Reproducibility of the 6-Minute Walk Test for Ambulatory Oxygen Prescription

Arjun B. Chatterjee a Richard W. Rissmiller c Kyle Meade d Connie Paladenech b
John Conforti a Norman E. Adair a Edward F. Haponik a Robert Chin, Jr. a

a Section of Pulmonary, Critical Care, Allergy and Immunologic Diseases, Wake Forest University School of Medicine, and d Department of Respiratory Care, Wake Forest University Baptist Medical Center, Winston-Salem, N.C., and c Department of Medicine, Carolinas Medical Center, and d Presbyterian Pulmonary and Critical Care, Charlotte, N.C., USA

Key Words
Ambulatory oxygen • Chronic obstructive pulmonary disease • Oxygen inhalational therapy • Walk test

Abstract
Background: Ambulatory oxygen is frequently prescribed for patients with chronic obstructive pulmonary disease (COPD) who have oxygen desaturation \( \leq 88\% \) during exercise. The 6-min walk test (6MWT) with continuous pulse oximetry monitoring is a common method to document this oxygen desaturation, but the reproducibility of this test in determining the need for ambulatory oxygen in patients with COPD is not well documented. Objective: The aim of this study was to establish the reproducibility of the 6MWT in determining the need for ambulatory oxygen prescription in stable COPD patients using the Centers for Medicare and Medicaid (CMS) criteria for ambulatory oxygen prescription. Methods: The study was designed as a prospective observational study in an academic health center and associated pulmonary rehabilitation program. Eighty-eight COPD patients referred to pulmonary rehabilitation underwent continuous pulse oximetry while performing standard 6MWT on 3 separate days. Results: Fifty-one (58\%) of these patients desaturated by continuous pulse oximetry to an \( \text{SpO}_2 \leq 88\% \) on at least one of the 6MWTs. Only 26 patients (30\%) demonstrated consistency in meeting the criteria for ambulatory oxygen set forth by the CMS on all three 6MWT with a \( \kappa \) statistic of 0.62. The percent agreement between 6MWTs for ambulatory oxygen prescription was 72\% and the paired observation was 51\%. Conclusions: The 6MWT distance is simple and widely used as a consistent measure of functional capacity in patients with COPD; however, the 6MWT oxygen saturation has only modest reproducibility in determining the need for ambulatory oxygen in stable COPD patients undergoing pulmonary rehabilitation.

Introduction

Since the demonstration of a survival benefit of oxygen administration in patients with arterial hypoxemia and chronic obstructive pulmonary disease (COPD) was shown in the 1980s [1, 2], it is estimated that over 600,000 patients with COPD receive this therapy [3]. Home oxygen therapy is the leading durable medical equipment expenditure for the Centers for Medicare and Medicaid Services (CMS), with estimated costs exceeding one billion dollars annually [3–5]. However, it is often prescribed inappropriately (not meeting set criteria in up to 40.5\% of patients) and with excessive flow rates based on suggest-
ed criteria [6]. In 1998, long-term oxygen therapy represented 28% of the total Medicare expenditure for durable medical equipment [4]. To be reimbursed for home oxygen therapy for respiratory-impaired patients, CMS has strict criteria that are also accepted by the Department of Veteran Affairs and most commercial health plans [4]. The criteria include exercise-induced oxygen desaturation to ≤88% or a reduction in the PaO2 to ≤55 mm Hg (7.3 kPa) [7].

The 6-min walk test (6MWT) is a standardized test commonly used to assess the need for ambulatory oxygen prescription in patients with cardiopulmonary disorders [8]. In COPD patients, the 6MWT has been used to evaluate responses to medical interventions, as a one-time measurement of functional status, as a predictor of morbidity and mortality and has been incorporated in a multidimensional tool to predict outcome in patients with COPD [8–10]. Despite the reproducibility of the distance walked during the 6MWT, the reproducibility of oxygen desaturation as an indication for oxygen prescription during the 6MWT in patients with COPD has not been evaluated rigorously. The financial implication of the use of 6MWT for ambulatory oxygen is substantial. We sought to determine the reproducibility of oxygen desaturation during the 6MWT for ambulatory oxygen prescription in patients with COPD.

