

Effect of encouragement on walking test performance

GORDON H GUYATT, STEWART O PUGSLEY, MICHAEL J SULLIVAN, PENELOPE J THOMPSON, LESLIE B BERMAN, NORMAN L JONES, ERNEST L FALLEN, D WAYNE TAYLOR

From the Department of Clinical Epidemiology and Biostatistics and the Department of Medicine, McMaster University, Hamilton, Ontario, Canada

ABSTRACT Walking tests, frequently used to document effects of treatment on exercise capacity, have never been standardised. We studied the effects of encouragement on walking test performance in a randomised study that controlled for the nature of the underlying disease, time of day, and order effects. We randomised 43 patients with chronic airflow limitation or chronic heart failure or both to receive or not receive encouragement as they performed serial two and six minute walks every fortnight for 10 weeks. Simple encouragement improved performance ($p < 0.02$ for the six minute walk), and the magnitude of the effect was similar to that reported for patients in studies purporting to show beneficial effects of therapeutic manoeuvres. Age and test repetition also affected performance. These results demonstrate the need for careful standardisation of the performance of walking tests, and suggest caution in interpreting studies in which standardisation is not a major feature of the study design.

In 1976 McGavin *et al* introduced the 12 minute walking test as a measure of exercise capacity for patients with chronic lung disease.¹ Since that time walking tests, which are simple, easy to administer, and use an exercise mode relevant to everyday activities have been widely used as measures of outcome in clinical trials. Interventions which have been investigated with walking tests in patients with chronic lung disease include exercise rehabilitation programmes,²⁻⁴ ventilatory muscle training,⁵⁻⁷ and the effects of oxygen^{8,9} and of various drugs, including carbimazole,¹⁰ prednisone,¹¹ naftidrofuryl,¹² bronchodilators,¹³ and sedatives.¹⁴ Although most investigators have used a 12 minute walking test, Butland and coworkers have recently shown that equivalent results may be obtained with six, and possibly two, minute walking times.¹⁵

Despite the extensive use of the test, surprisingly little effort has been put into standardising its performance. In the original description of the 12 minute walk, McGavin *et al* stated that encourage-

ment was given "as necessary." In subsequent experiments using walking tests as measures of exercise capacity encouragement was not held constant. We were concerned that the influence of encouragement might be sufficiently great to rival treatment effects, and therefore we investigated the impact of encouragement on two and six minute walking tests performance in patients with chronic lung and chronic heart disease.

Methods

PATIENTS

We recruited two groups of subjects who experienced fatigue or dyspnoea while performing activities of daily living. The first, with respiratory conditions, attended a regional referral centre for patients with pulmonary problems and had a best recorded forced expired volume in one second (FEV₁) less than 70% of the predicted value. The second group, patients with heart failure, who had been referred by local cardiologists, had impaired left ventricular function demonstrated by angiography, radionuclide scanning, or echocardiography. Exclusion criteria for both groups were as follows: (1) limitation of exercise performance as a result of factors other than fatigue or exertional dyspnoea,

Address for reprint requests: Dr Gordon Guyatt, McMaster University Health Sciences Centre, Room 3H7, 1200 Main Street, West, Room 3H7, Hamilton, Ontario, Canada.

Accepted 21 May 1984

The effect of encouragement on walking test performance

such as arthritis in the leg, angina, or claudication; (2) psychiatric disease preventing reliable performance of the walking test; (3) admission to hospital within the previous two months; (4) previous experience with walking tests.

STUDY DESIGN

We used a randomised block design, in which the blocking factors were level of encouragement, morning or afternoon test performance, and order of test performance. Each patient was tested six times at two week intervals. During each visit both a two minute and a six minute walking test were carried out, with a rest period of at least 20 minutes between tests. In all patients the initial test was administered without encouragement by the supervisor. Patients were then randomly allocated to receive encouragement for all or none of their subsequent five visits.

WALKING TEST PERFORMANCE

The study was conducted over 33 m of an enclosed corridor. Patients were instructed to walk from end to end, covering as much ground as they could during the allotted time. During the walks in which no encouragement was given (E-) the supervisor sat in a chair at one end of the course keeping time, but avoiding eye contact with the subject and maintaining silence. During walks in which encouragement was given (E+) the supervisor ensured that at 30 second intervals she was facing the subject, and she would then deliver one of a predetermined set of encouraging phrases such as "You're doing well" or "Keep up the good work."

