A simple and portable paced step test for reproducible measurements of ventilation and oxygen consumption during exercise

P W JONES, J M WAKEFIELD, E KONTAKI

From the Department of Medicine, Middlesex Hospital, London

ABSTRACT A single work rate paced step test is described. It was designed to allow the measurement of minute ventilation (\(\dot{V}_t\)) and oxygen consumption (\(\dot{V}_{O_2}\)), under standardised conditions, during exercise in the clinic and lung function laboratory and in field work. The subjects and the operator found the test simple to perform. The values for ventilation at a given oxygen consumption were similar to those from more complex and stressful tests and had a high degree of reproducibility. The ergometer was a 12 inch (30 cm) step with hand rails. The subjects stepped up and down, once every four seconds in response to a buzzer for 10 minutes. \(\dot{V}_t\) and \(\dot{V}_{O_2}\) were measured with a small portable device. In 53 normal subjects of mean weight 69 (range 49–107) kg, mean \(\dot{V}_{O_2}\) measured during the last 5 minutes of the test was 0.89 (range 0.53–1.52) l/min. Weight and height were independent contributors to the oxygen cost of stepping. The ventilatory equivalent for oxygen (\(\dot{V}_{O_2}\), which equals \(\dot{V}_t\) divided by \(\dot{V}_{O_2}\)) increased with age, and was similar to normal values for ventilation at 1.01/min \(\dot{V}_{O_2}\) obtained by workers who used multiple work rate tests on more complex equipment. The within subject variation in \(\dot{V}_{O_2}\), expressed as the coefficient of variation (cv) and calculated for each age decade, was 14%. This again compared very favourably with the cv for values of ventilation at 1.01/min \(\dot{V}_{O_2}\) from other studies. In 20 normal subjects daily testing showed significant falls in \(\dot{V}_{O_2}\) and \(\dot{V}_t\) over three days of 4% and 5-9% respectively; \(\dot{V}_{O_2}\) fell by 2%. Testing over four consecutive weeks showed small significant differences between weeks but no clear trend; the within subject weekly variation accounted for less than 3.5% of the total variance for each of the variables. Testing at least four months apart showed a 3.3% fall in the oxygen cost of stepping. No significant diurnal effects on the exercise responses were found.

The ventilatory response to exercise is altered by disease states; its relationship to changes in lung volumes and diffusing capacity is complex and may change during the natural history of the disease. Although exercise tests are becoming more widely used in the assessment of pulmonary disease, the use of the ventilatory response as a possible tool for monitoring disease progression or for measuring the effects of treatment remains undervaluated. One obstacle to this is the nature of the test, which frequently requires an unfamiliar form of exercise, relatively complex equipment, and well trained operators, and can usually only be performed within a specific laboratory or on a particular site. We considered that there was a need for a simple test that allows the reproducible measurement of ventilation and oxygen consumption, under controlled conditions of exercise and which could readily be performed in the clinic or a standard lung function laboratory, or in field work.

The paced step test was introduced for the assessment of respiratory insufficiency by Baldwin et al. The ventilatory response to this type of test was standardised by Hugh-Jones and Lambert in terms of external work. During exercise, however, breathing frequency increases as a function of metabolic demand and the oxygen uptake can be predicted only with limited accuracy from calculations of external work; as a result the ventilatory responses standardised in such a way may be quite variable (see Cotes, p 296). We have developed a paced step test in which minute ventilation (\(\dot{V}_t\)) and oxygen uptake (\(\dot{V}_{O_2}\)) were measured with a small portable device. For simplicity, a single work rate was used with the same stepping frequency and step height for all subjects.
To assess the reliability of this test and to validate our method of standardisation of the ventilatory response, we have studied normal subjects to estimate the size of the normal range of response and factors that effect it. We have also looked for any diurnal variation in the ventilatory response to this test and examined its short and medium term reproducibility.

