

The Chester step test—a simple yet effective tool for the prediction of aerobic capacity

Kevin Sykes*, Alison Roberts

Centre for Exercise & Nutrition Science, University College Chester, Parkgate Road, Chester CH1 4BJ, UK

Abstract

Objectives The assessment of cardiorespiratory fitness is becoming more commonplace in both community and occupational health settings. This study investigates the reliability and validity of the Chester step test, a submaximal test for the prediction of aerobic capacity, when compared to maximal oxygen uptake ($\dot{V}O_2\text{Max}$) measured during a treadmill test.

Design Participants completed a $\dot{V}O_2\text{Max}$ Treadmill Test using a standardised fast incremental ramp protocol designed to elicit exhaustion in 8–12 minutes. Following this, on separate days, subjects then completed the Chester step test (CST) on two occasions (CST1 and CST2). During the submaximal step test, subjects were asked to step on to and off a 30-cm step at a rate set by the metronome beat on an audio cassette. The initial step rate was 15 steps per minute and every 2 minutes the tempo increased by 5 steps per minute. The subject continued stepping until he/she reached 80% of their maximum predicted heart rate, or reported a rating of perceived exertion of 15 (hard) on the Borg scale, or reached the end of the 10-minute 5-stage test.

Setting Human Performance Laboratory, University College Chester

Participants Sixty-eight subjects (mean age 30.6 ± 9.7 years; range 18–52 years) with a wide range of ages and abilities

Main outcome measures Gas exchange and heart rate were measured continuously using a Polar heart rate monitor. Heart rate and rating of perceived exertion were recorded after each 2-minute stage.

Results Results revealed a high correlation ($r = 0.92$) between $\dot{V}O_2\text{Max}$ and CST1 ($P < 0.001$) and a standard error of the estimate of $3.9 \text{ ml O}_2/\text{kg}/\text{min}$, thus confirming the face validity of CST as a predictor of $\dot{V}O_2\text{Max}$. Using the method of analysis recommended by Bland and Altman (1986) [Lancet 5 (1986) 307] the mean difference between repeated predicted measures using CST was $-0.7 \text{ ml O}_2/\text{kg}/\text{min}$. The limits of agreement analysis also demonstrated that a measurement repeated on a separate day was within $4.5 \text{ ml O}_2/\text{kg}/\text{min}$ of the original predicted measurement. The Chester step test is therefore appropriate for use in situations where a change in aerobic capacity is expected to be more than $3.8 \text{ ml O}_2/\text{kg}/\text{min}$ higher or more than $5.2 \text{ ml O}_2/\text{kg}/\text{min}$ lower than the baseline measurement.

Conclusions The Chester step test was shown to be a valid test for the estimation of aerobic capacity within this group. The error of measurement is sufficiently small and suggests that this method is well suited to monitoring changes in aerobic capacity in rehabilitation settings.

© 2004 Chartered Society of Physiotherapy. Published by Elsevier Ltd. All rights reserved.

Keywords: Chester step test; Reliability; Validity; $\dot{V}O_2\text{Max}$

Background

Regular exercise, which improves cardiorespiratory fitness, is a powerful factor in enhancing health and wellbeing. Not only can those with active lifestyles gain substantial health benefits from physical activity but regular exercisers also tend to have healthier lifestyles compared to non-exercisers—being less likely to smoke, suffer fewer mi-

nor illnesses, make fewer visits to the doctor and generally eat more healthily.

The assessment of cardiorespiratory fitness commonly assessed by the measurement of maximum oxygen uptake ($\dot{V}O_2\text{Max}$) or aerobic capacity, can be a useful tool in promoting health. For example, it enables a baseline of fitness to be established, it enables an exercise programme to be more individually prescribed, it provides a system for monitoring change and it may be used as a health risk indicator [1]. In community and occupational health settings, the assessment of aerobic capacity is becoming more commonplace both in medical screening and as a measure and monitor of functional status [2].

* Corresponding author. Tel.: +44-1244-392734; fax: +44-1244-392735.

E-mail address: k.sykes@chester.ac.uk (K. Sykes).

