Low-technology exercise test in the preoperative evaluation of lung resection candidates

A. Brunelli, C. Pompili, M. Salati


Exercise tests are increasingly used in the preoperative functional evaluation of lung resection candidates. Low-technology exercise tests include six minute walking, shuttle walking and stair climbing.

Conflicting evidence has been reported regarding 6 minute walking test. This test should not be used to select patients for lung resection.

An incremental shuttle walk test is easier to replicate than 6 minute walking test. Most patients achieving 25 shuttles or 400 m have a maximum oxygen consumption measured at cardiopulmonary exercise test greater than 15 l/Kg/min. Although this test tends to underestimate performance at the lower range compared to peak oxygen consumption it can be used a screening test before lung resection. Excluding patients from operation based on this test alone is however not recommended and a formal cardiopulmonary exercise test should be always used in those walking less than 400 m.

Stair climbing has been extensively studied in thoracic surgery. Several studies have found that poor performance in this test is indicative of cardiopulmonary complications and mortality after lung resection. In particular, climbing less than 22 m represents very high risk, whereas climbing more than 22 m is associated with a favourable outcome. Recent guidelines recommend referring all patients climbing lower than 22 m to cardiopulmonary exercise test. Stair climbing can be used as a screening test in cases cardiopulmonary exercise test is not readily available. In general, patients climbing more than 22 m can proceed to surgery without further evaluation. Monaldi Arch ChestDis 2010; 73: 2, 72-78.

Keywords: Exercise test, Pulmonary resection, Preoperative evaluation, Stair climbing test, Shuttle walk test, 6 minute walk test.

Division of Thoracic Surgery, Ospedali Riuniti Ancona, Italy.

Correspondence: Dr. Alessandro Brunelli, Div Thoracic Surgery, Ospedali Riuniti Ancona, Via Conca 1, 60020 Ancona, Italy; e-mail: brunellialex@gmail.com

Introduction

Exercise testing has been recognised as pivotal in the preoperative functional evaluation of lung resection candidates. Recent European clinical guidelines have recommended the systematic performance of exercise tests to all patients with FEV1 or DLCO below 80% of predicted value. Ideally, a formal cardiopulmonary exercise test with direct measurement of oxygen consumption (VO2 peak) during incremental exercise on a cycle or treadmill should be performed. This high-technology test permits a precise assessment of the cardiovascular or ventilatory systems, and to identify the nature of the deficit if present.

However, CPET is not always readily available. A recent web-based survey among thoracic surgeons, chest physicians and oncologists which was sponsored by ERS and ESTS showed that only 10 to 30% of patients have a CPET before lung resection despite the fact that this technology is available in 75% of hospitals. This discrepancy may be partly explained by logistic and organisational local problems that may limit the access to the exercise lab (often used by cardiologists or other specialists). It may also reflect uncertainties concerning the real indications of this cost-intensive and labor-intensive high-tech test [1].

As an alternative to high-tech exercise tests in settings where these are not readily available low-technology tests have been proposed at least as a screening test.

The same web-based survey [1] showed that 65% of respondents perform one of the possible low-tech tests in at least 50% of candidates to lung resections. Most of them apply exercise testing in patients with some impairment of pulmonary function, while 18% declared to use them systematically in all patients as part of their work up, and as much as 20% stated they use low-tech tests as a surrogate to VO2 peak measurement as the last step to decide operability (a practice not recommended by the most recent guidelines) [2].
This review will focus on the most commonly used low-technology exercise tests used in the pre-operative evaluation of lung resection candidates.

### 6-12 Minute Walking Test (6MWT)

As the ability to ambulate, in most of cases, indicates functional status, the assessment of distance ambulated in a specific time period has been reported as a possible method of evaluating global fitness. The walking test was originally developed in the early 1960s by Balke. The aim was to evaluate functional capacity by simply measuring the distance walked during a defined period of time [3]. A 12-minute field performance test was then developed by Cooper to evaluate the level of physical fitness of healthy individuals [4]. This Author found that the correlation of this field-test data with the laboratory-determined oxygen-consumption data was 0.897.