Methods

MEASUREMENT OF VENTILATION AND OXYGEN CONSUMPTION

The Oxylog (PK Morgan)—a portable device weighing 2.6 kg—was used. This equipment consists of a mixing chamber and polarographic oxygen sensors and costs about £3000 in Britain. Respired gas is collected with a tight fitting face mask. Inspired air flow is measured with a turbine flow meter, which is calibrated and set at the factory. Minute inspired ventilation (Vi) corrected to BTPS is measured minute by minute. The mixed expired oxygen concentration is calculated from the difference between the outputs of two oxygen sensors, one of which samples the mixed expired gas and the other room air. Before each test the oxygen sensors were calibrated with room air. Setting of the zero oxygen level on fitting new sensors was performed with oxygen free nitrogen. A random check with nitrogen showed no zero drift. The life-span of the sensors is usually several months and they were not changed during the course of a study.

Within the Oxylog oxygen consumption (Vo2), corrected to STPD, is calculated from Vi and the mixed expired oxygen concentration by a microprocessor, which assumes a respiratory gas exchange ratio of 0.8. This assumption is necessary because inspired and not expired gas flow is measured. The effect of this on the measurement of oxygen uptake is small and can be calculated (see below under “Discussion”).

During the exercise test the Oxylog was placed beside the step and not carried by the subject.

ACCURACY OF MEASUREMENTS OF VENTILATION AND OXYGEN UPTAKE

In five subjects measurements made with the Oxylog were compared simultaneously with measurements made by means of computerised equipment (Gould 9000 IV) that incorporated a dry rolling seal spirometer and gas analysers. The latter equipment was calibrated before each test with a 31 precision syringe and gases of known composition, analysed with a Lloyd-Haldane apparatus. For the comparison between the two sets of equipment subjects performed a progressive exercise test of 25 watt increments every four minutes up to 100 watts on a cycle ergometer. Expired gas passed in series first to the Oxylog and then to the Gould equipment. Measurements made in the third and fourth minute of each work rate were recorded.

PACED STEP TEST

The subjects mounted a 12 inch (30 cm) platform once every four seconds in response to an audible timer. They were instructed to mount the platform with both feet on hearing the timer then immediately dismount and await the next signal. Arm rails were available for use if the subjects wanted them. Before the test the face mask was placed on the subjects and the straps were adjusted to allow a comfortable tight fit. Obvious leaks were corrected but deliberately no formal sensitive tests for small leaks were made. After the face mask had been fitted the subjects were asked to sit for four minutes and then told to begin exercising. Measurements of Vi and Vo2 were recorded every minute. No individuals with particular facial features (for example, with beards) were excluded from the study.

REPRODUCIBILITY OF MEASUREMENTS

1 Twenty normal subjects (seven men, three with beards) were recruited from hospital staff. They were not familiar with the mask or testing technique. Their mean age was 30 (range 22–39) years. These subjects were studied on three consecutive days at least two hours after a meal.

2 Eighteen of these subjects were also studied once a week for four consecutive weeks.

3 Seven subjects were studied on two separate days four months after the initial study.

INFLUENCE OF DIURNAL VARIATION

Seven normal subjects each performed two pairs of tests. Each pair consisted of a test between 8 and 10 am and a test between 5 and 6 pm. Each subject was tested once with the morning test as the first of the pair and once with the evening test first. The sequence in which these were performed (that is, morning–evening or evening–morning) was randomised between subjects. In the event, four subjects performed the morning–evening sequence first.

FACTORS AFFECTING THE RESPONSE TO THE STEP TEST

To examine factors that may create variations between subjects in the response to the test, a further 53 subjects (29 of them women) were each studied on one occasion. The subjects had no past or present history of lung disease, although a rigorous history of cough frequency was not taken. They were recruited from hospital staff and from a physiotherapy outpatient clinic. The subjects were recorded as smokers (current smokers or ex-smokers) and non-smokers
(that is, they had never smoked). Their heights, weights, and ages were recorded.

**ANALYSIS OF RESULTS**

The results were analysed by means of analyses of variance and covariance: a linear modelling package (GLIM 3.11, Royal Statistical Society, London) was used.