The classical method of measurement of cardiorespiratory fitness (or aerobic capacity) is by direct measurement of $\dot{V}O_2\text{Max}$, where the subject undergoes a maximal exercise test on a cycle or treadmill and oxygen consumption is measured directly. Whilst this is the gold standard, the equipment is expensive, requires a high level of technical expertise and supervision, is impractical in non-laboratory and field-test situations and is unsuitable for those individuals for whom exhaustive exercise is not recommended.

As a result numerous tests have emerged for the estimation of aerobic capacity. Some are field tests requiring maximum effort, for example the 20-m multistage shuttle run [3], whilst others are submaximal treadmill, cycle or bench-stepping tests with single stage or multistage protocols. For a full review, see [4].

The Chester step test (CST) [5] was designed as a submaximal, multistage test, where heart rate and exertion levels are monitored continuously. Since the test is stopped when the subject reaches approximately 80% of maximum heart rate, estimated as 220 minus the age of the subject ($220 - \text{age}$), and can be used with step heights of 15, 20, 25 and 30 cm, it is highly adaptable and applicable to a wide range of ages, abilities and conditions. It is inexpensive, easy to standardise, highly portable and safely controlled, therefore applicable for use by exercise and health professionals in a variety of workplace and community settings.

The physiological rationale is based on the linear relationship between workload, oxygen consumption and submaximal exercise heart rates, which enable a prediction of aerobic capacity to be made, either graphically or by a statistical line of best fit [1]. The accuracy of the American College of Sports Medicine's stair-stepping equation to predict oxygen cost ($\text{ml O}_2/\text{kg}/\text{min}$) at a given step height and rate was confirmed by Latin et al. [6].

Cook [7] investigated the relationship between direct measurement of $\dot{V}O_2\text{Max}$ during a treadmill test, with 26 highly active sports science subjects (mean age 21.4 ± 2.4 years), and with CST conducted on two separate occasions. Results from the Treadmill $\dot{V}O_2\text{Max}$ test ($60.5 \pm 9.96 \text{ ml O}_2/\text{kg}/\text{min}$) were not significantly different from the CST results (56.9 ± 7.95) ($P > 0.05$) and showed a correlation coefficient of $r = 0.92$. These results compared favourably with other step tests—e.g. [8–11] compared results from two established tests, the Astrand cycle test and the 20-m multistage shuttle run [3] in 20 active sport science students (mean age 21.5 ± 2.2 years). Results showed high correlation between CST and these other tests (CST/Astrand cycle test: $r = 0.94$, $P < 0.01$; CST/multistage shuttle run: $r = 0.81$, $P < 0.05$) with mean scores of CST ($54 \pm 9 \text{ ml O}_2/\text{kg}/\text{min}$), Astrand cycle test ($52 \pm 9 \text{ ml O}_2/\text{kg}/\text{min}$) and multistage shuttle run ($51 \pm 10 \text{ ml O}_2/\text{kg}/\text{min}$) with no significant differences observed between the means ($P > 0.05$). Whilst this previous study showed CST to have high concurrent validity with previously validated tests, the purpose

of the current study was to investigate the prediction accuracy of CST in relation to directly measured treadmill $\dot{V}O_2\text{Max}$ values taken on a wide range of ages and fitness levels.

Rationale for the study

To investigate the validity of CST by comparing the results obtained with the results of a Treadmill $\dot{V}O_2\text{Max}$ test and to estimate its reproducibility with apparently healthy subjects from a wide cross-section of ages and fitness levels.

Methods

Subjects

Sixty-eight subjects (mean age 30.6 ± 9.7 years; range 18–52 years) completed a Treadmill $\dot{V}O_2\text{Max}$ test on one occasion and the CST on a further two separate occasions. All subjects were volunteers, who were apparently healthy, free from medical contraindication to vigorous exercise and not taking any form of medication that would depress heart rate scores. All subjects signed an informed consent form.

Pre-test conditions

Subjects were asked to refrain from eating, smoking or drinking tea, coffee or alcohol for at least 2 hours before the measurement session. They were also requested not to take exercise for 24 hours prior to testing to ensure a consistent baseline activity level.

