Measurement of Symptoms, Lung Hyperinflation, and Endurance during Exercise in Chronic Obstructive Pulmonary Disease

DENIS E. O’DONNELL, MIU LAM, and KATHERINE A. WEBB
Respiratory Investigation Unit, Departments of Medicine, and Community Health and Epidemiology, Queen’s University, Kingston, Ontario, Canada

Changes in lung hyperinflation, dyspnea, and exercise endurance are important outcomes in assessing therapeutic responses in chronic obstructive pulmonary disease (COPD). Therefore, we studied the reproducibility of Borg dyspnea ratings, inspiratory capacity (IC; to monitor lung hyperinflation), and endurance time during constant-load symptom-limited cycle exercise in 29 patients with COPD (FEV$_1$ = 40 ± 2% predicted; mean ± SEM). Responsiveness was also studied by determining the acute effects of nebulized 500 µg ipratropium bromide (IB) or saline placebo (P) on these measurements. During each of four visits conducted over an 8-wk period, spirometry and exercise testing were performed before and 1 h after receiving IB or P (randomized, double-blinded). Highly reproducible measurements included: endurance time (intraclass correlation R = 0.77, p < 0.0001); Borg ratings and IC at rest, at a standardized exercise time (STD), and at peak exercise (R > 0.6, p < 0.0001); and slopes of Borg ratings over time, oxygen consumption ($\dot{V}O_2$), and ventilation (R > 0.6, p < 0.0001). Responsiveness was confirmed by finding a significant drug effect for: change (∆) in endurance time (p = 0.0001); ∆Borg$_{STD}$ and ∆Borg-time slopes (p < 0.05); and ∆IC at rest, at STD, and at peak exercise (p = 0.0001). With all completed visits, ∆Borg$_{STD}$ correlated better with ∆IC$_{STD}$ than any other resting or exercise parameter (n = 115, r = -0.35, p < 0.001). We concluded that Borg dyspnea ratings, and measurements of IC and endurance time during submaximal cycle exercise testing are highly reproducible and responsive to change in severe COPD. O’Donnell DE, Lam M, Webb KA. Measurement of symptoms, lung hyperinflation, and endurance during exercise in chronic obstructive pulmonary disease.

For patients with advanced chronic obstructive pulmonary disease (COPD), alleviation of symptoms such as dyspnea and leg discomfort, and improvement in activity level, become important therapeutic goals. It follows that, in clinical studies designed to evaluate the efficacy of therapeutic interventions in this population, reliable measurements of symptom intensity and of relevant physiological parameters should ideally be employed. The Borg category scale is being used increasingly to measure the intensity of dyspnea and leg discomfort during cardiopulmonary exercise testing in patients with pulmonary disease. However, the reliability (reproducibility) and responsiveness (ability to detect change) of this scale remain to be determined with precision in patients with advanced COPD. No definitive conclusions can be made from the existing literature because of the very small sample sizes in previous studies and because of interstudy differences in the sensations being rated (i.e., respiratory effort versus dyspnea), in the mode of exercise employed (i.e., treadmill versus cycle), and in the testing protocol (i.e., incremental versus submaximal constant-load) (1–4). There is also considerable variation in the time intervals over which reproducibility of this measurement has been assessed (1–4). Finally, we could find no specific information in the literature about the reproducibility of Borg ratings of exertional leg discomfort. Leg discomfort is the primary symptom that limits exercise in many patients with COPD (5) and is, therefore, an important outcome measure to examine when assessing the effects of interventions such as exercise training.

Both incremental testing and constant-load endurance testing have been used to assess therapeutic responses in patients with COPD. Both approaches have the potential to produce different, but complementary, information about changes in exercise performance following interventions. Increasingly, investigators are using endurance testing at a standardized submaximal work load to evaluate the impact of interventions such as exercise training, oxygen therapy, bronchodilator medication, or volume reduction surgery. The question arises, therefore, whether endurance times using symptom-limited constant-load exercise protocols are sufficiently reliable and sensitive to detect clinically important changes in COPD.

A variety of cardiopulmonary responses are traditionally measured during exercise testing for the assessment of the

(Received in original form April 2, 1998 and in revised form June 5, 1998)
Supported by Boehringer Ingelheim (Canada) Limited.
Dr. O’Donnell holds a career scientist award from the Ontario Ministry of Health.
Correspondence and requests for reprints should be addressed to Denis O’Donnell, Kingston General Hospital, Richardson House, 102 Stuart Street, Kingston, ON, K7L 2V7 Canada. E-mail: codonnell@post.queensu.ca
Internet address: www.atsjournals.org
dyspneic COPD patient. With many therapeutic interventions, dyspnea relief is multifactorial and includes alterations in ventilatory demand (6–8), in ventilatory mechanics (9–11), or in both. Repeated measurements of inspiratory capacity (IC) during exercise have been used by several investigators (12, 13), including ourselves (14), to track the behavior of dynamic end-expiratory lung volume (EELV) during exercise in COPD. There is increasing evidence that the IC during exercise, and the rate of change in IC during exercise (i.e., dynamic hyperinflation), are strong predictors of exertional dyspnea intensity and exercise intolerance (14). Thus, the availability of this simple method of tracking EELV during exercise could have wide clinical application, provided it is proved reliable.

