

## Exercise testing in clinical medicine

Euan A Ashley, Jonathan Myers, Victor Froelicher

**Exercise-induced changes in the electrocardiogram have been used to identify coronary artery disease for almost a century. Over the past decade, however, clinicians have increasingly focused on more expensive diagnostic tools believing them to offer improved diagnostic accuracy. In fact, by incorporating historical data, the simple exercise test can in most cases outperform the newer tests. The use of prediction equations and non-staged exercise protocols can improve the test still further, while advances in the use of the test for prognosis, with the discovery of novel risk factors and the addition of gas analysis, may in the future shift the primary emphasis away from diagnosis. Brief, inexpensive, and done in most cases without the presence of a cardiologist, the exercise test offers the highest value for predictive accuracy of any of the non-invasive tests for coronary artery disease.**

The exercise test has a long history in cardiovascular medicine. Although Willem Einthoven,<sup>1</sup> in 1908, was the first to document changes in the ST segment of the electrocardiogram (ECG) with exercise, it was not until 1932 that Goldhammer and Scherf<sup>2</sup> proposed exercise electrocardiography as a diagnostic tool for angina. Since then, the technique of exercising patients to stress the cardiovascular system has played a central part in the diagnostic work-up of coronary artery disease. However, a low reported sensitivity<sup>3</sup> has led some specialists, notably those in the USA, to turn away from the traditional exercise test in favour of more expensive investigations. We propose that this desertion is premature, because optimisation of the test can produce predictive values equal to those of the best and most expensive techniques. Indeed the test should be more widely used.

### Diagnosis

Meta-analysis of trials has shown that the exercise test has a specificity of around 80% and a sensitivity of around 70% for obstructive coronary disease confirmed by angiography.<sup>4</sup> However, many of these studies had methodological problems of limited challenge and work-up bias. The former takes place when patients already known to have coronary artery disease (eg, those with previous myocardial infarction) are used to challenge a diagnostic test. Work-up bias describes the situation in which the gold-standard test is applied only after a positive result from the test in question. In fact, in the only trial to date to clearly avoid work-up bias<sup>3</sup> specificity was 85% and sensitivity 45%. However, the application of computers yields multivariate prediction equations that can, by incorporating data obtained from patients outside the exercise-test setting, make the test as accurate as much more expensive tests (table 1 and panel 1). Although they may seem intimidating, programs are available for desktop computers and on the internet to do the calculations.

One difficulty with such equations is that they are hard to apply to populations other than those in which they were developed. One way of overcoming this drawback is to use a consensus approach, in which a final classification of low or high risk is made on the basis of a consensus among equations validated in different populations (eg, each equation classifies a given patient into low, intermediate, or high risk, and a given classification is accepted only if two of three equations agree).<sup>5</sup> Such a method predicts angiographic disease better than do cardiologists.<sup>6</sup> A variation on this theme is the use of prognostic scores, the most universal of which is the Duke treadmill score because it can be used both for prognosis<sup>7</sup> and diagnosis.<sup>8</sup> The score consists of three variables: the amount of ST depression, exercise capacity, and the presence of angina. This score can be calculated by a nomogram or by a computer (some of the commercial treadmill systems automatically calculate it).

A further advantage of the consensus or score approach is that, as opposed to yes or no answers, the separation of patients into groups according to risk suggests a management course: low-risk patients need no further testing at that time; those at high-risk need an invasive study; and those at intermediate-risk need another non-invasive study. On the assumption that the intermediate group is eventually diagnosed, data suggest that this test strategy would produce sensitivity and specificity values of around 90%.<sup>5</sup>

Another approach to improving test characteristics is to investigate different ECG criteria. Atwood and colleagues<sup>9</sup> took 100 computed measurements from digitised exercise ECG recordings and related them to angiographic data on the same patients. They found that computerised measurements at 3.5 min of recovery, from lead V5, 60 ms after the QRS complex (ST60), were better than all other single measurements. Further, prediction equations that included clinical and exercise-test data showed the greatest diagnostic power. In a new approach, Michaelides and colleagues examined 245 patients who underwent exercise testing with standard 12 leads, right ventricular leads, and thallium-201 scintigraphy. They found sensitivities of 66%, 92%, and 93%, and specificities of 88%, 88%, and 82%, respectively, for the detection of "any" coronary artery disease by angiography.<sup>10</sup> It should be noted, however, that their study was of a population with an unusually high prevalence of coronary disease. The QRS complex has also been a focus of investigation as a marker of

*Lancet* 2000; **356**: 1592–97

Department of Cardiovascular Medicine, University of Oxford, John Radcliffe Hospital, Oxford OX3 9DU, UK (E A Ashley MRCP); Cardiology Division, Stanford University and Palo Alto Veterans Administration Health Care System, Palo Alto, California, USA (J Myers PhD, Prof V Froelicher MD)