Treadmill $\dot{V}O_2\text{Max}$ test

The $\dot{V}O_2\text{Max}$ test was conducted on a Treadmill (HP Cosmos Pulsar, Nussdorf-Traunstein, Germany), using a standardised fast incremental ramp protocol, commencing on the horizontal and increasing the gradient by 1% every minute, at a running speed designed to elicit exhaustion in 8–12 minutes according to guidelines from the British Association of Sport and Exercise Sciences [12]. Subjects were familiarised with walking and running on a treadmill during a 5-minute warm-up period. Expired gases were measured by using a Quark breath-by-breath gas analysis unit (CosMed, Italy) with continuous monitoring of ventilation, carbon dioxide production, oxygen consumption and heart rate. $\dot{V}O_2\text{Max}$ was determined from the mean reading calculated from the final 30 s of exercise. Criteria for $\dot{V}O_2\text{Max}$ being reached were that two of the following should be met: (i) heart rate reached age-related maximum ($220 - \text{age}$), (ii) a respiratory exchange ratio of 1.15 was reached [12], and (iii) the subject reported a rating of perceived exertion of 19/20 on the Borg scale [13].

6	No exertion at all
7	
8	Extremely light
9	Very Light
10	
11	Light
12	
13	Somewhat Hard
14	
15	Hard (heavy)
16	
17	Very Hard
18	
19	Extremely Hard
20	Maximal exertion

Fig. 1. Rating of perceived exertion [13].

Chester step test

The test was conducted using a pre-recorded audiotape with instructions and timed metronome rhythms and a 30-cm step. The subject listened to the test instructions and commenced stepping to the metronome beat at 15 steps per minute for 2 minutes following which heart rate and rating of perceived exertion [13] (Fig. 1) were recorded (Level 1). The step rate then increased to 20 steps per minute for a further 2 minutes, when heart rate and rating of perceived exertion were again recorded (Level 2). The test continued in this progressive manner until the subject reached a heart rate of 80% of predicted maximum ($220 - \text{age}$). Providing the subject showed no overt signs of distress and that the rating of perceived exertion was below 15, the subject was permitted to finish the 2-minute stage and the test was then terminated. The maximum test duration was 10 minutes (i.e. Level 5). Aerobic capacity was then predicted by plotting the exercise heart rates on the prepared graphical datasheet (Appendix A), drawing a visual line of best fit between the data points, projecting the line up to maximum heart rate and estimating the corresponding aerobic capacity ($\text{ml O}_2/\text{kg}/\text{min}$) from the x -axis.

The step test was conducted on two separate days within 1 week (CST1 and CST2).

Results

The subjects were a group of apparently healthy males and females from a wide range of ages (18–52 years) and fitness levels ($25\text{--}68 \text{ ml O}_2/\text{kg}/\text{min}$) (Table 1).

Table 1
Descriptive statistics: mean (\pm S.D.)

Subjects	n	Age (year)	Weight (kg)	Height (cm)	BMI	$\dot{V}\text{O}_2\text{Max}$ ($\text{ml O}_2/\text{kg}/\text{min}$)	CST1 ($\text{ml O}_2/\text{kg}/\text{min}$)	CST2 ($\text{ml O}_2/\text{kg}/\text{min}$)
Males	47	33.2 (10.2)	80.2 (9.6)	179 (7.2)	25.0 (3.1)	54.5 (8.7)	53.2 (7.7)	53.9 (7.6)
Females	21	24.9 (4.6)	64.1 (8.0)	164.1 (5.5)	23.9 (3.3)	46.6 (10.3)	45.2 (9.5)	45.9 (9.6)
Total group	68	30.6 (9.7)	75.2 (11.8)	174.4 (9.6)	24.7 (3.2)	52.1 (9.9)	50.7 (9.0)	51.4 (9.0)

Table 2
Pearson's 'r' correlations between $\dot{V}\text{O}_2\text{Max}$ and CST1

Subjects	Variables	n	r	P
Males	$\dot{V}\text{O}_2\text{Max}$ vs. CST1	47	0.87	<0.001
Females	$\dot{V}\text{O}_2\text{Max}$ vs. CST1	21	0.95	<0.001
Whole group	$\dot{V}\text{O}_2\text{Max}$ vs. CST1	68	0.92	<0.001

Table 3
Standard error of the predicted estimate of aerobic capacity ($\text{ml O}_2/\text{kg}/\text{min}$)