To address these questions, we set the following study objectives. First, we wished to determine the reproducibility of Borg measurements of dyspnea and leg discomfort during constant-load cycle exercise tests performed over an 8-wk period in patients with stable advanced COPD. Second, we wished to evaluate the responsiveness of Borg dyspnea ratings by comparing the acute effects of a bronchodilator (nebulized ipratropium bromide [IB]) using a randomized, double-blind, two-period crossover design (Figure 1). The crossover design ensures equality in numbers of patients assigned each treatment, and allows each subject to serve as his or her own control.

After hospital/university research ethics approval was obtained, subjects gave informed consent and entered the study on a staggered basis. During a screening visit, subjects underwent thorough familiarization with all procedures and symptom rating scales, medical history, pulmonary function testing with reversibility to $\beta_2$-agonists, dyspnea evaluation, and symptom-limited incremental cycle exercise testing. Thirty-six patients were then randomized to receive either nebulized IB (500 $\mu$g) or P three times a day in unit dose vials for a 3-wk period. Each subject was crossed over to the alternate treatment after a 2- to 7-d washout period. Four experimental visits per subject were conducted at the beginning and end of each 3-wk period (Figure 1). In these visits, pulmonary function testing, dyspnea evaluation, and constant-load cycle exercise testing were performed immediately before and 1 h after the treatment corresponding to that being given during the coinciding 3-wk period.

Concomitant respiratory medications permitted for the duration of the study included regularly taken inhaled steroids and bronchodilators and a long history of cigarette smoking; (2) moderate to severe chronic breathlessness (modified Baseline Dyspnea Index [BDI] < 6) (15); (3) age 55 yr or older; and (4) clinically stable as defined by no changes in medication dosage or frequency of administration with no exacerbations or hospital admissions in the preceding 6 wk. Exclusions included: a history of asthma, atopy, or nasal polyps; other active lung disease or other significant disease that could contribute to dyspnea or exercise limitation; or oxygen desaturation to less than 80% during exercise on room air.

Study Design
Reliability (or reproducibility) was established by analyzing baseline measurements taken at repeated visits, i.e., a replication reliability study. Responsiveness to treatment was assessed by comparing the acute effects of two treatments (single-dose IB versus placebo [P]) using a randomized, double-blind, two-period crossover design (Figure 1). The crossover design ensures equality in numbers of patients assigned each treatment, and allows each subject to serve as his or her own control.

After hospital/university research ethics approval was obtained, subjects gave informed consent and entered the study on a staggered basis. During an initial screening visit, subjects underwent thorough familiarization with all procedures and symptom rating scales, medical history, pulmonary function testing with reversibility to $\beta_2$-agonists, dyspnea evaluation, and symptom-limited incremental cycle exercise testing. Thirty-six patients were then randomized to receive either nebulized IB (500 $\mu$g) or P three times a day in unit dose vials for a 3-wk period. Each subject was crossed over to the alternate treatment for an additional 3-wk period, after a 2- to 7-d washout period. Four experimental visits per subject were conducted at the beginning and end of each 3-wk period (Figure 1). In these visits, pulmonary function testing, dyspnea evaluation, and constant-load cycle exercise testing were performed immediately before and 1 h after the treatment corresponding to that being given during the coinciding 3-wk period.

Concomitant respiratory medications permitted for the duration of the study included regularly taken inhaled steroids and bronchodilators and a long history of cigarette smoking; (2) moderate to severe chronic breathlessness (modified Baseline Dyspnea Index [BDI] < 6) (15); (3) age 55 yr or older; and (4) clinically stable as defined by no changes in medication dosage or frequency of administration with no exacerbations or hospital admissions in the preceding 6 wk. Exclusions included: a history of asthma, atopy, or nasal polyps; other active lung disease or other significant disease that could contribute to dyspnea or exercise limitation; or oxygen desaturation to less than 80% during exercise on room air.

Study Design
Reliability (or reproducibility) was established by analyzing baseline measurements taken at repeated visits, i.e., a replication reliability study. Responsiveness to treatment was assessed by comparing the acute effects of two treatments (single-dose IB versus placebo [P]) using a randomized, double-blind, two-period crossover design (Figure 1). The crossover design ensures equality in numbers of patients assigned each treatment, and allows each subject to serve as his or her own control.

A fter hospital/university research ethics approval was obtained, subjects gave informed consent and entered the study on a staggered basis. During an initial screening visit, subjects underwent thorough familiarization with all procedures and symptom rating scales, medical history, pulmonary function testing with reversibility to $\beta_2$-agonists, dyspnea evaluation, and symptom-limited incremental cycle exercise testing. Thirty-six patients were then randomized to receive either nebulized IB (500 $\mu$g) or P three times a day in unit dose vials for a 3-wk period. Each subject was crossed over to the alternate treatment for an additional 3-wk period, after a 2- to 7-d washout period. Four experimental visits per subject were conducted at the beginning and end of each 3-wk period (Figure 1). In these visits, pulmonary function testing, dyspnea evaluation, and constant-load cycle exercise testing were performed immediately before and 1 h after the treatment corresponding to that being given during the coinciding 3-wk period.