STATISTICAL ANALYSIS

The results of this study were assessed by an analysis of variance of the difference scores between the first and subsequent walking tests, taking into account level of encouragement, time of day, and diagnostic group (cardiac or respiratory). Both main effects and interactions were examined. To ensure that the apparent effect of encouragement could not be explained by association with other variables, age, sex, and diagnostic group were correlated with the change in mean score from the first to the last four walks. Encouragement was then added as the final independent variable, and the additional portion of the variance explained by encouragement was calculated. Because Mungall and Hainsworth¹⁶ have previously found improvement over the first three of six repetitions of the 12 minute walk, we examined the effect of test repetition by comparing the scores from the first two visits with those from the last four.

Where appropriate, analyses were conducted for both two and six minute walking scores.

To assess the variability in walking test performance over time, the within person standard deviation for the last four visits (during which the effect of encouragement appeared to be constant) was calculated for both two and six minute walks and compared with the mean walking score for the last four visits and to the mean encouragement effect.

To determine whether scores on the two minute walk were comparable with scores on each two minute segment of the six minute walk, an analysis of variance of the scores on each of the four two minute segments was conducted, with encouragement as the other main factor.

Results

CHARACTERISTICS OF PATIENTS

Of the 57 patients who initially enrolled in the study, 43 (25 in the respiratory group and 18 in the cardiac group) completed the six visits. Of the 14 dropouts, seven had been randomised to the E+ and seven to the E- group; nine were from the cardiac group and five from the respiratory group. Seven dropped out for reasons unrelated to the study; one patient with heart failure declined because of increased symptoms after his first two walking tests, and six patients expressed dissatisfaction with the study.

Forty three subjects completed the study, 34 men and nine women with a mean age of 64.7 (SD 8.3) years. In the 25 respiratory patients the FEV₁ was 0.97 (0.25) litres and the vital capacity 2.4 (0.87) l. Of the 18 cardiac patients, five were in New York Heart Association class II, 12 in class III and one in class IV.

MAIN EFFECTS

In the six minute walk encouragement was associated with a significant increase in distance walked ($p < 0.02$) (fig 1). To calculate the encouragement

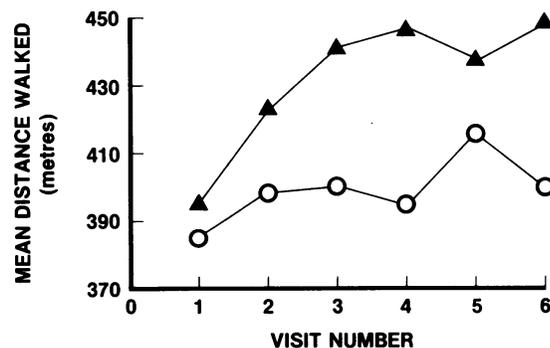


Fig 1 Mean distance walked in the six minute walking test by encouraged (\blacktriangle) and non-encouraged (\circ) groups of patients.

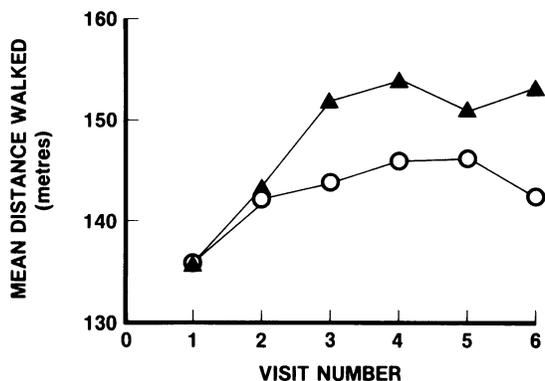


Fig 2 Mean distance walked in the two minute walking test by encouraged (▲) and non-encouraged (○) groups of patients.

effect we corrected the mean scores on visits 3–6 (after performance had plateaued) for baseline performance by subtracting the score for the first walk. The difference between corrected scores in the E+ and E– groups constituted the encouragement effect and averaged 30.5 m. A similar trend towards a positive effect of encouragement on distance walked was seen in the two minute walk (fig 2), but in the analysis of variance of the difference scores this did not reach significance. Time of day and order of test administration did not have any significant effect on change in walking test scores for either two or six minute walks, nor was there a trend associated with these variables.