Subsequently, a 6-minute walk was found to be equally as effective as the 12-minute walk, and this test was preferred to evaluate patients with respiratory problems for whom a 12-min walking test could be too exhausting [5]. More recently, Solway and coll [6] in a review of functional walking tests concluded that “the 6MWT is easy to administer, better tolerated, and more reflective of activities of daily living than the other walk tests”.

Contrary to other low-tech forms of exercise (such as the stair climbing test), the 6MWT assesses a submaximal level of functional capacity. Most patients do not achieve maximal exercise capacity during the 6MWT; instead, they choose their own intensity of exercise and are allowed to stop and rest during the test. Although most activities of daily living are performed at submaximal levels of exertion and this test may reflect the level of exercise for daily physical activities, this characteristic may detract from its ability to predict complication after pulmonary surgery.

The 6MWT should be performed in accordance with the American Thoracic Society guidelines [7]. According to these guidelines, this test should be performed indoors, along a long, flat, straight, enclosed corridor with a hard surface. The walking course must be 30 m in length and the length of the corridor should be marked every 3 m. The turnaround points should be marked with a cone (such as an orange traffic cone). A starting line, which marks the beginning and end of each 60 m lap, should be marked on the floor using brightly coloured tape. The patient is instructed to walk as far as he or she can during a six minute timeframe. Patients are permitted to slow down, to stop, and even to rest as necessary.

Bagg and coll [8] were among the first to evaluate this test in our specialty. In a small series of only 22 patients they were not able to find any difference between complicated and non-complicated patients in terms of distance walked in 12 minutes both in absolute terms and when standardized for effort. They concluded that the walking test did not add to pulmonary function tests during preoperative evaluation.

These results were confirmed some years later by Markos and coll [9], who did not find the distance walked in 12 minutes predictive of postoperative cardiopulmonary complications or mortality in 55 candidates to lung resection.

Holden and coll [10] evaluated three different types of exercise tests to predict death or mechanical ventilation after lung resection in 16 high-risk respiratory patients. They found that the distance walked in 6 minutes was predictive of mortality within 90 days of operation and a distance longer than 1,000 feet was predictive of survival with a sensitivity of 100%, a negative predictive value of 100%, and a positive predictive value of 85%.

These results seem to partially conflict with those reported by Pierce and coll [11]. These Authors found that the 6-minute walking distance was indicative of respiratory failure, but not of other complications or mortality, in a series of 54 candidates to lung resection for lung cancer.

In a series of 73 candidates to pulmonary resection (unpublished data), we were not able to find any difference in distance walked in 6 minutes in terms of cardiopulmonary morbidity rate (non-complicated 459 m vs. complicated 435 m, \( p=0.5 \)). Forty of these patients had also VO\(_2\) peak measured during cycle-ergospirometry. The correlation coefficient between VO\(_2\) peak measured during cycling and distance walked during the 6MWT was poor (\( r=0.41 \)). Moreover, 14 of 27 patients (51%) who walked less than 500 m had a VO\(_2\) peak greater than 15 ml/Kg/min. Conversely, only 4 patients who walked more than 500 m had a VO\(_2\) peak lower than 15 ml/Kg/min. In our small series, the 6MWT tended to underestimate performance compared to CPET. Patients walking less than 500 m can still have acceptable exercise tolerance levels according to VO\(_2\) peak measurement. The use of this test alone as a screening test before surgery may therefore lead to inappropriate exclusion of potential surgical candidates.

Based on the inconsistent scientific evidence reported in the literature, a joint task-force sponsored by the European Respiratory Society and the European Society of Thoracic Surgeons concluded that 6MWT should not be used to select patients for operation [2].

### Shuttle walk test (“SWT”)

The incremental shuttle walk test was originally devised by Singh and coll [12] to assess the exercise capacity of patients with chronic obstructive disease. The protocol requires the patient to walk up and down a 10 m course. The course is identified by two cones. The speed at which the patient is asked to walk is dictated by an audio signal played by a tape. The patient objective is to turn around the cone by the time he or she hears the au-
dio signal. Each minute the speed of walking is increased by a small increment typically 0.17 m/s. The first speed of walking is referred to as level 1, the second as level 2, and so on. Each level lasts for one minute and the tape continues for 10 levels. The number of shuttles (10 m in length) in each level is dictated by the walking speed at that level. The test is symptom-limited and ends when the patient is physically exhausted or too breathless to continue, or because he fails to complete a shuttle in the permitted timeframe.