Correspondence to: Dr Euan Ashley  
(e-mail: euan.ashley@cardiov.ox.ac.uk)

Grouping	Number of studies	Total patients	Sensitivity	Specificity	Predictive accuracy
<b>Meta-analysis of standard exercise ECG</b>	147	24 047	68%	77%	73%
Excluding MI patients	41	11 691	67%	72%	69%
Limiting work up bias	2	>1000	50%	90%	69%
<b>Meta-analysis of exercise test scores</b>	24	11 788	· ·	· ·	80%
<b>Probability stratification by exercise test scores</b>	1	2000	95%	90%	92%
<b>Thallium scintigraphy</b>	59	6038	85%	85%	85%
<b>SPECT without MI</b>	27	2136	86%	62%	74%
<b>Exercise echocardiography</b>	58	5000	84%	75%	80%
<b>Exercise echocardiography excluding MI patients</b>	24	2109	87%	84%	85%
<b>Non-exercise stress tests</b>					
Persantine thallium	11	<1000	85%	91%	87%
Dobutamine echocardiography	5	<1000	88%	84%	86%
<b>Cardiokymography</b>	1	617	71%	88%	79%
<b>Electron beam computed tomography</b>	5	2373	90%	45%	61%

MI=myocardial infarction. Values shown for non-exercise stress tests are probably overestimates as a result of work-up bias.

Table 1: Comparison of tests used in the diagnosis of coronary artery disease

myocardial ischaemia consequent on its use in scoring systems to predict infarct size. The Athens score is calculated from exercise-induced changes in the QRS complex, and a Dutch group have proposed a score based on composite changes in the Q, R, and S waves in two leads (aVF, V5). These investigators reported a sensitivity of 88.2% and a specificity of 84.8% for coronary artery disease, compared with values for ST depression of 54.9% and 83%, respectively. It appears that within this basic test there could yet be a measurement with even greater diagnostic power.

A question frequently raised is whether the test can predict disease in different sections of the population, such as in women or the elderly. In fact, the diagnostic characteristics are unchanged in women when age and symptoms move them into the intermediate probability of disease level. However, the lower prevalence of severe coronary artery disease and the fact that more women than men fail to reach maximum aerobic capacity during exercise testing leads to lower sensitivity for ST changes overall.<sup>11</sup> In relation to the elderly, Bayesian theory dictates that the high prevalence and greater severity of coronary stenoses in this group would increase the positive predictive value of testing but make it harder to rule out clinically significant disease. There is no reason to advise against exercise testing in the elderly if physical capability is intact. Goraya and colleagues<sup>12</sup> confirm the prognostic significance of the test in this population, particularly for the measurement of work load.

An important but under-appreciated caveat to the usefulness of exercise testing is the fundamental incapacity of ST depression to localise ischaemia. Although suggested by the recurrent finding of V5 as the lead of maximum predictive power irrespective of the coronary disease pattern, several direct investigations have corroborated the result.<sup>13</sup> The electrophysiological explanation for this apparent anomaly only came to light in 1998. Li and colleagues<sup>14</sup> showed, with a computer simulation and a sheep model of subendocardial ischaemia, that ST depression is generated at the lateral boundary between ischaemic and normal territories (elevation emanates from the ischaemic centre). Since epicardial boundaries of ischaemic territories are shared, and only those reaching the free wall contribute to external ST depression, this process cannot be said to localise ischaemia.

Some of the newer add-ons or substitutes for the exercise test can localise ischaemia as well as diagnose coronary disease when poor mobility precludes an exercise test or when the baseline ECG prohibits ST

analysis (such as in the presence of more than 1 mm of resting ST depression, left bundle-branch block, or Wolf-Parkinson-White syndrome). Table 1 presents some of the results from meta-analysis and multicentre diagnostic studies. Electron-beam computed tomography (EBCT) is a fast radiographic technique that can make a quantitative measurement of coronary artery calcification.<sup>15</sup> Cardiokymography is a simple motion sensor that allows recording of the movement of the left ventricle with a small transducer placed on the chest wall. Signal averaging has enhanced this older technology. Nuclear perfusion imaging includes both the early thallium radiographic images and the modern single-photon-emitted computed tomography (SPECT), which requires computer enhancement of the emissions of thallium and other agents. Since sensitivity and specificity are inversely related and altered by the chosen cut point for normal/abnormal, the predictive accuracy (percentage of patients correctly classified) is a convenient way to compare tests. For instance, although the sensitivity and specificity for exercise testing and EBCT are nearly opposite, their predictive accuracy is similar. Altering their cut points by lowering the amount of ST segment depression or raising the calcium score would result in similar sensitivities and specificities (although this assumes a disease prevalence of 50%). The clear message, however, is that although the newer technologies seem to have better diagnostic characteristics than the basic