Two other features affected walking test performance. A group effect was present such that respiratory patients showed greater improvement in walking score over time than did cardiac patients, irrespective of allocation to E+ or E– ($p < 0.05$). There was no interaction between group and encouragement, however, suggesting that the effect of encouragement was similar in the two groups. Age was the other main factor that had a significant effect on change in walking score over time: younger patients tended to improve their walking score more over the six walks than did older patients.

We examined the influence of test repetition on the distance walked by comparing the first two days' results with those of the last four visits. Performance improved in both two and six minute walks ($p < 0.0001$) and the effect of repetition was greater in encouraged than in non-encouraged patients in the two minute walk ($p < 0.02$), supporting the effect of encouragement on both tests (fig 1 and 2).

VARIABILITY

The within person standard deviations, the coefficients of variation, and the ratio of the within

Table 1 Variability of walking test scores in the groups given encouragement (e+) and not given encouragement (E–)

	Within person standard deviation (m)	Coefficient of variation*	Within person SD Encouragement effect
Two minute walk E+	8.217	0.054	—
E–	6.4	0.044	—
Total group	7.372	0.049	0.90
Six minute walk E+	22.83	0.052	—
E–	22.25	0.056	—
Total group	22.5581	0.053	0.74

*Coefficient of variation = within person standard deviation (SD) — mean absolute walking distance over the last four visits.

person standard deviations to the encouragement effects are presented for both two and six minute walks in table 1. Table 1 shows that variability in walking test score is approximately equal in the E+ and E– groups, and that the coefficient of variation is similar for the two and six minute walks. The coefficient of variation is most useful where the within person standard deviation increases in a linear fashion with the mean score; among our patients, however, there was no relation whatever between mean score and within person standard deviation. A more useful estimate of the utility of the two tests may be obtained by comparing the within person standard deviation with the treatment effect (in this case encouragement). The ratio of within person standard deviation to the treatment effect gives some idea of the responsiveness (that is, the ability to detect real change superimposed on random error) of the test: in this respect the six minute walk appears preferable (Table 1).

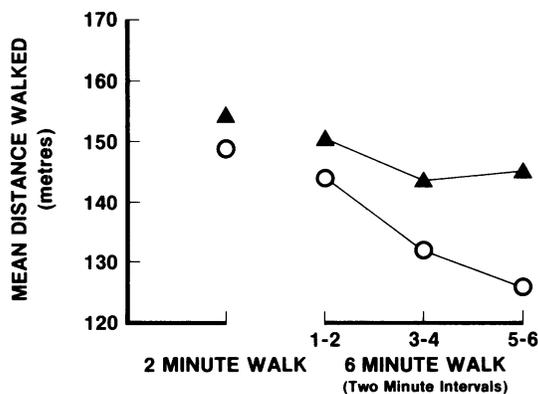


Fig 3 Mean distance walked in the two minute walking test and in consecutive two minute segments of the six minute walking test in encouraged (▲) and non-encouraged (○) groups of patients.

The effect of encouragement on walking test performance

An analysis of variance of the mean scores for the two minute walk and for each two minute interval of the six minute walk showed there was a highly significant difference between the distance walked during the four two minute intervals ($p < 0.0001$), and that there was a significant interaction between interval and encouragement ($p < 0.002$) (fig 3). Patients walked further during the two minute walk than during any two minute segment of the six minute walk, and walk scores decreased during the latter two intervals of the longer test. The effect of encouragement was greater during the second and third two minute intervals of the six minute walk than during the first; this suggests that the decrement in walk score which occurred after the first two minutes of the six minute walk was attenuated by supervisor encouragement.

Discussion

The purpose of this study was to examine different durations and modes of administration of walking tests, with a view to providing data which would guide the development of an optimal protocol for administering walking tests.