**Results**

**COMPARISON OF MEASUREMENTS OF $\dot{V}O_2$ AND $\dot{V}O_2$ ON TWO SETS OF EQUIPMENT**

Comparisons were made in five subjects up to a maximum $\dot{V}O_2$ of 331/min and maximum $\dot{V}O_2$ of 1.301/min. $\dot{V}O_2$ measured with the Oxylog was on average 0.83 (SEM 0.21) l/min less than the $\dot{V}O_2$ measured with the Gould equipment ($p < 0.001$). There was no significant difference in the measured $\dot{V}O_2$ (mean difference 16 (13) ml/min). Figure 1 shows the difference in $\dot{V}O_2$ between Oxylog and the Gould equipment and the $\dot{V}O_2$ and the R value measured on the Gould equipment at each work rate during cycle ergometry: the difference between the two measurements of $\dot{V}O_2$ did not change with increasing $\dot{V}O_2$ or with changing R value.

**TIME COURSE OF THE RESPONSES OF $\dot{V}I$ AND $\dot{V}O_2$ DURING 10 MINUTES OF STEPPING**

In the 20 subjects studied repeatedly $\dot{V}O_2$ reached 95% of the maximum value recorded in any one minute within five minutes of the onset of exercise and reached a plateau by the sixth minute, with no significant change with time thereafter ($p > 0.05$). In contrast, $\dot{V}I$ reached 95% of its maximum recorded value six minutes from the start of exercise and continued to increase throughout the exercise period, showing a significant trend against time over the last five minutes of the test period ($p < 0.01$). The mean results (not normalised) from the 20 subjects are shown in figure 2. The effect of age on the time course of the responses was assessed by comparing the responses from two cohorts drawn from the group of subjects who were studied on one occasion only. One cohort had an age range of 20–35 years ($n = 23$) and the other of 50–65 years ($n = 15$). No differences
Figure 3  Mean changes in ventilation, oxygen consumption, and ventilatory equivalent for oxygen during exercise in repeated testing. In daily testing (left hand side) significant falls in all three variables were seen. In weekly testing (right hand side) a significant trend for a reduction in ventilation and oxygen consumption was also seen. For details of statistical significance see text.

between these two subgroups were found in the time course of the responses of either \( V_t \) or \( \dot{V}O_2 \).

**Day to Day Variation**

In each of the 20 subjects measurements from the last five minutes of the exercise period were used in the analysis of variance. There was a significant fall in \( V_t \) and \( \dot{V}O_2 \) during repeated testing over three days (both \( p < 0.001 \); fig 3). The ventilatory equivalent for oxygen (\( \dot{V}I_o_2 \)), calculated from \( V_t \) divided by \( \dot{V}O_2 \), also fell (\( p < 0.01 \)). Although these trends were significant they were small. The falls by day 3 as a percentage of day 1 were: \( V_t \) 5-9%, \( \dot{V}O_2 \) 4%, and \( \dot{V}I_o_2 \) 2-0%.

**Week to Week Variation**

Measurements from the last five minutes of exercise were recorded from 18 subjects, tested once a week over four weeks (fig 3). There were significant differences between weeks in all three variables (\( p < 0.01 \)) and a significant covariance between the test number and both \( V_t \) and \( \dot{V}O_2 \) (\( p < 0.05 \) in each case). These differences and trends were very small; for each of the three variables the analysis of variance showed that this week to week variation accounted for less than 3.5% of the total within subject variation.

**Long Term Changes**

In seven subjects the mean \( \dot{V}O_2 \) values recorded from their first three tests were compared with the mean of the two tests performed after an interval of at least four months. There was a significant fall of 0.03 (SD 0.03) l/min (\( p < 0.01 \)); this corresponded to an overall fall of 3.3%. The mean change in \( V_t \) was 2.7% and the change in \( \dot{V}I_o_2 \) 2.0%; neither of these changes reached significance (\( p > 0.05 \)).

**Diurnal Variation**

In the seven subjects studied no significant differences (\( p > 0.05 \)) were found between morning and evening exercise responses: \( V_t \) was 20.8 morning, 20.71/min evening, SEM of difference 0.21/min; \( \dot{V}O_2 \) was 0.89 morning, 0.901/min evening, SEM of difference 0.01 l/min; \( \dot{V}I_o_2 \) was 23.5 morning, 23.2 evening, SEM of difference 0.2.