**Panel 1: Examples of multivariate equations showing weightings for various risk factors**

**Froelicher pretest**

$-2.1+(0.03 \times \text{age})-(0.4 \times \text{symptoms})+(0.8 \times \text{diabetes})+(0.4 \times \text{hypercholesterolaemia})+(0.01 \times \text{pack-years})+(0.7 \times \text{resting ST depression in mm})$

**Froelicher post-test**

$-1.2+(3.3 \times \text{pretest})+(0.5 \times \text{exercise ST depression in mm})+(0.6 \times \text{ST slope})-(0.16 \times \text{MET})-(0.5 \times \text{exercise angina})$

**Morise pretest**

$-3.6+(0.08 \times \text{age})-(1.3 \times \text{sex})+(0.6 \times \text{symptoms})+(0.7 \times \text{diabetes})+(0.3 \times \text{smoking})-(1.5 \times \text{body surface area})+(0.5 \times \text{oestrogen})+(0.3 \times \text{number of risk factors})-(0.4 \times \text{resting ECG})$

**Morise post-test**

$-0.12+(4.5 \times \text{pretest})+(0.37 \times \text{mm ST depression})+(1.0 \times \text{ST slope})-(0.4 \times \text{negative ST})-(0.016 \times \text{maximum heart rate})$

**Detrano**

$1.9+(0.025 \times \text{age})-(0.6 \times \text{sex})-(0.1 \times \text{symptoms})-(0.05 \times \text{MET})-(0.02 \times \text{maximum heart rate})+(0.36 \times \text{exercise-induced angina})+(0.6 \times \text{mm ST depression})$

Study	Number of patients	Type of patient	Follow-up	Variable predominantly associated with prognosis
Forslund, 2000	731	Stable AP	40 months	Low exercise capacity (during & after exercise); marked ST depression
Francis, 2000	303	Stable chronic heart failure	47 months	Peak VO <sub>2</sub> ; VE/VCO <sub>2</sub> slope
Hulsmann, 1998	226	Heart failure	12 months	Big endothelin; NYHA class; maximum workload; plasma ANP
Villella, 1999	6296	Post MI	6 months	Low pressure-rate product; low work capacity; symptomatic exercise-induced ischaemia
Blair, 1995	9777 (men)	General population	61 months	Maximum treadmill time
Snader, 1997	3400	Consecutive patients	24 months	Estimated functional capacity
Kjeldsen, 1997	1999 (men)	Healthy	16 years	Large SBP rise during exercise
Lauer, 1999	2953	Consecutive patients	38 months	Chronotropic response
Cole, 1999	2428	Consecutive patients	72 months	Heart-rate recovery
Go Raya, 2000	514 elderly 2593 younger	Consecutive patients	72 months	Workload

See "Further reading" on *The Lancet* website for study references in order. AP=angina pectoris; MI=myocardial infarction; SBP=systolic blood pressure.

Table 2: **Summary of prognostic studies**

exercise test, this advantage does not bear comparison with methods incorporating other clinical information such as the exercise-test scores.

### Prognosis and novel indications

The overriding paradigm in the diagnosis of cardiovascular disease over the past decades has been the presumptive identification of obstructive coronary disease in order to direct interventional angiography or bypass surgery. However, shortcomings in non-invasive tests such as the work-up bias mentioned above, combined with limitations fundamental to angiography<sup>16</sup> have led some to suggest that the principal place for the exercise test is in the assessment of prognosis (table 2). The appropriate direction of life-saving pharmacological tools in the form of statins, angiotensin-converting-enzyme inhibitors,  $\beta$ -blockers, and aspirin can be enhanced by tests capable of identifying high-risk populations. The financial implications in a novel risk factor age are clear.

#### Exercise and heart failure

The rise of exercise training as a treatment for heart failure<sup>17</sup> has led to renewed interest in the assessment of patients with reduced ventricular function. Most of the work in this area has focused on pretransplant patients, and, in particular, the ability of ventilatory and gas-exchange measures to stratify risk and predict outcome. During the 1990s, over 40 studies were published showing that peak oxygen consumption (VO<sub>2</sub>) was a significant univariate or multivariate predictor of risk in heart failure, and this variable is now regarded as one of the most potent prognostic markers in this condition. A landmark 1991 study<sup>18</sup> engendered the notion of a cutpoint value for peak VO<sub>2</sub> (14 mL/kg per min), with patients who attained values above this point displaying 1 year and 2 year survival rates similar to those in transplanted patients, whereas those below this value had significantly worse survival rates.

Although appealing to clinicians, the idea of a cutpoint has been questioned. One study showed that ventilatory variables (VE/VCO<sub>2</sub>) and a chronotropic index (panel 2)

### Panel 2: Key concepts

#### MET

Metabolic unit. 1 MET equates to resting metabolic rate (taken to be 3.5 mL/kg per min)

#### Work-up bias

A biostatistical error caused when study participants are chosen for the gold-standard test on the basis of the test in question. To avoid this, participants need to agree to undergo both tests irrespective of the outcome of the first.