Encouragement by the supervisor had a substantial positive effect (mean 30.5 m) on patients' scores in the six minute walk ($p < 0.02$); in the two minute walk there was a similar though not significant trend (fig 1 and 2). The potential importance of this may be appreciated from a comparison of the present results with the differences in walking test scores previously attributed to pharmacological or physiological interventions. Since most previous studies have used the 12 minute walking test, to make such a comparison we must estimate what our treatment effect might have been had we doubled the duration of the walk. We are guided in making this estimate both by previous work, which has shown that after an initial burst in the first two minutes, the distance walked in each subsequent two minute segment of the 12 minute walk is remarkably constant,¹⁵ and by our finding (fig 3) that the encouragement effect was greater in the second and third two minute segments than in the first. These data suggest that a reasonable estimate of the treatment effect that would have been observed had the duration of the walk been 12 minutes is twice the treatment effect which we observed. Given this assumption, the encouragement effect we found is comparable to the treatment effect in eight of 10 "positive" studies which have used the walking test as a measure of outcome (table 2). Thus it is possible that differential encouragement may tend to bias walking test scores in favour of the actively treated group, irrespective of treatment efficacy. An impor-

Table 2 *Effect of various treatments on walking test performance in previous studies*

Reference	Treatment	Size of effect (m)
Present investigation McGavin <i>et al</i> ²	Encouragement	61*
	Exercise rehabilitation programme	83
Cockcroft <i>et al</i> ³	Exercise rehabilitation programme	77
Sinclair and Ingram ⁴	Exercise rehabilitation programme	187
Belman and Mittman ⁵	Ventilatory muscle training	130
Pardy <i>et al</i> ⁶	Ventilatory muscle training	74
Pardy <i>et al</i> ⁷	Ventilatory muscle training	84
Woodcock <i>et al</i> ⁸	Oxygen carried by assistant	70
	Oxygen carried by patient	48
	Predosing with oxygen	56
	Oxygen	53
Leggett and Flenley ⁹		
Leitch <i>et al</i> ¹³	Aminophylline	51
	Salbutamol	53
	Aminophylline and salbutamol	48
Woodcock <i>et al</i> ¹⁴	Promethazine	32

*Outcome of six minute walk, effect doubled to approximate that of 12 minute walk.

tant biasing effect of encouragement is very unlikely if the test supervisor is blind to subject allocation; the test supervisor was not, however, blind to allocation in seven of the 10 positive studies cited.^{2-7,9} These results suggest that in future trials using walking tests encouragement should be rigorously standardised, and that the test supervisor should be blind with respect to subjects' treatment group."

The fact that walking test score improved more over time in younger than in older patients was an unexpected finding. It suggests another possible bias which could lead to a spurious conclusion of treatment efficacy—if, for example, treatment groups were unbalanced with respect to age, with more younger people having the experimental intervention.

Important issues in administration of the walking test include the duration of the test, whether encouragement should be uniformly withheld or uniformly given, and the number of practice sessions which are necessary before an intervention can be tested. Butland and colleagues¹⁵ showed high correlations between results of the two, six, and 12 minute walks, suggested that 12 minutes is an unnecessarily long duration and that the two minute walk is less powerful in discriminating between subjects, and concluded that "the six, minute walk may represent a sensible compromise." As a measure of outcome in clinical trials, the responsiveness of the test—that is, its ability to detect clinically important change—is more important than the ability to discriminate between subjects. Responsiveness is directly proportional to the treatment effect of a

given intervention, and inversely proportional to the within person standard deviation. The ratio of within person standard deviation to treatment effect was lower for the six minute than for the two minute walk in the present study (table 1), suggesting the six minute walk is the more appropriate test for evaluating treatment effects.

One criterion for deciding on whether or not to give encouragement to clinical trial participants during performance of tests of exercise capacity is the variability in score over time. Within subject variation was similar during E+ and E- trials (table 1). So long as there is consistency, therefore, the decision on whether or not to provide encouragement would appear to be arbitrary.

Two previous studies which addressed walking test reproducibility suggested that after the initial two walks, in which a training effect is observed, scores obtained on subsequent repetitions do not vary systematically.^{15 16} As depicted in figures 2 and 3, the results of the present investigation are consistent with this conclusion. Thus when walking tests are being used to measure exercise performance in experimental studies two practice tests should be conducted before baseline measurements are obtained.