The SWT test appears more productive and more closely associated with VO₂ peak estimated by formal exercise testing than six-min walking test [12].

Indeed it has been estimated by regression analysis that 25 shuttles on shuttle walk test indicate a VO₂ peak of 10 ml/Kg/min [13] and therefore this cut-off has been suggested as a screening selection criteria for surgery.

More recently Win and coll [14] did not find any statistically significant difference in the shuttle walk distance between patients with and without complications after lung resection. They concluded that this test should not be used to predict the surgical outcome in lung cancer patients. The same group demonstrated that SWT distance significantly correlated with oxygen uptake (VO₂) peak (r=0.67) [15]. All 55 patients who walked more than 400 m during the SWT had a VO₂ peak measured at CPET (using treadmill) greater than 15 ml/Kg/min. Seventy patients failed to achieve 400 m. Yet only 22 of them had a VO₂ peak below 15 ml/Kg/min and nine of 17 patients who walked less than 250 m had a VO₂ peak greater than 15 ml/Kg/min. These findings seem to demonstrate that SWT tends to underestimate exercise capacity at the lower range compared with peak oxygen consumption measured with CPET. These authors suggested that patients achieving 400 m on the shuttle walk test do not require formal measurement of oxygen consumption.

In a recent paper, Benzo and colleagues [16] assessed the VO₂ SWT correlation by directly measuring the VO₂ during the shuttle test with a portable gas analyser in a group of 50 stable COPD patients. They found a high correlation between each level of the SWT and the VO₂ consumption. The cut-off of 25 shuttles had positive predictive value for predicting a VO₂ peak > 15 ml/Kg/min of 90%.

**Stair climbing**

The stair climbing test has been used by surgeons for decades. Van Nostrand and colleagues were the first to study this test for preoperative evaluation. They found that patients unable to climb two flights of stairs had a 50% mortality rate after pneumonectomy [17].

Stair climbing test can be performed as a submaximal or symptom-limited test. In the submaximal test an objective is set (i.e. 2 or 3 flights of stairs). If the patient achieves the objective he/she is deemed fit for surgery. Most of the studies on stair climbing test are however performed in a symptom-limited fashion. The patients are instructed to climb the maximum number of steps and stop only for exhaustion, limiting dyspnea, chest pain or leg fatigue. Patients are usually monitored throughout the test by means of a pulse oximeter and accompanied by a trained physician. Verbal encouragement and a continuous interaction to assess symptoms are used. The result of the test is variably expressed either in number of steps/flights or metres climbed. The expression in metres is probably to be favoured as it provides a form of standardisation when results are compared in different settings. In fact either the number of steps per flight or the height of the steps can vary from hospital to hospital. Usually the height of a step varies from 0.15 to 0.17 m.

Stair climbing tests are regarded as a constant work rate exercise test and this is at variance with the incremental shuttle walk test and other more sophisticated incremental protocols on treadmill and cycle. It is likely that the performance expressed during the stair climbing test reflects a mixed aerobic-anaerobic capacity rather than a pure aerobic capacity.

In some studies, the time to complete the test is recorded and used to estimate the VO₂ peak based on modified equations [18, 19].

Similar to all the other low-tech tests, stair climbing is an economical and widely applicable form of exercise. It is simple and rapid, requires few personnel, expertise, and equipment. Moreover, it has been reported that the stair climbing test yielded greater values of VO₂ peak than cycle ergometry [10, 20, 21]. This finding may be explained by the fact that stair climbing is a more stressful exercise than cycling, involving a larger muscle mass, and it is extremely motivating for the patients who are pushed to reach a visible objective represented by the next landing.

Pollock and coll [21] found that patients with chronic obstructive pulmonary disease were able to generate a VO₂ peak by stair climbing that was equivalent to that measured by cycle ergometry. Interestingly, in that study, a climb of one flight of stairs generated a VO₂ peak of 7.9 ml/Kg/min, whereas climbing 4.6 flights (83 steps) generated a VO₂ peak of 20 ml/Kg/min or higher.