Subjects	CST1
Males	4.3
Females	3.1
Whole group	3.9

Table 4
Test–retest analysis of CST1 and CST2 revealing the bias and 95% limits of agreement Bland and Altman (1986) [14]

Variables	N	Bias ($\text{ml O}_2/\text{kg}/\text{min}$)	95% limits of agreement ($\text{ml O}_2/\text{kg}/\text{min}$)
CST1 vs. CST2	68	−0.7	4.5

Table 2 illustrates the high overall correlation ($r = 0.92$; $P < 0.001$) between $\dot{V}\text{O}_2\text{Max}$ and the results of the CST, with the marginally higher values for females ($r = 0.95$; $P < 0.001$) than for males ($r = 0.87$; $P < 0.001$). The regression equation ($\dot{V}\text{O}_2\text{Max} = 0.964 \times 1.007(\text{CST})$; $P < 0.0005$) further illustrates the ability of CST to predict $\dot{V}\text{O}_2\text{Max}$ (Fig. 2), whilst the standard error of the predicted estimate for CST1 was $\pm 3.9 \text{ ml O}_2/\text{kg}/\text{min}$ (Table 3).

The test–retest repeatability of CST was found to be good. Using the method of analysis recommended by Bland and Altman [14] the mean difference between repeated predicted measures was $-0.7 \text{ ml O}_2/\text{kg}/\text{min}$. The analysis also demonstrated that a measurement repeated on a different day was within $4.5 \text{ ml O}_2/\text{kg}/\text{min}$ of the original predicted measurement (Table 4).

Discussion

These results demonstrated that the CST is a valid predictor of aerobic capacity in males and females from a wide range of ages and fitness levels. However, the overall standard error of estimate of $\pm 3.9 \text{ ml O}_2/\text{kg}/\text{min}$ means that its accuracy of prediction, in subjects with aerobic capacity

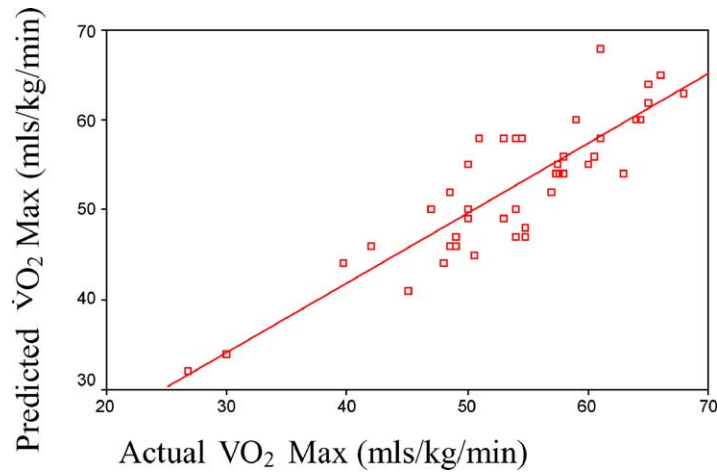


Fig. 2. Relationship between actual and predicted $\dot{V}O_2$ Max during the Chester step test.

values ranging from 25 to 68 ml O_2 /kg/min, is approximately 5–15%. The estimate was found to be slightly more accurate in females than males (standard error of estimate for males and females was ± 4.3 and ± 3.1 ml O_2 /kg/min, respectively). This margin of error is in agreement with earlier work [11] and illustrates the point that CST is best used in situations where an estimate of aerobic capacity is required and not when an exact measure is needed.

The test duration varied from 4 to 10 minutes, depending on the fitness of the individual. Some of the less fit subjects reached 80% of their predicted maximum heart rate at the end of Level 2 (4 minutes) whereupon the test was terminated, whilst some of the fitter individuals were able to complete all five levels (10 minutes) without reaching this target heart rate. There were no observable differences in prediction accuracy between those completing two or three stages and those completing all five stages of the test.

Potential sources of error included (i) the use of a visual line of best fit, (ii) the prediction of maximum heart from $220 - \text{age}$, (iii) the curvilinear relationship of the oxygen consumption-heart rate-workload at near-maximal effort, (iv) accurate reading and recording of heart rate, and (v) the subject's ability to maintain the correct stepping tempo and technique, affecting mechanical efficiency.