We have shown that the level of encouragement provided by the test supervisor to patients undergoing walking tests can have a large effect on subjects' performance. While six minutes appears to be an optimal duration for the test, there is no compelling reason to either give or not give encouragement, so long as tests are conducted in a consistent manner. Finally, we have confirmed the results of previous work which has suggested that two practice walks are necessary because of an initial training effect. The results of this study suggest the need for careful standardisation and blinding of test supervisors when exercise capacity is being measured in patients with chronic respiratory or chronic heart disease.

We would like to acknowledge the assistance of Drs MT Newhouse, JLC Morse, and FE Hargreave in recruiting patients for this study and the assistance of Ms Marie Townsend in data analysis. The work was supported in part by the St Joseph's Hospital Foundation, Hamilton, Ontario, and the Medical Research Council of Canada.

References

- ¹ McGavin CR, Gupta SP, McHardy GJR. Twelve-minute walking test for assessing disability in chronic bronchitis. *Br Med J* 1976;i:822-3.
- ² McGavin CR, Gupta SP, Lloyd EL, McHardy GJR. Physical rehabilitation for the chronic bronchitic: results of the controlled trial of exercises in the home. *Thorax* 1977;32:307-11.
- ³ Cockcroft AE, Saunders MJ, Berry G. Randomised controlled trial of rehabilitation in chronic respiratory disability. *Thorax* 1981;36:200-3.
- ⁴ Sinclair DJM, Ingram CG. Controlled trial of supervised exercise training in chronic bronchitis. *Br Med J* 1980;i:519-21.
- ⁵ Belman MJ, Mittman C. Ventilatory muscle training improves exercise capacity in chronic obstructive pulmonary disease. *Am Rev Respir Dis* 1980;121:273-80.
- ⁶ Pardy RL, Rivington RN, Despas PJ, Macklem PT. Inspiratory muscle training compared with physiotherapy in patients with chronic airflow limitation. *Am Rev Respir Dis* 1981;123:421-5.
- ⁷ Pardy RL, Rivington RN, Despas PJ, Macklem PT. The effects of inspiratory muscle training on exercise performance in chronic airflow limitation. *Am Rev Respir Dis* 1981;123:426-33.
- ⁸ Woodcock A, Gross ER, Geddes DM. Oxygen relieves breathlessness in "pink puffers." *Lancet* 1981;i:907-9.
- ⁹ Leggett RJE, Flenley DC. Portable oxygen and exercise tolerance in patients with chronic hypoxic cor pulmonale. *Br Med J* 1977;ii:84-6.
- ¹⁰ Butland RJA, Pang JA, Geddes DM. Carbimazole and exercise tolerance in chronic airflow obstruction. 1982;37:64-7.
- ¹¹ O'Reilly JF, Shaylor JM, Harrison BDW. Week-to-week reproducibility of the twelve-minute walking test and its use in assessing the effects of oral steroid treatment in chronic airways obstruction. *Thorax* 1980;35:718.
- ¹² Williams AJ, Osman J, Skinner C. Effects of naftidrofuryl on breathlessness and exercise tolerance in chronic bronchitis. *Thorax* 1982;37:617-9.
- ¹³ Leitch AG, Morgan A, Ellis DA, Bell G, Haslett C, McHardy GJR. Effect of oral salbutamol and slow-release aminophylline on exercise tolerance in chronic bronchitis. *Thorax* 1981;36:787-9.
- ¹⁴ Woodcock AA, Gross ER, Geddes DM. Drug treatment of breathlessness: contrasting effects of diazepam and promethazine in pink puffers. *Br Med J* 1981;283:343-6.
- ¹⁵ Butland RJA, Pang J, Gross ER, Woodcock AA, Geddes DM. Two-, Six-, and Twelve-Minute walking test in respiratory disease. *Br Med J* 1982;284:1007-8.
- ¹⁶ Mungall IPF, Hainsworth R. Assessment of respiratory function in patients with chronic obstructive airways disease. *Thorax* 1979;34:254-8.