#### Chronotropic index

During an exercise test, heart-rate (HR) reserve used divided by metabolic reserve used

HR reserve used

$$\frac{HR_{\text{stage}} - HR_{\text{rest}}}{HR_{\text{peak}} - HR_{\text{rest}}} \times 100$$

Metabolic reserve used

$$\frac{MET_{\text{stage}} - MET_{\text{rest}}}{MET_{\text{peak}} - MET_{\text{rest}}} \times 100$$

fared better than peak VO<sub>2</sub> in predicting death,<sup>19</sup> whereas another study<sup>20</sup> recorded similar discriminatory power for each of seven cutpoints between 10 mL/kg per min and 17 mL/kg per min. In fact, an earlier publication from the same group<sup>21</sup> showed that peak VO<sub>2</sub>, expressed as a continuous variable, outperformed clinical measures, right heart catheterisation data, exercise time, and other exercise test data in predicting outcome. Such observations are very important in the current health-care climate, because they help direct scarce resources (eg, donor hearts) towards patients who are most likely to benefit.

As a result of these studies, the American College of Cardiology and American Heart Association have recommended the inclusion of gas analysis in exercise tests done to assess transplant patients and to differentiate cardiac exercise intolerance or dyspnoea from pulmonary causes.<sup>22</sup> Much evidence suggests an important role for abnormally heightened ventilation, both as a marker of the severity of heart failure and of prognosis.<sup>23</sup> This heightened ventilation has been expressed as an increase in the slope of the relation between minute ventilation and CO<sub>2</sub> production (VE/VCO<sub>2</sub> slope), a variable that has become common as an expression of the response to exercise in heart failure. Because this ratio is inversely related to cardiac output and parallels the degree of dead-space ventilation, it signifies the extent to which reduced pulmonary perfusion underlies the excess ventilatory response to exercise in part through a mismatching of ventilation and perfusion in the lung.

Data have suggested, however, that these techniques are underused. One survey found that less than 3% of exercise tests done in the Veterans Health Care system in the USA included gas-exchange measurements, despite more than 15% of the patients meeting the class-I criterion (panel 3). The promising data from randomised controlled trials of exercise training in heart failure suggest a class IIb indication for the use of exercise testing with gas analysis in the cardiac patient. Furthermore, the rising prevalence of heart failure demands an increasing role for this method of assessment for diagnosis, prognosis, or assessment of therapy indications.

#### Heart rate

Exercise chronotropic incompetence has been known for some time to predict all-cause mortality in healthy populations.<sup>24</sup> When combined with a measure of exercise capacity as the chronotropic index (panel 2) it is predictive of all-cause mortality independent of thallium

**Panel 3: Indications for the use of gas-exchange measurements in exercise testing**

Criteria	Methodology
<b>Class I</b> Disorders for which there is evidence and/or general agreement that given procedure or treatment is useful and effective	Assessment of exercise capacity and response to therapy in patients with heart failure who are being considered for heart transplantation  Assistance in the differentiation of cardiac versus pulmonary exercise-induced dyspnoea or impaired exercise capacity when cause is uncertain
<b>Class II</b> Disorders for which there is conflicting evidence and/or divergence of opinion about the usefulness/efficacy of procedure or treatment	
<b>IIa</b> Weight of evidence/opinion in favour of usefulness/efficacy	Assessment of exercise capacity when indicated for medical reasons in patients in whom subjective assessment of maximum exercise is unreliable
<b>IIb</b> Usefulness/efficacy less well established by evidence/opinion	Assessment of patient's response to specific therapeutic interventions in which improvement of exercise tolerance is important goal or end point  Determination of the intensity of exercise training as part of comprehensive cardiac rehabilitation
<b>Class III</b> Disorders for which there is evidence and/or general agreement that procedure/treatment is not useful/effective and in some cases may be harmful	Routine use to assess exercise capacity

ischaemia.<sup>25</sup> In fact, the risk associated with relative bradycardia was equal to that of a perfusion defect, and the effects were additive. The mechanism for these associations is not clear. Control of exercise heart rate involves both sympathetic and parasympathetic input. Some form of autonomic paresis, which does not necessarily require localised perfusion defects, could accompany the chronic sympathetic overactivation of the heart-failure syndrome. We know for certain that  $\beta$  receptors are downregulated in heart failure, that autonomic dysfunction implies cardiovascular risk,<sup>26</sup> and that chronotropic incompetence is predicted by increased left ventricular mass and cavity size.<sup>27</sup> Irrespective of the mechanism, the finding calls for attention to maximum heart rates. Ironically, the tests of many patients who fail to reach their age-predicted maximum heart rate are labelled non-diagnostic. This factor, together with the finding that the inclusion of all patients irrespective of maximum heart rate actually improves test characteristics for the diagnosis of coronary artery disease,<sup>28</sup> argues strongly for a role for this simple measurement in routine exercise testing.