Concomitant respiratory medications permitted for the duration of the study included regularly taken inhaled steroids and bronchodilators and a long history of cigarette smoking; (2) moderate to severe chronic breathlessness (modified Baseline Dyspnea Index [BDI] < 6) (15); (3) age 55 yr or older; and (4) clinically stable as defined by no changes in medication dosage or frequency of administration with no exacerbations or hospital admissions in the preceding 6 wk. Exclusions included: a history of asthma, atopy, or nasal polyps; other active lung disease or other significant disease that could contribute to dyspnea or exercise limitation; or oxygen desaturation to less than 80% during exercise on room air.

Study Design
Reliability (or reproducibility) was established by analyzing baseline measurements taken at repeated visits, i.e., a replication reliability study. Responsiveness to treatment was assessed by comparing the acute effects of two treatments (single-dose IB versus placebo [P]) using a randomized, double-blind, two-period crossover design (Figure 1). The crossover design ensures equality in numbers of patients assigned each treatment, and allows each subject to serve as his or her own control.

After hospital/university research ethics approval was obtained, subjects gave informed consent and entered the study on a staggered basis. During an initial screening visit, subjects underwent thorough familiarization with all procedures and symptom rating scales, medical history, pulmonary function testing with reversibility to $\beta_2$-agonists, dyspnea evaluation, and symptom-limited incremental cycle exercise testing. Thirty-six patients were then randomized to receive either nebulized IB (500 $\mu$g) or P three times a day in unit dose vials for a 3-wk period. Each subject was crossed over to the alternate treatment for an additional 3-wk period, after a 2- to 7-d washout period. Four experimental visits per subject were conducted at the beginning and end of each 3-wk period (Figure 1). In these visits, pulmonary function testing, dyspnea evaluation, and constant-load cycle exercise testing were performed immediately before and 1 h after the treatment corresponding to that being given during the coinciding 3-wk period.

Concomitant respiratory medications permitted for the duration of the study included regularly taken inhaled steroids and bronchodilators and a long history of cigarette smoking; (2) moderate to severe chronic breathlessness (modified Baseline Dyspnea Index [BDI] < 6) (15); (3) age 55 yr or older; and (4) clinically stable as defined by no changes in medication dosage or frequency of administration with no exacerbations or hospital admissions in the preceding 6 wk. Exclusions included: a history of asthma, atopy, or nasal polyps; other active lung disease or other significant disease that could contribute to dyspnea or exercise limitation; or oxygen desaturation to less than 80% during exercise on room air.

Study Design
Reliability (or reproducibility) was established by analyzing baseline measurements taken at repeated visits, i.e., a replication reliability study. Responsiveness to treatment was assessed by comparing the acute effects of two treatments (single-dose IB versus placebo [P]) using a randomized, double-blind, two-period crossover design (Figure 1). The crossover design ensures equality in numbers of patients assigned each treatment, and allows each subject to serve as his or her own control.

After hospital/university research ethics approval was obtained, subjects gave informed consent and entered the study on a staggered basis. During an initial screening visit, subjects underwent thorough familiarization with all procedures and symptom rating scales, medical history, pulmonary function testing with reversibility to $\beta_2$-agonists, dyspnea evaluation, and symptom-limited incremental cycle exercise testing. Thirty-six patients were then randomized to receive either nebulized IB (500 $\mu$g) or P three times a day in unit dose vials for a 3-wk period. Each subject was crossed over to the alternate treatment for an additional 3-wk period, after a 2- to 7-d washout period. Four experimental visits per subject were conducted at the beginning and end of each 3-wk period (Figure 1). In these visits, pulmonary function testing, dyspnea evaluation, and constant-load cycle exercise testing were performed immediately before and 1 h after the treatment corresponding to that being given during the coinciding 3-wk period.

Concomitant respiratory medications permitted for the duration of the study included regularly taken inhaled steroids and bronchodilators and a long history of cigarette smoking; (2) moderate to severe chronic breathlessness (modified Baseline Dyspnea Index [BDI] < 6) (15); (3) age 55 yr or older; and (4) clinically stable as defined by no changes in medication dosage or frequency of administration with no exacerbations or hospital admissions in the preceding 6 wk. Exclusions included: a history of asthma, atopy, or nasal polyps; other active lung disease or other significant disease that could contribute to dyspnea or exercise limitation; or oxygen desaturation to less than 80% during exercise on room air.

Figure 1. The study design illustrates the sequence and timing of visits. The timing of events and measurements is also shown for experimental visits.
lators other than anticholinergic medications. Use of medications remained stable throughout the study treatment periods. Patients requiring additional medication or changes in medication for greater than 2 d were withdrawn from the study. Before each visit, patients were asked to withdraw from inhaled β₂-agonists, anticholinergics, and theophyllines for at least 4 h, 12 h, and 24 h prior to testing, respectively. Subjects avoided caffeine and heavy meals at least 4 h prior to testing, and avoided alcohol and major physical exertion entirely on the day of each visit. All visits were conducted at the same time of day for each subject.