**Patients and Methods**

We enrolled patients with stable COPD actively participating in a Pulmonary Rehabilitation Program at our institution. The study protocol was approved by the Institutional Review Board of the Wake Forest University Baptist Medical Center. Patients were acclimated to and performed 6MWTs on 3 different days during attendance at pulmonary rehabilitation sessions. Clinical stability of the patients was assessed subjectively with the level of the patient’s respiratory symptoms (cough and sputum production, level of fatigue, modified Borg score and relative well-being) and objectively with vital signs, SpO2 monitoring, a subject-reported general sense of well-being (7-point visual analog scale), and a respiratory system physical examination by a respiratory therapist before the 6MWT using the American Association of Cardiovascular and Pulmonary Rehabilitation guidelines [11]. The test was deferred if the patient was not felt to be at his/her baseline. Relative contraindications to performance of the 6MWT included: resting heart rate >120 beats/minute, systolic blood pressure >180 mm Hg, and diastolic blood pressure >100 mm Hg.

6MWTs were performed following American Thoracic Society (ATS) guidelines except that a rectangular level track with a distance of 61 meters per lap was used [8]. Patients were instructed to walk laps at their own pace, attempting to cover as much ground as possible in the allotted 6 min accompanied by a trailing respiratory therapist. SpO2, blood pressure, and pulse rate were assessed at the start and completion of each test. Pulse rate and continuous oxygen saturation were recorded at each minute interval. Walks were initiated on room air, and if SpO2 values fell consistently below 88% for a minimum of ≥5 s with a reliable and true pulse oximetry signal as judged by the respiratory therapist, the test was interrupted and O2 was administered via nasal cannula and titrated to maintain an SpO2 >90%.

SpO2 was measured and the results printed using a Nellcor N-20P pulse oximeter. The Nellcor N-20P samples data points that are averaged over 5–7 s. Printouts are done every 30 s representing the average of the 5–7 s closest to the 30-second interval. In the 70–100% oxygen saturation range on this device, the published standard deviation is ±2%. We minimized SpO2 artifact by assuring a good signal at baseline, ensuring proper waveform during the 6MWT, requiring a persistent drop ≤88% for a minimum of 5 s and a record for documentation (average of ~30 s). There are no published data establishing the criterion standard for the duration of oxygen saturation ≤88% nor does CMS require a duration time for ambulatory oxygen qualification. Oxygen desaturation was defined as a 4% drop in oxygen saturation. Qualification for ambulatory oxygen was defined as a nadir oxygen saturation ≤88%. Baseline demographics were gathered from the information available through the Pulmonary Rehabilitation Program at the Wake Forest University Baptist Medical Center. All patients enrolled had the diagnosis of COPD and determination of either FEV1 or FEV1 % of predicted (FEV1 %); however, lung volumes and DLCO were not available for each patient.

**Statistical Analysis**

The χ statistic was computed using the methods described by Fisher and van Belle [12] with reference to Fleiss [13] and Reynolds [14]. The χ statistic is a measure of the reproducibility of a test; χ ≥0.7 is considered a good test while a poorly functioning test has χ <0.3. The χ statistic is used to compare two different measurements to assess reproducibility of the measurements, particularly when no gold standard is available. Since we used a series of three tests, we constructed a three-dimensional array to represent all of the variations in the data. Using the techniques described, we modeled the three-dimensional array to generate expected observation rates for use in the χ analysis. The χ statistic was computed for our primary measure of SpO2 ≤88% and for desaturation ≥4%.