The heart-rate fall after exercise has also been suggested as an important prognostic marker (heart-rate recovery).<sup>29</sup>

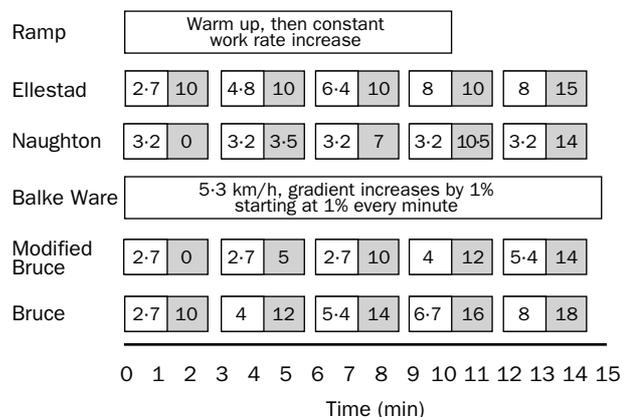
One study found that a delayed decrease in heart rate during the first minute after graded exercise was a powerful and independent predictor of all-cause mortality in 2428 patients followed for 6 years.<sup>30</sup> A potential difficulty with the measurement of heart-rate recovery for prognosis, however, is the use of a cool-down period, which is thought to negatively impact the diagnostic sensitivity of the test.

**Hypertension**

The clinical significance of an exaggerated blood-pressure rise to exercise has been recognised and debated for over 15 years. But there is now a clear indication that such a rise represents a risk factor for new-onset hypertension. The Framingham Offspring study<sup>31</sup> showed that an exaggerated diastolic blood-pressure rise was associated with a two-fold to four-fold risk of new-onset hypertension. A raised recovery systolic blood pressure was also predictive of hypertension in men alone. However, the data in relation to angiographic disease and mortality are less clear cut. Some investigators reported a relation between exercise systolic blood pressure and mortality significantly greater than that with resting blood pressure.<sup>32</sup> McHam and colleagues<sup>33</sup> showed that although a delayed decline in systolic blood pressure during recovery is associated with severe angiographic coronary disease, exercise hypertension is not associated with an increased mortality rate. This finding may be partly explained by the different exercise methods used in these studies. Some workers prefer bicycle ergometry to treadmill testing for exercise blood-pressure measurements because the arm is kept more still. McHam and colleagues<sup>34</sup> also showed that, as in heart-failure patients, exercise hypertension is negatively associated with myocardial perfusion abnormalities. What is clear is that for the moment the importance of exercise hypertension remains contentious.

**Exercise capacity**

The measurement of exercise capacity, made more difficult by staged exercise protocols, is often lost in the focus on the ST segment. In fact, exercise capacity turns out to be a strong independent risk factor for all-cause and cardiovascular mortality<sup>35-37</sup> and one, moreover, that can be altered by training. Blair and colleagues<sup>36</sup> showed that for every minute of increase in maximum treadmill time there was a corresponding 7.9% decrease in the risk of mortality.<sup>36</sup> Another study that examined several factors including SPECT thallium-perfusion scanning, found



**Figure 1: Summary of commonly used treadmill protocols**

The figures in the shaded boxes are the treadmill gradient percentages and the figures to the left are the treadmill speeds in km/h. All protocols except the Ramp and Balke-Ware use 3 min incremental stages.

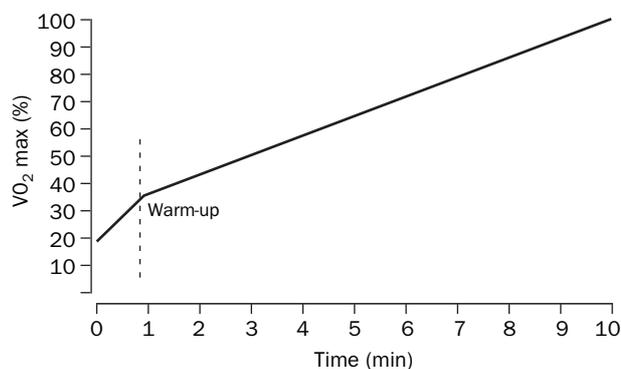


Figure 2: **Schematic representation of the ramp treadmill test**  
After a 1 min warm-up at 2.0 km/h and zero gradient, the rate of change of speed and grade is individualised to yield a work rate corresponding to the estimated maximum exercise capacity (y axis) in about 10 min.

that the strongest predictor of all-cause mortality was estimated fair or poor functional capacity (estimated simply from age and sex).<sup>37</sup> These analyses show exercise capacity to be a powerful prognostic indicator, and one that, by contrast with the measures discussed above in relation to heart rate, implies a mechanism for its own amelioration. Goldstein and Holmboe<sup>38</sup> have emphasised the importance of the potential to modify poor prognostic indicators, if we are to argue for measuring them.