**Factors Affecting the Responses to Stepping**

The mean age of the 53 subjects was 50 (range 19-77) years. There were five or more subjects per decade below 60 years of age, but only four aged over 70 years. Mean weight was 69 (range 49-107) kg, mean height 1.69 m (range 1.40-1.84 m). Twenty-one subjects were smokers and 32 non-smokers. Mean 18 during the last five minutes of the step test was 25.6 (range 22.4-38.4) l/min and mean \( \dot{V}O_2 \) was 0.89 (range 0.53-1.52) l/min. Expressed as a proportion of the basal metabolic rate (defined as 1 met = 3.5 mls O\(_2\)/kg body weight), the average work rate imposed by the test was 3.7 (SD 0.6) mets.

Over the age range 20-60 years (for comparison with the values given by Jones and Campbell) \( \dot{V}O_2 \) for men (\( n = 16 \)) was 28.2 (SD 5.0) and for women (\( n = 26 \)) 28.4 (4.2).

**Factors affecting \( \dot{V}O_2 \)**

There was a highly significant correlation between \( \dot{V}O_2 \) and weight (\( r^2 = 0.57, p < 0.001 \)) and, although there was a significant interaction between height and weight, there was also a positive correlation (\( p < 0.025 \)) between height and \( \dot{V}O_2 \) that was independent of the weight effect. The relationship between weight, height, and \( \dot{V}O_2 \) could be described using the equation:

\[
\dot{V}O_2 (l/min) = a + b \text{ weight (kg)} + c \text{ height (m)},
\]

where \( a = -0.74 \) (SEM 0.30), \( b = 0.010 \) (0.002), and \( c = 0.54 \) (0.22).

After removal of the height and weight effects, sex, smoking, and age did not significantly contribute to the variance in \( \dot{V}O_2 \).
The factors affecting $V_I$
The factors affecting $V_I$ were similar to those affecting $V_{O_2}$. The correlation with weight was significant ($r^2 = 0.43$, $p < 0.001$). Height did not have a significant independent effect and age, sex, and smoking were also without effect.

Factors affecting $V_{I_{O_2}}$
There was a highly significant covariance between $V_{I_{O_2}}$ and age ($r^2 = 0.21$, $p < 0.001$), although the slope of this relationship was very low (fig 4). Age did not significantly interact with any other factor in determining the $V_{I_{O_2}}$, and none of the other factors individually contributed more than 3-2% to the total variance. $V_{I_{O_2}}$ could be related to age with the equation:

$$V_{I_{O_2}} = a + b \text{ age},$$

where $a = 23.7$ (SEM 1.6) and $b = 0.134$ (0.036).

Discussion
This exercise test was designed to be simple and to be used on a wide population of adult patients. The selection of any one work rate is a compromise between setting an unsustainable work rate for certain subjects and an insufficient load for others. The stress produced by the test was relatively modest; 3.7 mets is equivalent to walking at 3-4 mph or the performance of housework. The small standard deviation (0.6 mets) shows that although the subjects' weights varied over a twofold range the relative stress for each subject was well standardised and all these normal subjects were able to complete 10 minutes of stepping regardless of age or physical characteristics.

The subjects were allowed to use the handrails if they wished. All work against gravity means energy consumption (and therefore oxygen uptake), so the use of handrails did not remove the metabolic load on the subjects. For a given level of oxygen uptake arm exercise provokes a slightly greater ventilatory response than leg exercise, but the use of handrails was allowed because patients with lung disease frequently preferred using them during this test.

In a step test the external work is the product of the step height, the frequency of stepping, the subject's weight, and a constant. We found, however, that the subjects' height was also positively correlated with oxygen consumption (a measure of internal work), independently of its relationship to weight. Shahnavaz has shown that the oxygen cost of stepping varies as a function of the ratio of leg length to platform height. He found that the optimum ratio was 50-55% and that in any given subject, at a given work rate, ratios less than this were associated with an increase in oxygen consumption. Our observation of differences between subjects accords with his findings, since total body height is quite closely related to leg length, and our 12 inch (30 cm) step was invariably less than half the length of the subjects' legs.

These observations illustrate the well recognised importance of measuring, rather than calculating, the oxygen consumption when ventilation is being standardised. In our study, although weight and height were statistically significant predictors of oxygen consumption, together they still accounted for only 60% of the total variance in measured oxygen uptake. Prediction equations for other forms of exercise are similarly imprecise; for example, with cycle ergometry, 2SD at 1.01/min oxygen uptake would be 18% (from Cotes, equation 12.11).