**Treatment**

Treatments to be compared in this trial were administered by inhalation via face mask from a H u d s o n Up-Draft Nebulizer U nit (M od e l No. 1712; H u d s o n Respiratory Care Inc., Temecula, C A ) at a flow of 8 L/min over a period of approximately 15 min. The total volume of nebulized solution was 5 ml for all patients; IB 300 μg was compared with a placebo which consisted of sterile 0.9% sodium chloride solution (EELV (24), changes in IC reflect changes in end-expiratory lung volume measurements of IC at rest, every 3 min during exercise, and at peak exercise). The study was originally designed to look at both acute and chronic responses to bronchodilator therapy in severe COPD. It was estimated that a sample size of at least 28 subjects was required for each treatment arm of the study: calculations were based on the change in chronic dyspnea (TDI) as the primary endpoint, our own laboratory values for a similar COPD population, α = 0.05, β = 0.20, and a two-tailed test of statistical significance. The focus of this part of the study, however, was to look at the acute responses to IB or placebo; specifically, by Borg symptom ratings, IC measurements, and exercise endurance time.

Results are reported as means ± SEM. The conventional level of statistical significance of 0.05 was used for all analyses. Baseline demographics, lung function, and exercise capacity measurements were compared between the two treatment sequence groups using unpaired t tests. Exercise responses were studied using linear regression analysis of data sets from each individual, with slopes and intercepts expressed as means of these regression lines. Standardized exercise time (STD) was equal to the time of the highest equivalent amount of work or time completed in all experimental exercise tests for each subject, i.e., the time of the shortest exercise test rounded down to the nearest minute.

The intraclass correlation coefficient of reliability (R) (25) was employed to evaluate the reliability of using: (1) the Borg scale for the quantification of exertional dyspnea intensity; (2) the Borg scale for the quantification of exertional leg discomfort; (3) IC measurements for the evaluation of operational lung volumes at rest and during exercise; and (4) symptom-limited peak time for the assessment of submaximal exercise endurance. Various forms of measuring exertional symptom intensity and IC were considered in the study, i.e., slopes and intercepts, values at STD, and peak values. A 5% increase toward unity, error constitutes a smaller portion of the observed measurement; R > 0.4 indicates poor reliability, 0.4 < R < 0.75 fair to good reliability, and R > 0.75 excellent reliability (25). The coefficient of variation (CV), which describes the variance involved in a measurement relative to the mean of the measurement, was also calculated within-subjects for the above measurements.

A analysis of variance models incorporating the repeated measures crossover design of the study (26) were applied to assess the responsiveness of exercise endurance, dyspnea, and IC measurements to bronchodilator treatment. Possible carryover and period effects were first considered and assessed in the models. To examine the strength of the relationship between acute changes in exertional dyspnea and lung hyperinflation in response to therapy, Pearson’s correlations were performed with data from all completed tests. In this analysis, change (Δ) in BorgT10 was the dependent variable and independent variables included changes in concurrent exercise measurements of lung hyperinflation (IC, IRV), ventilation (Ve, Ve/VC), ratio of minute ventilation to oxygen consumption (Ve/O2), ratio of minute ventilation to carbon dioxide production (Ve/CO2), breathing pattern (F, VT, Vt/IC), gas exchange (VCO2/VO2, SaO2), cardiovascular function (heart rate, blood pressure), as well as changes in resting pulmonary function (FEV1, FVC, IC, MIP). Changes in endurance time were similarly analyzed.

**RESULTS**

**Subjects**

Of the 36 randomized patients, there were 29 evaluable patients (Table 1). Seven patients were prematurely withdrawn...
from the study because of hospital admission with respiratory tract infection (n = 2), hospital admission with congestive heart failure (n = 1), repeated attacks of respiratory panic (n = 1), increased breathlessness requiring a change in treatment for greater than 2 d (n = 1), facial skin irritation from the nebulizer (n = 1), and noncompliance due to lack of time (n = 1). One of the 29 evaluable patients experienced an acute exacerbation of COPD requiring additional medication 7 d before the last visit and did not complete the exercise testing in this visit.

Statistical Considerations

The two treatment sequence groups were comparable at study entry for demographics, lung function, and exercise capacity: n = 13 were in the treatment sequence that received IB in the first two experimental visits and placebo in the last two visits, n = 16 received treatments in the reverse order. Stability of baseline spirometry was verified before making inferences from significance tests on treatment effects: predose measurements of FEV₁ and FVC were highly repeatable (Table 2), ensuring that the level of airflow limitation was constant for the duration of the study. Before establishing the significance of direct treatment effects, various carryover and period effects were also ruled out.

With the change in the slope of Borg dyspnea ratings over time as the primary endpoint, a significant drug effect was observed in this study (Table 3). The power for detecting this response was 92%: calculated with values obtained in response to IB for the 29 subjects, α = 0.05, and a two-tailed test of significance.

**Measurements of Exercise Performance**

Exercise performance for constant-load testing was assessed by looking at symptom-limited peak endurance time, peak V̇O₂ and peak V̇E; each of these measurements showed excellent reproducibility with an R > 0.75 (Table 2). The responsiveness of a measurement was confirmed by the occurrence of a significant drug effect in response to therapy (Table 4); of the above measurements, the only one showing a drug effect in response to treatment was the endurance time (p = 0.0001). Therefore, the outcome measure looking at exercise performance that was both reliable and responsive to change was the exercise endurance time.

