 92.0 (275–375)   | 96.0 (210–274)    | 100.0 (<210)      | NS      | 0.783                |
| Maximum heart rate (beats/min)                | 72.0 (154–204)                  | 92.0 (144–153)    | 95.8 (131–143)   | 96.2 (123–13)     | 95.8 (<122)       | NS      | 0.752                |
| Target heart rate (%)                         | 79.2 (100.6–127)                | 87.5 (93.0–100.5) | 96.0 (85.0–92.9) | 96.0 (77.8–84.9)  | 91.7 (<77.7)      | NS      | 0.677                |
| Increase in systolic blood pressure (mm Hg)   | 88.0 (71–127)                   | 92.0 (49–70)      | 88.0 (36–48)     | 91.7 (25–35)      | 92.0 (<25)        | NS      | 0.536                |
| Duke treadmill score                          | 72.7 (–4 to 3)                  | 88.5 (–7 to –5)   | 93.8 (–11 to –8) | 95.2 (–14 to –12) | 100.0 (<–14)      | NS      | 0.793                |
| Age (years)                                   | 83.3 (40.1–55.6)                | 76.0 (55.6–63.4)  | 95.8 (63.7–67.6) | 96.0 (67.6–70.7)  | 100.0 (71.3–79.8) | NS      | 0.733                |
| Prognostic coronary heart disease             |                                 |                   |                  |                   |                   |         |                      |
| Maximum heart rate (beats/min)                | 40.0                            | 52.0              | 70.8             | 50.0              | 70.8              | NS      | 0.587                |
| Time of last abnormal ECG in recovery (s)     | 33.3                            | 50.0              | 64.0             | 66.7              | 70.8              | 0.024   | 0.663                |
| Time of first normal ECG in recovery (s)      | 29.4                            | 36.8              | 56.3             | 64.7              | 66.7              | NS      | 0.670                |
| Duration of exercise test (s)                 | 32.0                            | 52.0              | 44.0             | 76.0              | 79.2              | NS      | 0.689                |
| Age (years)                                   | 41.7                            | 36.0              | 83.3             | 64.0              | 54.2              | NS      | 0.600                |
| Duke treadmill score                          | 45.5                            | 57.7              | 53.1             | 52.4              | 73.9              | NS      | 0.600                |

Data in parentheses indicate the range for each quintile.

\*The time of the first normal ECG was available for only 87 subjects. For subjects for whom this variable was unknown, the time of the last abnormal ECG was at least 253 s. The mean (SD) interval between the time of the last abnormal and the first normal ECG was 59 (40) s.

depression. As moderate risk has been equated with a four year survival of between 92–95%<sup>12</sup> and a probability of prognostic disease of 31%,<sup>3</sup> it may be argued that all 518 should have been further investigated. However, only 38% of the moderate group without ST depression were selected for further investigation and only 7.8% of these had prognostic disease while 63% had normal coronary arteries. Similarly, not all men at moderate risk were selected for further investigation and, although the proportion with prognostic disease was higher in men than in women, it was still less than that predicted by the Duke score. Therefore, although exercise test scores have been promoted as a way of reducing “overuse of invasive procedures”,<sup>13</sup> they may substantially increase them. The main difference between our approach and that taken by such scores is that we have focused on a specific group of patients, namely women with positive exercise tests. Our findings have been based on diagnostic data for the majority of these patients.

### Using exercise test variables to identify women at low and high risk for CHD

This study shows that none of the exercise test variables should be viewed as dichotomous from a diagnostic point of view. For instance, most research studies regard the achievement of at least 1 mm of ST depression as a positive test. In this study, though, fewer than half of all women with such ST depression had CHD, but the majority of women with greater than 2.5 mm of ST depression had significant disease and a significant number of women with greater than 2 mm of ST depression had prognostic disease (table 2). It seems more appropriate to consider variables to have values that can identify high, intermediate, and low risk groups. The clinical usefulness of such values is increased if they can identify groups at very low or very high risk and if they occur commonly. In this study, only 10% of women had at least 2 mm of ST depression. Conversely, ST recovery time, exercise test duration, and maximum heart rate could not individually identify groups at very low or very high risk for significant disease (table 3). It was found that the best

method for excluding significant disease was to use all three of these variables in combination. Although this group of women was relatively small, it would still have been possible to reduce the number of normal coronary angiograms by 21% (table 4). Using all three variables in combination was also the best method for identifying a group at very high risk.

### Coronary risk factors in women with true and false positive exercise tests

As with other studies, in this study we found that youth in women was associated with false positive exercise tests and, in fact, false positive results were predominant up to the relatively advanced age of 65 (table 3). Although maximum heart rate and exercise test duration may be expected to be associated with age, both of these variables discriminated for CHD independently of age. We found that conventional coronary risk factors, apart from hypercholesterolaemia, were as prevalent in women with false as in those with true positive exercise tests. One reason may be that primary care physicians are more likely to refer patients for investigation if they do have risk factors. Although hypercholesterolaemia was more common in women with true positive exercise tests, its use in improving the positive predictive value of exercise testing in women is likely to be limited because the prevalence of hypercholesterolaemia in women without CHD was also high.