The percent agreement measures the percent of the observations in which positive and negative measurements agree, and the paired observation measures the percent of the observations in which positive tests agree. To evaluate if there was any difference in the populations when separated by number of positive tests, we categorized the population based on the number of tests in which they desaturated by 4% or became hypoxemic to a level <88%. This allowed the creation of ‘desaturation categories’ with 1, 2, or 3 positive tests and the creation of ‘hypoxemia categories’ with 1, 2, or 3 positive tests. We compared age, gender, FEV1%, FEV1/FVC ratio, and total lung capacity % of predicted between all groups using the two-sample, 2-tailed Student’s t test with unequal variance.

Statistical calculations for χ were calculated using Excel (Excel 2008, Microsoft, Redmond, Wash., USA). All other statistical calculations were done using JMP7 (SAS Institute, Cary, N.C., USA; version 7).
Results

Ninety-seven pulmonary rehabilitation participants were screened and 88 participants with COPD were enrolled in the study with three 6MWTs performed. The median age was 72 years, with 37 males, a median FEV₁ of 1.21, a mean FEV₁% of 52%, and a median FEV₁/FVC ratio of 52%. DLCO was available for 34 patients with a mean of 11.1 ml/min/mm Hg. DLCO % of predicted (DLCO%) was available for 37 patients with a mean of 55%. The baseline SpO₂ was 96% (SD 1.98).

All three 6MWT were completed with a median of 11.5 days (interquartile range: 13.5). In subcategory comparisons between patients who never desaturated and those who desaturated on all three 6MWTs and the various combinations thereof, there was no difference in age, FEV₁, FEV₁%, FEV₁/FVC ratio, DLCO, DLCO%, baseline SpO₂, and days between walks. The characteristics of the patients are shown in table 1.

Fifty patients (58%) desaturated to ≤88% in at least one of the three 6MWTs. Twelve (13%) met the CMS criteria in one of three walks, 13 (15%) met the criteria on two of three walks and only 26 (30%) met the criteria on all three 6MWTs. Seventy-six patients (86%) desaturated by 4% in at least one test, with 55 (63%) of these patients showing this level of desaturation on all three walks (table 1). The percent agreement between tests for CMS qualification for ambulatory oxygen prescription was 72%, the paired-observation percent agreement was 51% and the κ statistic was 0.62 (table 2).

Although the patients were instructed to walk at their own natural pace during each of their 6MWT, we compared the time to desaturate by 4% as a surrogate marker for effort in those that desaturated to ≤88%, and there was no significant difference in the time to desaturate by 4% in each of the 6MWT including those that desaturated to ≤88%. As an additional metric of stability over the period of the measurements, the distance walked on serial 6MWT was not statistically significantly different, nor was the overall personal sense of health (symptom score) between serial evaluations.

Discussion

This study demonstrates that the 6MWT is not necessarily robust in determining the need for ambulatory oxygen based on CMS criteria. The modest κ value as well as the low agreement and paired-observation test between the three 6MWT to determine the need for ambulatory oxygen using the current CMS criteria suggest that the reproducibility of 6MWT may not be consistent enough for ambulatory oxygen therapy prescription in

Table 1. Characteristics of the patients enrolled in the study

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>n (%)</th>
<th>Mean ± SD</th>
<th>Median</th>
<th>IQR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patients</td>
<td>88</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male gender</td>
<td>37 (42)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>88</td>
<td>72</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td>FEV₁, liters</td>
<td>75</td>
<td>1.21</td>
<td>0.55</td>
<td></td>
</tr>
<tr>
<td>FEV₁, % of predicted</td>
<td>73</td>
<td>52 ± 19.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FEV₁/FVC</td>
<td>60</td>
<td>0.52</td>
<td>0.26</td>
<td></td>
</tr>
<tr>
<td>DLCO, ml/min/mm Hg</td>
<td>34</td>
<td>11.1 ± 3.43</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DLCO, % of predicted</td>
<td>37</td>
<td>55 ± 14.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline SpO₂, %</td>
<td>88</td>
<td>96 ± 1.98</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Days between walks</td>
<td>88</td>
<td>11.5</td>
<td>13.5</td>
<td></td>
</tr>
<tr>
<td>Symptom rating</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Walk 1</td>
<td>88</td>
<td>3</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Walk 2</td>
<td>88</td>
<td>3</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Walk 3</td>
<td>88</td>
<td>3</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>6MWT distance, m</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Walk 1</td>
<td>88</td>
<td>326 ± 105</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Walk 2</td>
<td>88</td>
<td>358 ± 109</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Walk 3</td>
<td>88</td>
<td>362 ± 117</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