**Measurements of Symptom Intensity**

Repeated measurements of dyspnea at rest, at STD, and at peak exercise were shown to be highly reproducible (Figure 2, Table 2), as were Borg dyspnea ratings expressed as slopes over exercise time, V̇O₂, and V̇E (Table 3) and their respective y-intercepts. However, significant drug effects were only

**Table 1**

<table>
<thead>
<tr>
<th>SUBJECT CHARACTERISTICS*</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Male:Female</strong></td>
</tr>
<tr>
<td><strong>Age, yr</strong></td>
</tr>
<tr>
<td><strong>Height, cm</strong></td>
</tr>
<tr>
<td><strong>Weight, kg</strong></td>
</tr>
<tr>
<td><strong>Body mass index, kg/m²</strong></td>
</tr>
<tr>
<td><strong>Smoking history, pack-yr</strong></td>
</tr>
<tr>
<td><strong>Modified Baseline Dyspnea Index</strong></td>
</tr>
<tr>
<td><strong>Reversibility to β₂-agonists, ΔFEV₁ in L (%)</strong></td>
</tr>
</tbody>
</table>

**Pulmonary function, % predicted**

| FEV₁, L       | 1.05 ± 0.07 (40 ± 2) |
| FVC, L        | 2.24 ± 0.14 (59 ± 3) |
| FEV₁/FVC, %   | 47 ± 2 (69 ± 3) |
| MIP at FRC, cm H₂O | 53 ± 7 (66 ± 11) |
| MIP at RV, cm H₂O | 60 ± 9 (65 ± 9) |
| MEP, cm H₂O   | 91 ± 6 (50 ± 3) |

**Symptom-limited maximal exercise, % predicted**

| Work rate, watts | 54 ± 4 (39 ± 3) |
| Heart rate, beats/min | 107 ± 3 (64 ± 2) |
| V̇O₂, L/min       | 1.09 ± 0.08 (61 ± 4) |
| V̇E, L/min        | 34.9 ± 2.5 (72 ± 4) |
| %S₂O₂            | 93 ± 1 |

* Values are means ± SEM.
found for changes in Borg at STD (p < 0.01) (Table 3) and Borg-time slopes (p < 0.05) (Figure 3, Table 4). As with Borg dyspnea ratings, Borg-derived measurements of exertional leg discomfort were also highly reproducible (Figure 2, Tables 2 and 3).

**Measurements of Operational Lung Volumes**

Measurements of IC at rest showed excellent reproducibility (R > 0.75), as did resting spirometric measurements of FEV₁ and FVC (Table 2). Measurements of IC at STD and at peak exercise were also highly reproducible (Figure 2, Table 2). IRV measurements were also reproducible at rest, at STD, and at peak exercise (Table 2). For the assessment of the extent of dynamic hyperinflation (DH) during exercise, slopes of IC over time, VO₂, or VE showed very poor reliability (R < 0.4). However, when DH was measured more simply as the change in IC from rest, fair to good reliability was found for changes occurring at STD and at peak exercise (Table 2).

Significant drug effects were found for changes in resting FEV₁ and FVC, and for changes in measurements of both IC and IRV evaluated at rest, at STD, and at peak exercise (Table 4). Slopes of IC relative to time (Figure 3), VO₂ or VE were not responsive to drug therapy in this study.

**Correlates of Improvement**

For all completed tests (n = 115), the change in exertional dyspnea intensity at a standardized time correlated better with the concurrent change in IC, expressed in absolute terms (r = −0.34, p < 0.0005) or as a percentage of predicted normal (r = −0.35, p < 0.0005), than with any other resting or exercise parameter. Other significant correlates of ΔBorgSTD included resting changes in FEV₁ (r = −0.26, p = 0.006) and IC (r = −0.25, p = 0.007), and concurrent changes in exercise breathing pattern: ΔF (r = 0.31, p = 0.001) and ΔVE/IC (r = 0.34, p < 0.0005).

Improvement in exercise endurance time correlated best with resting spirometric changes (p < 0.0005): ΔFVC (r = 0.43), ΔFEV₁ (r = 0.40), change in peak expiratory flow rate

**TABLE 3**

<table>
<thead>
<tr>
<th>Group</th>
<th>Mean</th>
<th>Within-Subject SD</th>
<th>Within-Subject CV (%)</th>
<th>Intraclasse R</th>
<th>Drug Effect p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dyspnea-time, Borg · min⁻¹</td>
<td>0.66</td>
<td>0.24</td>
<td>36.9</td>
<td>0.832</td>
<td>0.019</td>
</tr>
<tr>
<td>Dyspnea-Vo₂, Borg · (L/min)⁻¹</td>
<td>6.62</td>
<td>3.69</td>
<td>55.7</td>
<td>0.713</td>
<td>0.109 NS</td>
</tr>
<tr>
<td>Dyspnea-Ve, Borg · (L/min)⁻¹</td>
<td>0.24</td>
<td>0.15</td>
<td>60.3</td>
<td>0.615</td>
<td>0.158 NS</td>
</tr>
<tr>
<td>Leg-time, Borg · min⁻¹</td>
<td>0.78</td>
<td>0.41</td>
<td>52.7</td>
<td>0.641</td>
<td>0.001</td>
</tr>
<tr>
<td>Leg-Vo₂, Borg · (L/min)⁻¹</td>
<td>7.88</td>
<td>4.01</td>
<td>50.9</td>
<td>0.731</td>
<td>0.108 NS</td>
</tr>
<tr>
<td>Ve-time, L/min · min⁻¹</td>
<td>2.71</td>
<td>1.10</td>
<td>40.3</td>
<td>0.463</td>
<td>0.029 NS</td>
</tr>
<tr>
<td>Vo₂-time, L/min · min⁻¹</td>
<td>0.10</td>
<td>0.04</td>
<td>41.2</td>
<td>0.394</td>
<td>0.346 NS</td>
</tr>
<tr>
<td>Vo₂, L/min · (L/min)⁻¹</td>
<td>28.0</td>
<td>4.66</td>
<td>16.7</td>
<td>0.471</td>
<td>0.224 NS</td>
</tr>
</tbody>
</table>