Several studies have shown the capability of the stair climbing test to predict post-operative complications in lung resection candidates. Olsen and colleagues [18] were the first to hint at a possible role for this test in the preoperative workup. In their preliminary study of 54 patients undergoing thoracotomy they found that the ability to climb 3 flights of stairs (corresponding in their setting to 13 m) most clearly separated those patients with longer hospital stay, postoperative intubation, and greater frequency of complications. The small sample size and heterogeneity of surgical patients (from exploratory thoracotomy to pneumonectomy) precluded further sub-group analyses and the role of this test in predicting fatal cardiopulmonary morbidity.
As mentioned above, Holden and coll [10] studied 16 patients deemed at high respiratory risk and evaluated 3 different types of exercise tests. Eleven patients had no or minor complications and 5 patients died. Patients with minor complications had a longer distance walked (6MWT) and a higher altitude climbed on the stair climbing test compared to those patients who ultimately died. VO₂ peak measured at cycle-ergometry was not different (measured in only 10 patients). They concluded that a stair climb of greater of 44 steps (7.5 m in their setting) was predictive of a successful surgical outcome.

Pate and coll [22] were the first to express the performance of the stair climbing test in metres rather than steps in the attempt to provide a form of standardisation. They operated on 12 patients with borderline pulmonary function based on their performance assessed at CPET (VO₂ peak > 10 ml/Kg/min). Patients with complications climbed an average of 2 m less than those without complications. All patients but one (who eventually experienced respiratory insufficiency) were able to climb more than 3 flights of stairs, confirming the high correlation between this test and VO₂ peak measured at CPET.

Girish and coll [23] found in a group of thoracic and upper abdominal surgical patients that the inability to climb two flights of stairs was associated with a positive predictive value of 82% for the development of post-operative complications. The number of post-operative days in the hospital decreased proportional with a patient’s increased ability to climb stairs.

Subsequently, we reported our preliminary experience on 160 lung resection candidates [19]. We found that the altitude reached during pre-operative stair climbing test was the only independent predictor of cardiopulmonary complications after controlling the effect of other variables by logistic regression analysis. In particular, only 6.5% of subjects climbing more than 14 metres developed complications. However, 29% of subjects who climbed between 12 and 14 metres, and 50% of those climbing less than 12 metres, had complications. This progressive increase of cardiopulmonary morbidity rate with the reduction of the altitude climbed preoperatively, indirectly demonstrates that stair climbing is a stressful test capable of revealing severe deficits in maximum aerobic capacity. At the same time, we were able to prove that stair climbing was superior with respect to spirometric variables in discriminating between complicated and non-complicated patients. In fact, in a high-risk group of 17 patients with percent predicted FEV₁ less than 35% or percent predicted DLCO less than 35% or both, and who were judged to be operable on the basis of their satisfactory performance at the stair climbing test, we observed only four complications and no mortality. Based on these findings we concluded that this type of exercise should be considered as a routine screening test in all lung resection candidates.

We recently reported on a series including 640 major anatomic lung resections (lobectomy and pneumonectomy) [24]. The population included 41% of patients older than 70 years of age, 49% of...
patients with some form of cardiac co-morbidity and 24% of patients with moderate to severe COPD. This case-mix underlines the concept that virtually all patients can proceed to this test before operation. The average duration of the test was 111 seconds and the average altitude climbed was 19 m. We did not observe any significant complication related to the test. The altitude reached on the pre-operative stair climbing test was reliably associated with postoperative cardiopulmonary complications, mortality and hospital costs. Compared to patients climbing more than 22 m those unable to climb 12 m had an incidence of cardiopulmonary complications and mortality of 2.5- and 13-fold higher, respectively. The mortality rate in this low-performing group was as high as 13% (figure 1). When the analysis was restricted to those 73 patients with prohibitive respiratory reserve (% FEV1 < 40% or % DLCO < 40% or both), we found that the altitude reached at stair climbing tests retained its predictive value. In this high-risk group, we did not observe any death among those climbing more than 22 m, but we had a 20% mortality rate (2 of 10 patients) among those not reaching 12 m.