≤88%
| Any 6MWT                         | 51    |           |        |     |
| One 6MWT                         | 12    |           |        |     |
| Two 6MWTs                        | 13    |           |        |     |
| Three 6MWTs                      | 26    |           |        |     |

>4%;
| Any 6MWT                         | 76    |           |        |     |
| One 6MWT                         | 5     |           |        |     |
| Two 6MWTs                        | 16    |           |        |     |
| Three 6MWTs                      | 55    |           |        |     |

Interquartile range (IQR) is the 25th percentile subtracted from the 75th percentile. Median and IQR are provided for data that were not normally distributed. Means ± SD are provided for normally distributed data. Symptom rating is a subjective self-measure of a subject’s overall health on a 7-point visual analog scale.

Table 2. Agreement characteristics of the 6MWT results for ≥4% desaturation and for pulse oximetry ≤88% in a 6MWT

<table>
<thead>
<tr>
<th></th>
<th>κ statistic</th>
<th>Agreement</th>
<th>Paired observation</th>
</tr>
</thead>
<tbody>
<tr>
<td>≥4%</td>
<td>0.52</td>
<td>76%</td>
<td>73%</td>
</tr>
<tr>
<td>≤88%</td>
<td>0.62</td>
<td>72%</td>
<td>51%</td>
</tr>
</tbody>
</table>
patients with COPD. This study was not designed to determine the cost/benefit ratio of ambulatory oxygen but to determine the limitations of a commonly used test that is used to prescribe ambulatory oxygen therapy. If consistent reproducible desaturation to $\leq 88\%$ during exercise were required for ambulatory oxygen prescription, then two thirds of our patients would not have qualified for ambulatory oxygen. Conversely, if just one 6MWT desaturation to $\leq 88\%$ were required to meet criteria, then 40% of the patients who might benefit from ambulatory oxygen would not qualify.

Oxygen desaturation during activity can be documented by a variety of tests using $\text{SpO}_2$ including simple walking, stair climbing, 6MWT, the shuttle-walk test, cardiac stress testing or cardiopulmonary exercise testing. The 6MWT is a convenient, safe, simple test widely used to monitor functional capacity. The timed walks are remarkably consistent in terms of the distance walked [15–17]. However, exercise-induced oxygen desaturation during a 6MWT does not correlate well with the distance walked [15]. In two different studies of the effect of supplemental oxygen on exercise training in COPD patients who were not hypoxemic at rest, oxygen supplementation resulted in improved dyspnea scores, health-related quality of life and peak exercise tolerance. However, there was no difference noted in mean $\text{SpO}_2$ or in 6MWT distance either before or after exercise [18, 19].

Poulain et al. [20] compared 6MWT to cardiopulmonary exercise testing using a cycle ergometer to detect oxygen desaturation in 80 clinically stable COPD patients. Twenty-three patients (29%) had oxygen desaturation, defined as a decrease in oxygen desaturation $>4\%$ during the last 3 min of exercise in the 6MWT alone. Sixteen patients (20%) desaturated in both tests and 41 (51%) did not desaturate in either. To assess reproducibility, two 6MWTs were done 15 min apart. In 10 patients who desaturated in both initial 6MWT, a repeat 6MWT was performed in $6\pm1$ days and appeared to confirm oxygen desaturation. The authors concluded that oxygen desaturation during a 6MWT exercise was reproducible, not protocol dependent and could be used to prescribe supplemental oxygen [20]. By contrast, we examined oxygen desaturation to a level of $\leq 88\%$, the endpoint CMS uses to reimburse for home oxygen, and found that when 6MWTs were done on different days, this criterion is not highly reproducible in stable COPD patients. Although 86% of our COPD patients experienced $>4\%$ desaturation during exercise, only 30% showed consistent reductions to levels satisfying CMS criteria with only a modest $\kappa$ statistic for test quality.