* Note: all y- intercepts were highly reproducible and did not change with treatment.

**TABLE 4**

<table>
<thead>
<tr>
<th>Ipratropium Bromide (500 μg)</th>
<th>Preperiod</th>
<th>Postperiod</th>
<th>Drug Effect p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rest FEV₁, L</td>
<td>0.24 ± 0.04†</td>
<td>0.17 ± 0.04†</td>
<td>0.0001</td>
</tr>
<tr>
<td>FVC, L</td>
<td>0.47 ± 0.07†</td>
<td>0.32 ± 0.07†</td>
<td>0.0001</td>
</tr>
<tr>
<td>IC, L</td>
<td>0.45 ± 0.10†</td>
<td>0.33 ± 0.10†</td>
<td>0.0001</td>
</tr>
<tr>
<td>IRV, L</td>
<td>0.43 ± 0.08†</td>
<td>0.33 ± 0.09†</td>
<td>0.0001</td>
</tr>
<tr>
<td>Dyspnea, Borg</td>
<td>–0.3 ± 0.1†</td>
<td>–0.3 ± 0.1†</td>
<td>0.0001</td>
</tr>
<tr>
<td>Isotonic exercise Dyspnea, Borg</td>
<td>–0.6 ± 0.2†</td>
<td>–0.3 ± 0.2</td>
<td>0.0001</td>
</tr>
<tr>
<td>Leg discomfort, Borg</td>
<td>–0.3 ± 0.2</td>
<td>0.1 ± 0.2</td>
<td>0.03 ± 0.3</td>
</tr>
<tr>
<td>IC, L</td>
<td>0.39 ± 0.07†</td>
<td>0.29 ± 0.11</td>
<td>0.0001</td>
</tr>
<tr>
<td>IRV, L</td>
<td>0.32 ± 0.06†</td>
<td>0.23 ± 0.10</td>
<td>0.0001</td>
</tr>
<tr>
<td>Vo₂, L/min</td>
<td>–0.02 ± 0.02</td>
<td>–0.01 ± 0.02</td>
<td>0.0001</td>
</tr>
<tr>
<td>Ve, L/min</td>
<td>–0.3 ± 0.7</td>
<td>0.6 ± 0.7</td>
<td>0.01 ± 0.03</td>
</tr>
<tr>
<td>F, breaths/min</td>
<td>–2.2 ± 0.6†</td>
<td>–0.7 ± 0.5</td>
<td>0.02 ± 0.5</td>
</tr>
<tr>
<td>Peak exercise Exercise time, min</td>
<td>2.8 ± 0.9†</td>
<td>1.1 ± 0.4†</td>
<td>0.0001</td>
</tr>
<tr>
<td>Dyspnea, Borg</td>
<td>–0.4 ± 0.3</td>
<td>–0.1 ± 0.2</td>
<td>0.01 ± 0.2</td>
</tr>
<tr>
<td>Leg discomfort, Borg</td>
<td>0.1 ± 0.2</td>
<td>0.4 ± 0.2†</td>
<td>0.02 ± 0.2</td>
</tr>
<tr>
<td>IC, L</td>
<td>0.41 ± 0.08†</td>
<td>0.30 ± 0.11†</td>
<td>–0.08 ± 0.07</td>
</tr>
<tr>
<td>IRV, L</td>
<td>0.28 ± 0.07†</td>
<td>0.24 ± 0.11</td>
<td>–0.08 ± 0.05</td>
</tr>
<tr>
<td>Ve, L/min</td>
<td>0.01 ± 0.02</td>
<td>0.00 ± 0.02</td>
<td>0.04 ± 0.03</td>
</tr>
</tbody>
</table>

* Values are means ± SEM.
† p < 0.01, significant change from baseline.
‡ p < 0.05.
The increase in exercise endurance also correlated significantly with the reduction in the slope of exertional Borg dyspnea ratings over time ($r = -0.25, p = 0.007$) and in standardized Borg dyspnea ratings ($r = -0.32, p < 0.0005$).

**DISCUSSION**

The novel findings of our study are as follows. Borg ratings of dyspnea and leg discomfort were highly reproducible when measured during constant-load cycle exercise in patients with stable advanced COPD, at least over an 8-wk time period. The Borg dyspnea ratings measured with this protocol were also highly responsive. Exercise time was a reliable measure of exercise endurance in COPD, being both reproducible and responsive to change. Serial IC measurements during exercise were reproducible and responsive, and changes in IC after bronchodilator therapy correlated well with changes in exertional dyspnea.