### Protocol

Experience and history show that the pervasiveness of an idea, method, or product relies only in part on its intrinsic quality. Rather more, the acceptance of an idea relies on its extrinsic survival potential, a complex attribute with temporal, logistical, and experiential dimensions.<sup>39</sup> In applied exercise testing, these features are apparent in relation to the choice of treadmill protocol. When treadmill and cycle ergometer testing was first introduced into clinical practice, practitioners adopted protocols used by major researchers such as Balke, Astrand, Bruce, and Ellestad (figure 1). However, over time, those described by Bruce came to predominate. A survey of 71 cardiology divisions within the US Veterans Health Care system showed that the Bruce protocol was used by 82%,<sup>40</sup> a number that is broadly in line with the clinical experience of the authors and several other surveys. This use is in contrast with the recommendations of major academic bodies<sup>22,41</sup> that advocate gradual and individualised protocols. Their reasons are clear: work-rate increments that are too large or rapid result in lower sensitivity for detecting coronary disease,<sup>42</sup> a less reliable test for studying the effects of therapy,<sup>43</sup> and a tendency to overestimate exercise capacity.<sup>43,44</sup> The advantage of the experience and data acquired with the Bruce protocol is outweighed by the many patients unable to complete even stage 1 (about 5 METS, see panel 2), the difficulty in interpreting gas-exchange measurements below maximum with large and unequal increments in work, and the poorer estimates of exercise capacity achieved compared with other protocols. In addition, since many clinical laboratories do not employ gas analysis routinely, a clear relation between those variables that are measured and oxygen uptake is desirable.

Ramp testing is an alternative to the Bruce treadmill protocol, which was first proposed in 1991<sup>44</sup> and is now taking on a higher profile (figure 2). This protocol, aided by a pretest activity questionnaire, aims to bring the patient from rest to maximum exercise through a linear

increase in work over about 10 min. The rationale for this length of test has been well validated. Studies suggest that tests individualised to last 10 min produce the highest values for oxygen uptake,<sup>45</sup> better differentiation of treatment effects in clinical trials,<sup>46</sup> and a closer relation of work rate to oxygen uptake.<sup>43</sup> In addition, avoidance of fatigue and a focus on total work done rather than on exercise time (a highly variable measure), is provided by this approach.

Despite these potential advantages there have been few direct comparisons of the ramp protocol with others. Of these, three groups of researchers investigated cardiovascular patients whereas two studied healthy women and obese women. Two studies used a ramped Bruce protocol (ramping between the stages of the classic Bruce) whereas the others used individualised protocols. Taken overall, the findings from these studies suggest that a ramped protocol is preferred and better tolerated by patients,<sup>44,47</sup> produces an oxygen uptake-to-work ratio closer to unity (in some patients<sup>44</sup> but not those over 60 years), and may result in higher values for metabolic equivalents and exercise duration. Although this last finding was reported in only one study, such an observation has relevance for the predictive characteristics of the test. The higher workloads achieved by the same patients on the treadmill compared with the cycle ergometer are associated with improved exercise-test sensitivity for coronary artery disease.<sup>48</sup> It makes intuitive sense, given two tests in the same patient, that the one capable of eliciting a higher oxygen uptake would represent a truer examination of cardiopulmonary function (rather than the limitations of local muscle fatigue).

The generalist is increasingly called upon to begin the work-up of the cardiac patient and, therefore, potential enhancement of non-invasive tests should be welcomed. As such, developments summarised in this paper give the test a new lease of life. The inertia that sees the majority of centres still using unequal stage and rapidly incremented exercise protocols should be overcome. Well-validated prediction equations and novel diagnostic markers which lend increased value for predictive-accuracy should be adopted. Finally, major advances in the use of the test for prognosis, with the discovery of novel risk factors and the addition of gas analysis, may yet shift the primary emphasis away from diagnosis. The exercise test is inexpensive, brief, requires minimum equipment and space, and can be done in most cases safely without the presence of a cardiologist. Far from fading away, the technique first proposed in 1932<sup>2</sup> as an aid to the diagnosis of angina is in the process of being reborn.

### References

- 1 Einthoven W. Weiteres über das elektrokardiogramm. *Arch f d ges Physiol* 1908; **122**: 517.
- 2 Goldhammer S, Scherf D. Elektrokardiographische untersuchungen bei kranken mit angina pectoris ("ambulatorischer Typus"). *Ztschr f klin Med* 1932; **122**: 134.
- 3 Forelicher VF, Lehmann KG, Thomas R, et al. The electrocardiographic exercise test in a population with reduced workup bias: diagnostic performance, computerized interpretation, and multivariable prediction. *Ann Intern Med* 1998; **128**: 965-74.
- 4 Gianrossi R, Detrano R, Mulvihill D, et al. Exercise induced ST depression in the diagnosis of coronary artery disease: a meta-analysis. *Circulation* 1989; **80**: 87-98.
- 5 Do D, West JA, Morise A, Atwood E, Froelicher V. A consensus approach to diagnosing coronary artery disease based on clinical and exercise test data. *Chest* 1997; **111**: 1742-49.
- 6 Lipinski M, Atwood E, Osterberg L, Franklin B, West J, Froelicher V. Can physicians diagnose coronary disease as well as scores? *Circulation* 1999; **100**: I505.