Using a statistical line of best fit removes the potential variability in drawing a visual line of best fit [15]. During the testing, the observer was required to hold the Polar heart rate watch rather than attaching it to the wrist of the subject. It was observed that in a small number of instances the pre-recorded instructions, announced after each 2 minutes stage, caused the heart rate of to momentarily elevate. It was important therefore, for the tester to carefully monitor the heart rate to ensure the correct reading was recorded. It was also important that the tester ensured that the subject maintained the correct stepping rate, since deviations from these pre-set rhythms affect the exercise oxygen cost. Whilst most subjects had no difficulty in keeping to the metronome beat, a small number needed more close mon-

itoring and the tester also verbally emphasised the correct rhythm. The pre-recorded instruction to change the lead leg during stepping proved to be straightforward for all subjects to follow. No subject reported any joint pains or untoward distress whilst undertaking the step test and no subject felt unable to continue the test until the end of the stage during which 80% of predicted maximum heart rate had been reached nor reported a rating of perceived exertion greater than 15.

CST is well suited to mass screening events, such as estimating the overall fitness level of a population and its high test-retest reliability, means that a single test is likely to give a valid and meaningful result. CST is also appropriate for use in monitoring changes in aerobic capacity following a rehabilitation programme, particularly in situations where the change is expected to be more than 3.8 or less than 5.2 ml O_2 /kg/min from the baseline measurement.

Conclusions

The Chester step test was shown to provide a valid test for the estimation of aerobic capacity within this group, who were representative of males and females from a wide range of ages and fitness levels. The error of measurement is sufficiently small and suggests that this method is well suited to monitoring changes in aerobic capacity in rehabilitation settings.

Key Messages

- The Chester step test is a valid and reliable tool for the assessment of aerobic capacity.
- It is an inexpensive, easily standardised and portable test, well suited for use in clinical and non-clinical situations where aerobic capacity needs to be assessed and monitored.

References

- [1] American College of Sports Medicine. ACSM's guidelines for exercise testing and prescription. 6th ed. Baltimore: Williams & Wilkins; 2000.
- [2] Sykes K. Promoting health through physical activity. *J Ass Chart Physio Women's Health* 2001;88:6–8.
- [3] Leger LA, Mercier D, Gadoury C, Lambert J. The multistage 20-metre shuttle run test for aerobic fitness. *J Sports Sci* 1988;6:93–101.
- [4] Heyward VH. Advanced fitness assessment and exercise prescription. Human Kinetics; 2002.
- [5] Sykes K. Capacity assessment in the workplace: a new step test. *J Occup Health* 1995;1:20–2.
- [6] Latin RW, Berg K, Kissinger K, Sinnett A, Parks L. The accuracy of the ACSM stair-stepping equation. *Med Sci Sports Ex* 2001;10:1785–7.
- [7] Cook R. An assessment of the reliability and validity of the Chester step test for the estimation of aerobic capacity. M.Sc. Thesis, University College Chester, 1996.
- [8] Siconolfi SF, Garber CE, Lasater TM, Carleton RA. A simple, valid step test for estimating maximal oxygen uptake in epidemiological studies. *Am J Epidemiol* 1985;121(3):382–90.
- [9] Johnson J, Siegel D. The use of selected submaximal step tests in predicting changes in maximal oxygen intake of college women. *J Sports Med* 1981;21:259–64.
- [10] Shaprio A, Shoenfeld Y, Shaprio Y, Sports Med J. Recovery heart rate after submaximal work. *Phys Fitness* 1976;16:57–9.
- [11] Stevens N, Sykes K. Aerobic fitness testing: an update. *J Occup. Health* 1996;48(12):436–8.
- [12] Bird S, Davison R, editors. *Physiological testing guidelines*. 3rd ed. Loughborough: British Association of Sport & Exercise Sciences; 1997.
- [13] Borg GAV. *The RPE scale—a method for measuring perceived exertion*. Stockholm: Borg Perception; 1994.
- [14] Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;5:307–10.
- [15] Sykes K. *Chester step test*. Wrexham: Assist Publications; 2003.

Available online at www.sciencedirect.com

SCIENCE @ DIRECT®