**Exertional Symptom Intensity**

Despite increasing evidence to the contrary, there is prevailing skepticism about the reliability of simple category scaling to quantify symptom intensity and response to therapy. Lack of reproducibility of such subjective measurements could arise for a variety of reasons including: (1) random error; (2) inherent lability of the measured characteristic; (3) deficiencies of the measuring instrument; and (4) misunderstanding on the part of subjects or interviewers with respect to the precise question being asked. To evaluate reproducibility of Borg measurements of exertional symptoms, we calculated $R$ (25) in a large study sample of patients with stable COPD. This correlation expresses the relative magnitude of the main components of the variance of a measurement (i.e., variability among the steady-state values and variability of random errors) such that most of the untoward effects of unreliability are expressible as functions of it (25). We found good to excellent reliability ($R > 0.6$) for Borg ratings of dyspnea and leg discomfort, indicating that error constituted only a small portion of the observed measurement on repeated testing.

A prerequisite for the determination of reproducibility of symptom measurement is clinical stability. In our study, stability was evident from the high level of reproducibility of FEV$_1$ and other spirometric measurements over the 8-wk observation period. Borg ratings at the symptom-limited peak of exercise were highly reproducible in our patients, as others have found for peak Borg ratings following incremental cycle or treadmill exercise in COPD (1, 2, 27). However, peak Borg ratings were poorly responsive and, therefore, not appropriate in evaluating symptom responses to therapy, because such ratings before and after therapeutic interventions are often similar despite improvements in exercise endurance.

For this reason, comparison of submaximal Borg ratings at a standardized time, work rate, or metabolic load, or comparison of the slopes and intercepts of Borg ratings expressed as a function of time, VO$_2$, work rate, or ventilation, may be preferable in evaluating responses to therapy. However, in previous studies, submaximal dyspnea ratings during both constant-load treadmill testing and incremental cycle ergometry showed poor reproducibility in COPD (27). Belman and coworkers (3) demonstrated a reduction in Borg ratings on sequential testing during submaximal treadmill exercise and postulated...
desensitization to dyspnea as an explanation. However, we found that submaximal Borg values showed good reproducibility ($R = 0.58$) when tested repeatedly at a constant work load of approximately 50 to 60% of their predetermined maximum (Figure 2). The measurement of a continuum of subjective ratings as a function of independent physiologic stimulus (i.e., $V_{O_2}$, $V_E$) provides added information in assessing dyspneic patients. Mador and associates (27), however, showed large intrasubject variation in the slopes of Borg-time, Borg-work rate, and Borg-$V_{O_2}$, while employing an incremental cycle exercise protocol in COPD. With the constant-load protocol, we found that Borg-time and Borg-$V_{O_2}$ slopes and their intercepts showed excellent reproducibility; Borg-$V_E$ slopes were somewhat less reproducible (Table 3). Borg-time slopes also demonstrated greater responsiveness to acute bronchodilator therapy than Borg expressed as a function of $V_{O_2}$ or $V_E$ (Table 3). Our results, therefore, support the finding of Mahler and coworkers (28) who also found Borg-time slopes to be more reproducible and responsive than Borg-$V_{O_2}$ or Borg-$V_E$ slopes in asthmatic patients during incremental cycle exercise testing. We conclude that the measurement of Borg-time slopes is the most reliable method of evaluating symptomatic responses in patients with stable COPD during endurance exercise testing.

Several recent studies have shown that perceived leg discomfort, either exclusively or partially, limits exercise performance in COPD. Whereas dyspnea is more likely to be the primary limiting symptom in more advanced COPD (5, 8, 29), leg discomfort has been shown to be the predominant limiting symptom in mild to moderate disease (5). A n assessment of the intensity of exertional leg discomfort and its contribution to exercise limitation is, therefore, an integral component of the evaluation of disabled COPD patients who are, for example, entering pulmonary rehabilitation programs. Our study showed that the slopes of Borg ratings of leg discomfort over time and $V_{O_2}$ are both highly reproducible ($R > 0.6$). Moreover, in two previous controlled studies in patients with severe COPD, we have shown that Borg-time slopes for exertional leg discomfort were also responsive to change when using constant-load protocols similar to the one used here (6, 8). Thus, Borg-time slopes of leg discomfort fell significantly after exercise training, but not after a control period, in COPD (8). Similarly, slopes of exertional leg discomfort over time fell significantly while breathing added oxygen (60%), but not room air, in patients with advanced COPD (6).

Exercise Performance

There is no consensus with respect to which measurement of exercise performance should be used in the clinical evaluation of symptomatic patients with advanced COPD. Endurance tests, such as the 6-min walk or constant-load treadmill or cycle testing, have widespread use, and there is anecdotal information that such tests may be more responsive than incremental tests. Our study shows that endurance times on sequential cycle tests over a moderate time interval are highly reproducible ($R = 0.78$) and sufficiently sensitive to detect a clinically relevant increase (i.e., $32 \pm 9\%$ improvement, $p < 0.001$) after

![Figure 3. Measurements of dyspnea intensity, intensity of leg discomfort, inspiratory capacity, and ventilation are shown over time during constant-load exercise before (dashed lines) and after (solid lines) either nebulized 500 $\mu$g ipratropium bromide (IB) or saline placebo (P). Borg-time slopes fell significantly after IB ($p < 0.05$) with no change in y-intercepts; inspiratory capacity was increased at any given time after IB ($p < 0.001$), although slopes over time were unchanged. Values shown were calculated at the first visit of each 3-wk period (mean ± SEM; n = 29).]
bronchodilator therapy compared with placebo. It is noteworthy that, while \( V_o2 \) and \( V_e \) at the breakpoint of symptom-limited submaximal exercise were also highly reproducible, each of these measurements were insensitive to the effects of bronchodilator therapy.