- 7 Mark DB, Shaw L, Harrell FE Jr, et al. Prognostic value of a treadmill exercise score in outpatients with suspected coronary artery disease. *N Engl J Med* 1991; **325**: 849–53.
- 8 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.
- 9 Atwood J, Do D, Froelicher V. Can computerization of the exercise test replace the cardiologist? *Am Heart J* 1998; **136**: 543–52.
- 10 Michaelides A, Psoadakis Z, Dilaveris P, et al. Improving detection of coronary artery disease by exercise electrocardiography with the use of right precordial leads. *N Engl J Med* 1999; **340**: 340–45.
- 11 Hlatky MA, Pryor DB, Harrell FE Jr, Califf RM, Mark DB, Rosati RA. Factors affecting sensitivity and specificity of exercise electrocardiography: multivariable analysis. *Am J Med* 1984; **77**: 64–71.
- 12 Goraya TY, Jacobsen SJ, Pellikka PA, et al. Prognostic value of treadmill exercise testing in elderly persons. *Ann Intern Med* 2000; **132**: 862–70.
- 13 Mark D, Hlatky M, Lee K, et al. Localizing coronary artery obstructions with the exercise treadmill test. *Ann Intern Med* 1987; **106**: 53–55.
- 14 Li D, Li CY, Yong AC, Kilpatrick D. Source of electrocardiographic ST changes in subendocardial ischemia. *Circ Res* 1998; **82**: 957–70.
- 15 O'Rourke RA, Brundage BH, Froelicher VF, et al. American College of Cardiology/American Heart Association Expert Consensus document on electron-beam computed tomography for the diagnosis and prognosis of coronary artery disease. *Circulation* 2000; **102**: 126–40.
- 16 Topol EJ, Nissen SE. Our preoccupation with coronary luminology. The dissociation between clinical and angiographic findings in ischemic heart disease. *Circulation* 1995; **92**: 2333–42.
- 17 Coats AJ. Exercise training for heart failure: coming of age. *Circulation* 1999; **99**: 1138–40.
- 18 Mancini DM, Eisen H, Kusmaul W, Mull R, Edmunds LH Jr, Wilson JR. Value of peak exercise oxygen consumption for optimal timing of cardiac transplantation in ambulatory patients with heart failure. *Circulation* 1991; **83**: 778–86.
- 19 Robbins M, Francis G, Pashkow FJ, et al. Ventilatory and heart rate responses to exercise: better predictors of heart failure mortality than peak oxygen consumption. *Circulation* 1999; **100**: 2411–17.
- 20 Myers J, Gullestad L, Vagelos R, et al. Cardiopulmonary exercise testing and prognosis in severe heart failure: 14 mL/kg/min revisited. *Am Heart J* 2000; **139**: 78–84.
- 21 Myers J, Gullestad L, Vagelos R, et al. Clinical, hemodynamic, and cardiopulmonary exercise test determinants of survival in patients referred for evaluation of heart failure. *Ann Intern Med* 1998; **129**: 286–93.
- 22 Gibbons RJ, Balady GJ, Beasley JW, et al. ACC/AHA guidelines for exercise testing: executive summary: a report of the American College of Cardiology/American Heart Association task force on practice guidelines. *Circulation* 1997; **96**: 345–54.
- 23 Chua TP, Ponikowski P, Harrington D, et al. Clinical correlates and prognostic significance of the ventilatory response to exercise in chronic heart failure. *J Am Coll Cardiol* 1997; **29**: 1585–90.
- 24 Ellestad MH, Wan MK. Predictive implications of stress testing: follow-up of 2700 subjects after maximum treadmill stress testing. *Circulation* 1975; **51**: 363–69.
- 25 Lauer MS, Francis G, Okin PM, Pashkow FJ, Snader CE, Marwick TH. Impaired chronotropic response to exercise stress testing as a predictor of mortality. *JAMA* 1999; **281**: 524–29.
- 26 Schwartz PJ, La Rovere MT, Vanoli E. Autonomic nervous system and sudden cardiac death. Experimental basis and clinical observations for post-myocardial infarction risk stratification. *Circulation* 1992; **85** (suppl 1): 177–91.
- 27 Lauer MS, Larson MG, Evans JC, Levy D. Association of left ventricular dilatation and hypertrophy with chronotropic incompetence in the Framingham Heart Study. *Am Heart J* 1999; **137**: 903–09.
- 28 Gauri A, Raxwal V, Fearon W, Froelicher V. Do beta blockers and inadequate heart rate response impair the diagnostic accuracy of the standard exercise test? *Circulation* 1999; **100**: 1505.