In a study assessing nocturnal desaturation in COPD patients over 3 nights, Lewis et al. [21] found that significant nocturnal desaturation (defined as the time spent with oxygen saturation $<90\%$ on home overnight pulse oximetry $\geq 30\%$ of the night) changed in 9 of 26 study patients (35%) between nights 1 and 2. They concluded that a 1-night study to detect significant nocturnal desaturation might be inadequate in COPD patients [21]. Eaton et al. [22] found that the 6MWT had excellent within-subject reproducibility for distance in patients with fibrotic idiopathic interstitial pneumonia, but the degree of oxygen desaturation at the end of exercise had poor reproducibility. Similarly, our study suggests only a modest reproducibility of desaturation to CMS criteria in repeated 6MWTs in patients with COPD for ambulatory oxygen prescription despite significant financial implications [5, 7, 23–25].

Long-term oxygen therapy for COPD patients with $\text{PaO}_2 <55$ mm Hg (7.5 kPa) at rest prolongs survival, reduces the frequency of hospitalization, decreases development and progression of pulmonary hypertension, and improves activities of daily living and quality of life [1, 2, 26]. The long-term benefits of ambulatory oxygen use are less evident, yet supplemental oxygen is often prescribed to patients with stable COPD who desaturate with exercise. Two Cochrane Reviews, one addressing short-term ambulatory oxygen and one addressing long-term ambulatory oxygen therapy for patients with COPD, concluded that short-term ambulatory oxygen benefited patients with COPD by increasing exercise endurance (distance, time, and number of steps) and maximal exercise capacity (exercise time and work rate), but there was lack of evidence to support long-term ambulatory oxygen therapy for patients with COPD [27, 28]. Bradley et al. [29] performed a meta-analysis of randomized controlled trials examining the short-term benefit of ambulatory oxygen in patients with COPD. They concluded that although ambulatory oxygen improved exercise performance, the clinical significance of the degree of improvement is not clear. Oxygen therapy during exercise does appear to allow for exercise training in hypoxemic and non-hypoxemic patients with COPD by improving endurance and the breathing pattern during high-intensity training in pulmonary rehabilitation programs [18]. It has not been determined whether patients with COPD who desaturate intermittently during exercise have outcomes different from patients who desaturate consistently. However, a recent report from the National Emphysema Treatment Trial noted an increased mortality in those patients with a higher prevalence of exercise de-
saturation who were non-hypoxemic at rest but using continuous oxygen therapy than in non-hypoxemic patients not using oxygen from the data gathered in that trial, but it was unclear whether continuous oxygen was of benefit to the former group [30]. An ongoing workshop sponsored by CMS and the National Heart, Lung and Blood Institute examining the benefits of long-term oxygen therapy in patients with COPD may help answer its role in less hypoxemic patients [5].

Possible limitations to this study are the study population, the lack of arterial blood gas sampling, the variable length of time between the 6MWTs, the rectangular shape of our 6MWT track, and the unvalidated symptom score used in our pulmonary rehabilitation program. Although this study may not be applicable to patients with COPD outside a pulmonary rehabilitation center, it does suggest that the oxygen desaturation in a 6MWT is not robust in this more controlled group. A study in an unselected cohort of patients with COPD should be done to see if the lack of reproducibility would be duplicated in this group.