Evaluation of Respiratory Mechanics

During cardiopulmonary exercise testing, measurement of the ventilatory response is deemed to be particularly relevant with respect to the evaluation of dyspneic patients and the assessment of therapeutic responses. We have previously shown, for example, that after exercise training (8) or oxygen therapy (6) in patients with severe COPD, relief of exertional dyspnea was explained primarily by reduction in the slopes of ventilation over time. Dyspnea relief or improved exercise tolerance in response to other interventions such as bronchodilator therapy (12) or volume reduction surgery (9, 30), however, may occur in the absence of a change in ventilation. In these circumstances, improvements in ventilatory mechanics, specifically reduced dynamic lung hyperinflation, are likely to be important contributing factors. Serial IC measurements throughout exercise allow noninvasive assessment of ventilatory mechanics (without the need for esophageal balloon placement) and permit us to track the behavior of operational lung volumes. To the extent that TLC does not change appreciably during exercise (24), changes in IC accurately reflect changes in dynamic EELV. In similar patients with COPD, we and others (31, 32) have previously confirmed that peak values of inspiratory esophageal pressure (an estimate of effort) did not change during repeated IC measurements, thereby validating the use of IC measurements throughout exercise.

Inspiratory capacity is determined by the degree of hyperinflation, inspiratory muscle strength, and the extent of intrinsic mechanical loading of these muscles. The IC also provides information about the position of tidal volume on the respiratory system’s pressure–volume (P–V) curve: the smaller the IC, the closer \( V_r \) to TLC and the upper alinear extreme of the P–V relation. IC also represents the operating limits for volume expansion during exercise: the greater the \( V_r/IC \) ratio, the greater the mechanic constraints on volume expansion and the greater the intensity of dyspnea (14, 31).

A nother reason for rigorous evaluation of IC measurements is the recent emergence of commercial software packages for clinical cardiopulmonary exercise testing that measure tidal flow–volume loops throughout exercise. Such programs use IC measurements to place tidal flow–volume loops on the volume axis within their maximal envelopes. However, there is little or no information about the reproducibility and responsiveness of repeated IC measurements during exercise. Our results showed that IC measurements at rest, at a standardized level of submaximal exercise, and at peak exercise were all highly reproducible (Figure 2) and responsive to change (Figure 3); as were other IC–derived measurements such as IRV or EILV. Although changes in IC from rest to a standardized level of exercise were acceptably reproducible, they were not responsive to change, and exercise slopes of IC were neither reproducible nor responsive. This is not surprising, because changes in IC throughout exercise are variable and often alinear. Moreover, in response to bronchodilator therapy, the slope of IC over time may remain unaltered and merely be shifted downward in parallel (Figure 3). Therefore, independent measurements of IC, rather than slope determinations, should be used preferentially to assess the effects of an intervention on dynamic hyperinflation. It is noteworthy that changes in IC measured during exercise correlated more strongly with changes in instantaneous Borg dyspnea ratings than any other physiologic parameter measured at rest or during exercise.

After acute bronchodilator therapy, change in dyspnea occurred in the absence of any change in ventilation, indicating the importance of changes in ventilatory mechanics. Bronchodilators acutely reduced the level of expiratory flow limitation, and thus hyperinflation, for any given ventilation. For these patients with advanced COPD and minimal demonstrable reversibility to \( \beta_2 \)-agonists, the reduction in dyspnea intensity after high-dose anticholinergic therapy also correlated with the change in resting IC and FEV1 measurements. However, in 17% of the sample, IC changed in the presence of minimal or no change in resting FEV1. Similarly, a study by Belman and associates (12) has shown that dyspnea relief following albuterol therapy in COPD was explained by significant changes in operational lung volumes and not by changes in the resting FEV1. In our studies, an increase in IC or reduction in dynamic hyperinflation of approximately 0.3 to 0.4 L during exercise translated into a clinically significant improvement in exercise endurance and reduction in exertional dyspnea. It should be emphasized that while the dose of nebulized IB (i.e., 500 \( \mu \)g) used in this study represented a standardized single dosage for patients with advanced COPD, it likely exceeds the dosage of this drug that is generally prescribed using a metered dose inhaler (MDI) and spacer device. In the absence of dose–response data for anticholinergics comparing the two modes of delivery, it is not possible to determine whether similar therapeutic responses could have been achieved in this population using lower doses of ipratropium delivered by MDI.

In conclusion, our study confirms the availability of reliable evaluative instruments to quantify symptom intensity, exercise endurance, and the level of dynamic hyperinflation in patients with advanced COPD. Accurate assessment of these important clinical outcomes is integral to the effective management of symptomatic patients with advanced pulmonary disease.

References