- 29 Imai K, Sato H, Hori M, et al. Vagally mediated heart rate recovery after exercise is accelerated in athletes but blunted in patients with chronic heart failure. *J Am Coll Cardiol* 1994; **24**: 1529–35.
- 30 Cole CR, Blackstone EH, Pashkow FJ, Snader CE, Lauer MS. Heart-rate recovery immediately after exercise as a predictor of mortality. *N Engl J Med* 1999; **341**: 1351–57.
- 31 Singh JP, Larson MG, Manolio TA, et al. Blood pressure response during treadmill testing as a risk factor for new-onset hypertension: the Framingham heart study. *Circulation* 1999; **99**: 1831–36.
- 32 Kjeldsen SE, Mundal R, Sandvik L, Erikssen G, Thaulow E, Erikssen J. Exercise blood pressure predicts cardiovascular death and myocardial infarction. *Blood Press Monit* 1997; **2**: 147–53.
- 33 McHam SA, Marwick TH, Pashkow FJ, Lauer MS. Delayed systolic blood pressure recovery after graded exercise: an independent correlate of angiographic coronary disease. *J Am Coll Cardiol* 1999; **34**: 754–59.
- 34 Campbell L, Marwick TH, Pashkow FJ, Snader CE, Lauer MS. Usefulness of an exaggerated systolic blood pressure response to exercise in predicting myocardial perfusion defects in known or suspected coronary artery disease. *Am J Cardiol* 1999; **84**: 1304–10.
- 35 Morris CK, Ueshima K, Kawaguchi T, Hideg A, Froelicher VF. The prognostic value of exercise capacity: a review of the literature. *Am Heart J* 1991; **122**: 1423–31.
- 36 Blair SN, Kohl HW 3rd, Barlow CE, Paffenbarger RS Jr, Gibbons LW, Macera CA. Changes in physical fitness and all-cause mortality: a prospective study of healthy and unhealthy men. *JAMA* 1995; **273**: 1093–98.
- 37 Snader CE, Marwick TH, Pashkow FJ, Harvey SA, Thomas JD, Lauer MS. Importance of estimated functional capacity as a predictor of all-cause mortality among patients referred for exercise thallium single-photon emission computed tomography: report of 3,400 patients from a single center. *J Am Coll Cardiol* 1997; **30**: 641–48.
- 38 Goldstein RE, Holmboe ES. Prognostic indicators for coronary artery disease: ready for the bedside? *JAMA* 1999; **281**: 565–66.
- 39 Lynch A. Thought contagion: how belief spreads through society: 1. New York: Basic Books, 1996.
- 40 Myers J, Voodi L, Umann T, Froelicher V. A survey of exercise testing: methods, utilization, interpretation and safety in the VAHCS. *J Cardiopulmonary Rehab* 2000; **20**: 251–58.
- 41 American College of Sports Medicine. Guidelines for exercise testing and prescription, 6th edn. Baltimore: Lippincott, Williams & Wilkins, 2000.
- 42 Panza JA, Quyyumi AA, Diodati JG, Callahan TS, Epstein SE. Prediction of the frequency and duration of ambulatory myocardial ischemia in patients with stable coronary artery disease by determination of the ischemic threshold from exercise testing: importance of the exercise protocol. *J Am Coll Cardiol* 1991; **17**: 657–63.
- 43 Tamesis B, Stelken A, Byers S, et al. Comparison of the Asymptomatic Cardiac Ischemic Pilot and modified Asymptomatic Cardiac Ischemia Pilot versus Bruce and Cornell exercise protocols. *Am J Cardiol* 1993; **72**: 715–20.
- 44 Myers J, Buchanan N, Walsh D, et al. Comparison of the ramp versus standard exercise protocols. *J Am Coll Cardiol* 1991; **17**: 1334–42.
- 45 Buchfuhrer MJ, Hansen JE, Robinson TE, Sue DY, Wasserman K, Whipp BJ. Optimizing the exercise protocol for cardiopulmonary assessment. *J Appl Physiol* 1983; **55**: 1558–64.
- 46 Redwood DR, Rosing DR, Goldstein RE, Beiser GD, Epstein SE. Importance of the design of an exercise protocol in the evaluation of patients with angina pectoris. *Circulation* 1971; **43**: 618–28.
- 47 Will PM, Walter JD. Exercise testing: improving performance with a ramped Bruce protocol. *Am Heart J* 1999; **138**: 1033–37.
- 48 Hambrecht RP, Schuler GC, Muth T, et al. Greater diagnostic sensitivity of treadmill versus cycle exercise testing of asymptomatic men with coronary artery disease. *Am J Cardiol* 1992; **70**: 141–46.

A list of further reading appears on *The Lancet's* website: [www.thelancet.com](http://www.thelancet.com)