### Conclusions

It is possible to improve the positive predictive value of exercise testing in women and to reduce the rate of normal coronary angiograms by consideration of the maximum heart rate, exercise test duration, and ST recovery time to identify a group of women with ST depression at low risk for prognostic CHD. In contrast, the positive predictive value of exercise testing is always higher in men and use of these variables cannot be used to exclude prognostic disease in men.

### Authors' affiliations

**Y K Wong, S Dawkins, R Grimes, F Smith, K D Dawkins, I A Simpson,** Wessex Cardiac Unit, Southampton University Hospital, Tremona Road, Southampton, Hampshire, UK

### REFERENCES

- 1 **Wong Y**, Rodwell A, Dawkins S, *et al*. Sex differences in investigation results and treatment in subjects referred for investigation of chest pain. *Heart* 2001;**85**:149–52.
- 2 **Sketch MH**, Mohiuddin SM, Lynch JD, *et al*. Significant sex differences in the correlation of electrocardiographic exercise testing and coronary arteriograms. *Am J Cardiol* 1975;**36**:169–73.
- 3 **Shaw LJ**, Peterson ED, Shaw LK, *et al*. Use of a prognostic treadmill score in identifying diagnostic coronary disease subgroups. *Circulation* 1998;**98**:1622–30.
- 4 **Pratt CM**, Francis MJ, Divine GW, *et al*. Exercise testing in women with chest pain. Are there additional exercise characteristics that predict true positive test results? *Chest* 1989;**95**:139–44.
- 5 **Berman JL**, Wynne J, Cohn PF. A multivariate approach for interpreting treadmill exercise tests in coronary artery disease. *Circulation* 1978;**58**:505–12.
- 6 **Goldschlager N**, Selzer A, Cohn K. Treadmill stress tests as indicators of presence and severity of coronary artery disease. *Ann Intern Med* 1976;**85**:277–86.
- 7 **Sheffield LT**, Holt JH, Reeves TJ. Exercise graded by heart rate in electrocardiographic testing for angina pectoris. *Circulation* 1965;**32**:622–9.
- 8 **Ramamurthy G**, Kerr JE, Harsha D, *et al*. The treadmill test: where to stop and what does it mean? *Chest* 1999;**115**:1169.
- 9 **Morise AP**, Detrano R, Bobbio M, *et al*. Development and validation of a logistic regression-derived algorithm for estimating the incremental probability of coronary artery disease before and after exercise testing. *J Am Coll Cardiol* 1992;**20**:1187–96.
- 10 **Mark DB**, Hlatky MA, Harrell FE, *et al*. Exercise treadmill score for predicting prognosis in coronary artery disease. *Ann Intern Med* 1987;**106**:793–800.
- 11 **Morise AP**, Lauer MS, Froelicher VF. Development and validation of a simple exercise test score for use in women with symptoms of suspected coronary artery disease. *Am Heart J* 2002;**144**:818–25.
- 12 **Mark DB**, Shaw L, Harrell FE, *et al*. Prognostic value of a treadmill exercise score in outpatients with suspected coronary artery disease. *N Engl J Med* 1991;**325**:849–53.
- 13 **Marcus R**, Lowe III R, Froelicher VF, *et al*. The exercise test as gatekeeper. *Chest* 1995;**107**:1442–6.

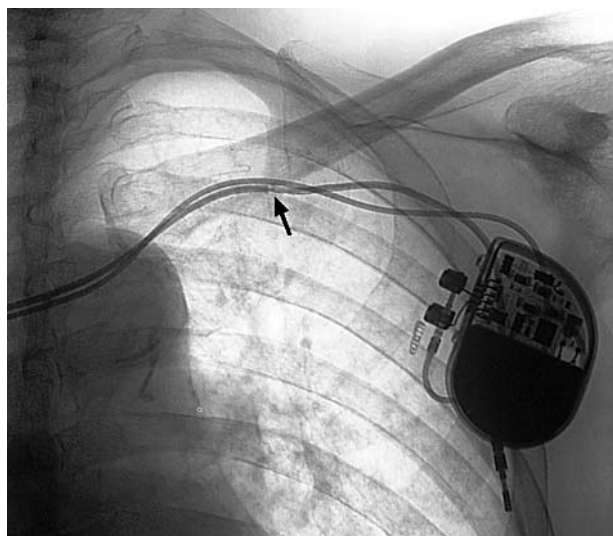
## IMAGES IN CARDIOLOGY

### Complications of pacemaker leads

A patient with dilated cardiomyopathy and heart failure was referred for implantation of a biventricular pacemaker. He had received a VVI pacemaker seven years ago for atrial fibrillation with bradycardia. A new ventricular lead was implanted after two years due to isolation failure of the first lead (below left). Both leads were implanted by subclavian puncture. Implantation of the coronary sinus lead was attempted using a left sided approach, but had to be aborted due to occlusion of the innominate vein, with venous drainage via a network of collaterals (below right). The

procedure was successfully undertaken using right sided axillary vein puncture.

This case illustrates how insulation failure or lead fracture may result from subclavian crush. Friction with the clavicle may be avoided by introducing leads more distally by cephalic vein cutdown or axillary vein puncture. The second complication was venous occlusion, which most often involves the subclavian or innominate veins, and is most common in patients with multiple leads.



**H Burri  
H Sunthorn  
D Shah**

burri-haran@diogenes.hcuge.ch