The measurements of oxygen uptake obtained with the Oxylog are open to slight error because of the assumption, used by the microprocessor in the equipment, that the respiratory gas exchange ratio (R) is a constant with a value of 1.0. The error due to this assumption can be easily calculated: if the true R value were 0.8 the oxygen uptake would be underestimated by 3.5%, and if the R value were 0.9 the underestimate would be 1.8%. In practice, the reported differences between Oxylog measurements and those made simultaneously on conventional equipment at oxygen uptakes below 1.51/min are about 4%, according to the calculations of Belyavin et al. These authors concluded that not all the error could be ascribed to the "R value effect" and inspection of their data shows that the error increased slightly above 1.51/min oxygen uptake. In our comparison no significant differences were found between the two measurements, although the Oxylog appeared slightly to overestimate rather than underestimate.
VO₂. We could find no significant trend for changing errors in Oxylog measurement of VO₂ with change in R (fig 1) and conclude that other measurement errors in the two sets of equipment were greater than the "R value effect." Consideration of this effect of R is important because it changes during exercise (as shown in fig 1). At the start of exercise VO₂ reaches a steady state within 3-4 minutes and in this step test VO₂ as measured by the Oxylog remained constant after four minutes of exercise. In contrast, we recorded a 5% rise in Vl beyond this period. Carbon dioxide production (VCO₂) follows ventilation very closely, so it will have increased by up to 5% (less if alveolar carbon dioxide fell with the increase in breathing). This will have caused the R value (VCO₂/VO₂) to rise by ≤0.05 at the work rate and VO₂ of the step test. The resulting change in the Oxylog measurement of VO₂ will have been an increase of 0.9%, almost undetectable. These errors in VO₂ measurement with the Oxylog, both theoretical and measured, are acceptable and are very much less than the error to be expected when, for example, VO₂ is calculated from treadmill speeds and gradients (Cotes,5 p 303).

To standardise the measurement of ventilation during exercise it is usual to express it as a function of the oxygen consumption, since breathing increases with the increase in metabolism. The regression line between ventilation and oxygen uptake has a positive intercept on the ventilation axis, so the ratio between the two will vary and decrease with increasing levels of oxygen consumption (see Cotes5 for full exposition). For this reason, the preferred practice is to express ventilation at a given level of VO₂ (typically 1.0l/min). Ventilation at this level is estimated by interpolation between measurements at levels of work above and below the desired oxygen uptake. This was not possible with our single work rate test, so ventilation was expressed as the ratio of the measured oxygen consumption—the ventilatory equivalent for oxygen or VlO₂. For the reasons outlined above the VlO₂ should have greater variability than measurements of ventilation standardised to a fixed value of VO₂. The size of this effect should not be overestimated, however. The normal data for cycle ergometry in our laboratory show that an increase in VO₂ from 0.75 to 1.0l/min will cause the calculated VlO₂ to fall by 3%. In subjects who greatly overventilate because of pulmonary fibrosis an equivalent change in VO₂ would produce a 7% fall in VlO₂, (calculated from data of Spiro et al1). A 3.3% increase in VO₂ in the fixed rate step test, however, could be produced only by a commensurate increase in body weight. Thus in a given patient over a period of time, only very large changes in weight would appreciably alter the VlO₂ even in patients with a very disturbed ventilatory response to exercise. If the variability in the oxygen uptake contributed significantly to the variation in VlO₂ between subjects, we would have expected to find a significant correlation between VlO₂ and weight, since the greater the weight the greater would have been the VO₂. No significant correlation between weight and VlO₂ was found; indeed, weight accounted for only 0.8% of the total variance in VlO₂.

Our technique of standardising exercise ventilation should also have slightly widened the range of values for the normal ventilatory response to exercise, but we have little evidence that this occurred. The mean value for the standard deviation of VlO₂ calculated for each age decade was 3.8 and the range of these deviations was 2.4–6.0. The mean coefficient of variation per decade was 14% (range 9.19%). This compares very favourably with the SD of 4.4l/min ventilation at 1.0l/min oxygen consumption, with a coefficient of variation of 16%, calculated from the normal range data presented by Spiro et al.1 We therefore conclude that although the use of the ventilatory equivalent for oxygen may in theory widen the range of normality, other biological factors appear to predominate in determining intersubject variation in the ventilatory response to exercise. We suggest therefore that the VlO₂ appears to provide a valid and reliable way of standardising the response to this test.