Arterial blood gas analysis is not required by CMS to document exercise desaturation and the use of SpO₂ monitoring for this purpose represents common clinical practice. CMS does not require a specific duration of oxygen desaturation to limit artifactual transient oxygen desaturation, and we were much more rigorous in our definition and documentation of oxygen desaturation than required by CMS. However, the accuracy of pulse oximetry during exercise has been questioned. Escourrou et al. [31] studied 101 patients with various diagnoses (33 had COPD) during exercise. Three different pulse oximeters were used, and patients exercised on a bicycle. With exercise, PaO₂ measurements and SpO₂ readings correlated well (r = 0.88). In another study, 8 patients with severe COPD underwent exercise testing on a bicycle with both blood gas analysis and pulse oximetry. SpO₂ was moderately accurate compared to arterial blood oxygen saturation [32]. We cannot exclude that a more technologically advanced oximeter or the accuracy (or inaccuracy) of the pulse oximetric readings may have played a role in the suboptimal reproducibility of the 6MWT in our study; however, we minimized this potential error using experienced respiratory therapists to differentiate true from spurious oximetric values and requiring a persistence of the drop in pulse oximetry readings. Our definition of a significant drop in SpO₂ to ≤88% with exercise is more stringent than the current CMS criteria.

While we designed the study and strove to minimize the days between the tests in clinically stable patients already enrolled in pulmonary rehabilitation, completion of all three 6MWT took a mean of 21 days with a standard deviation of 31 days and an interquartile range of 14 days. We compared our results with an analysis of the subgroup that completed all three 6MWTs within 14 days and found no difference in baseline characteristics, symptom ratings, 6MWT distance or frequency of desaturation. The κ value increased to a marginally acceptable value of 0.71. While κ did improve, the borderline value is still concerning due to the significant financial impact and unclear clinical benefit of ambulatory oxygen prescription in COPD patients with exercise desaturation.

In patients with severe COPD (FEV₁ <50%), frequently repeated pulmonary rehabilitation resulted in improved 6MWT distance compared to annual rehabilitation; however, there was no discussion of SpO₂ changes [33]. Our patients were enrolled in continuing pulmonary rehabilitation, and their stability is reflected by the reproducible 6MWT distance, times to desaturation, and symptom scores. This study was designed to assess the reproducibility of a 6MWT for exercise oxygen desaturation; therefore, we cannot comment on the efficacy of ambulatory oxygen in COPD patients with exercise desaturation.

The track used for the 6MWT was rectangular. The 2002 ATS Guidelines for the 6MWT recommend a straight rather than continuous track based on a paper presented at the Annual Conference of the American College of Chest Physicians in 2000 [34]. In the only published study addressing this issue from 2003, there was a 10% increase in the 6MWT distance walked on continuous compared to straight tracks in the National Emphysema Treatment Trial. This study did not report data on SpO₂ [8, 34].

The seven-point visual analog scale that our pulmonary rehabilitation program uses to assess a sense of general well-being in our patients is not validated. The data are collected to add to clinical evaluation of stability. Although the symptom score had a median of 3 for each 6MWT, it only adds a small value to the sense of stability. The consistent 6MWT distance walked adds relatively more evidence that the patients were stable between 6MWTs.

**Conclusions**

Our preliminary study suggests that oximetric desaturation during 6MWTs used for the evaluation of the need for ambulatory oxygen in stable COPD patients partici-
pating in pulmonary rehabilitation may not be consistently reproducible with only a 51% agreement based on CMS criteria between three 6MWTs. Although there are predictors (DLCO and resting SpO₂) that may suggest exercise-induced oxygen desaturation in obstructive lung disease, these were not associated with desaturation in all three tests [35, 36]. The inconsistency of the 6MWT for exercise-induced SpO₂ desaturation has important implications for the treatment of individual COPD patients, the overall costs of care, and utilization of health care resources. Performance of multiple 6MWTs in patients with stable COPD, if the clinical suspicion is high, may be necessary to accurately identify patients at risk for consistent oxygen desaturation with exercise. In light of the recent evidence that exercise-induced oxygen desaturation during a 6MWT in patients with COPD may have prognostic value, our study lays the groundwork for further studies into this question [30, 37].

References