In a more recent study [25], we measured the oxygen consumption in 109 patients performing the stair climbing test before lung resection by using a portable gas analyser. We found a high correlation between altitude climbed and VO2 peak measured during the test (r = 0.7). Fifty-five percent of patients unable to reach 12 m had a VO2 peak below 15 ml/Kg/min, whereas 98% of patients climbing more than 22 m had a VO2 peak greater than 15 ml/Kg/min. The cut-off of 22 m had a positive predictive value of 86% to predict a VO2 peak of 15 ml/Kg/min.

These results corroborate recent recommendations from ERS-ESTS guidelines [2] stating that patients must be referred to formal CPET in case of a poor performance (< 22 m) at stair climbing test whenever this test is used as a first line functional screening test.

Another form of expressing the performance of the stair climbing test may be the speed of ascent. This parameter has been suggested as a possible way of standardising this test across different institutions and infrastructures [26]. In 44 patients with FEV1 < 80% of predicted, Koegelenberg and colleagues [26] found a linear correlation between speed of ascent and VO2 peak measured at cycle-ergometry. All 24 patients with a speed greater than 15 m/min had a VO2 peak > 20 ml/Kg/min. Further studies are needed to assess the association between speed of ascent and development of post-operative complications. However, if these findings are confirmed in future investigation and in different case-mixes, speed of ascent could be used as an additional parameter to assess physical fitness before operation.

Another interesting parameter which has been investigated in the literature is the exercise oxygen desaturation. Nevertheless, controversy still exists on the best definition (< 90% or greater than 4%), and its association with postoperative complications [27-30].

In a recent analysis on 536 patients performing a maximal stair climbing tests prior to their major lung resection, we [30] found that the value of 4% exercise oxygen desaturation (EOD) during stair climbing with respect to rest oxygen saturation was a reliable predictor associated with the occurrence of postoperative complications after logistic regression analysis. Patients with EOD had a mortality rate of 8% versus 3% in those without EOD. In particular, those patients climbing less than 12 m and developing EOD had a mortality rate as high as 29%. These findings confirm a previous investigation from Ninan and coll [28] who found that an exercise oxygen desaturation of 4% or greater in a standardised stair-step apparatus predicted longer intensive care unit stay and incidence of major cardiopulmonary morbidity. An older study from Rao and colleagues [27] found that exercise oximetry measured during a 150 m walk around the ward predicted better than spirometry home oxygen re-

### Table 1. - Characteristics of different types of low-technology exercise tests according to published reports

<table>
<thead>
<tr>
<th></th>
<th>6MWT</th>
<th>Shuttle walking test</th>
<th>Stair climbing test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Economic</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Rapid</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Reproducible</td>
<td>Yes</td>
<td>Yes</td>
<td>No- Structure dependent</td>
</tr>
<tr>
<td>Correlated with VO2 peak</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Predictive of surgical outcome</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Safe</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Easy to understand</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Familiar to patients</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Maximal effort</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
</tbody>
</table>
requirements, need of admission to the intensive care unit, prolonged hospital stay, and respiratory failure. Oximetry identified 50% of the patients who died.

Conclusions

Low-technology exercise tests are simple and reliable tests capable of detecting major problems in the cardiopulmonary system, which may occur to postoperatively. Although these tests, particularly stair climbing test and SWT, appear to be valuable instruments in assessing the level of performance they can not precisely identify the nature of the possible deficit in the oxygen transport system. In this regard, these tests will never entirely replace the more sophisticated cardiopulmonary exercise tests with direct measurement of expired gases. CPET remains the best test to precisely assess the extent and cause of a possible reduced aerobic reserve by the provision of direct and derived parameters. Corrective measures and treatments can be implemented accordingly to improve patients’ fitness and reduce surgical risk.

For this reason, low-tech tests can only be used as first line screening tests [2]. Only those patients with a very good performance on the stair climbing test (i.e. > 22 m) or, to a lesser extent the SWT (> 400 m) can proceed safely to surgery. All the others need to be referred to a formal CPET for a more detailed definition of the status of their exercise capacity and more precise surgical risk-stratification.

References

4. Cooper KH. A means of assessing maximal oxygen intake: correlation between field and treadmill testing. JAMA 1968; 203: 201-204.
26. Koegelenberg CF, Diacon AH, Irani S, Bolliger CT.