The ventilatory response to stepping, as expressed in terms of VlO₂, increased significantly with age. This is in keeping with the findings from studies that have used conventionally standardised responses.12–15 Sex and smoking were without effect on VlO₂. The values of VlO₂ (obtained at a mean VO₂ of 0.89l/min) were a little higher than our own and other published values1–8 for normal ventilation during cycle ergometry. Jones and Campbell10 suggest a normal value of 261l/min for ventilation at 1.0l/min VO₂ in healthy men 20–60 years and 27–51l/min for women. Our values in subjects over the same age range were only a little higher. One factor responsible for this may have been the time during the exercise at which measurements were made. Oxygen uptake follows changes in work rate during exercise more quickly than does ventilation16; many measurements of ventilation and oxygen uptake are made during non-steady-state one minute increment progressive exercise tests, so normal values obtained in this way1 would produce a lower value of ventilation for a given oxygen uptake than more steady-state measurements such as ours. It is usually recognised that below the "anaerobic threshold" ventilation reaches a steady state after three to five minutes, but we noted a continued slight drift upwards after this time with the step test, although not during walking with Oxylog.10 This may have been due to lactic acid accumulation in the blood. Differences in blood lactate concentrations
for a given level of oxygen uptake between cycle ergometry and treadmill exercise have been described, but we know of no comparable studies concerning stepping. The upward drift of ventilation over the last five minutes was less than 5% of the total response so we thought that the greater precision achieved by estimating the response to exercise from the mean of the last five one minute measurements outweighed any advantage accruing from the use of a single maximum value or the value from the last minute.

The reproducibility of the measurements of ventilation and oxygen consumption from the step test on repeated testing was good and better than the reproducibility in normal subjects measured during repeated incremental exercise tests with more complex equipment. Small leaks around the face mask may have occurred, but we did not formally test for these because it would not usually be possible to do so under the conditions for which the test was designed. The stability of the results and the similarity between our values for ventilation and other published values suggest that any leaks were small and not important. Small but statistically significant differences were found in repeated testing over several weeks. Subtle changes in the pattern of stepping may have occurred for the small fall in oxygen consumption during both the day to day and the weekly testing, and, although a long-term drift in the characteristics of the oxygen sensors cannot be excluded, a random check on the zero level showed no drift. The size of this fall during several months was, however, very small. Possibly small weight changes occurred in the subjects, but weight changes were not recorded. The fall in ventilation seen with repeated daily testing appears to have been due to a fall in oxygen consumption rather than a result of the subject’s acclimatisation to the face mask, because the size of the change in the ventilatory equivalent for oxygen and its variance were rather less than the equivalent changes in ventilation and oxygen consumption. There are few appropriate published data with which to compare our findings on the reproducibility of exercise test results and possible trends in repeated testing. A 12% fall in ventilatory response between the first two cycle ergometer tests of a series has been reported in a study on a small number of subjects who were breathing into a mouthpiece with a noseclip, but using similar techniques and more subjects we have not observed such a change. The changes seen with repeated weekly step testing are unlikely to pose a problem of interpretation in intervention studies, but the short term changes seen with repeated daily testing must be taken into consideration when acute experiments are being planned.

In conclusion, the single work rate, paced step test provides a standardised and reproducible technique for the measurement of ventilation and oxygen consumption at an oxygen consumption around 1.0 l/min in adult subjects of average build. It is simple and portable, requires little training of patient or operator, and produces values for ventilation, expressed as a function of oxygen consumption, very similar to those from incremental tests using more complex and less mobile equipment and with a higher level of reproducibility.

We thank Dr A J Sargeant (Polytechnic of North London) for the loan of the Oxylog. EK was in receipt of a Greek state scholarship. PWJ was a Wellcome senior research fellow in clinical science.

References

Paced step test for reproducible measurements of ventilation and oxygen consumption during exercise


