Reliability and validity of the test of incremental respiratory endurance measures of inspiratory muscle performance in COPD

Magno F Formiga1,2
Kathryn E Roach1
Isabel Vital3
Gisel Urdaneta3
Kira Balestrini3
Rafael A Calderon-Candelario3,4
Michael A Campos3,4,*
Lawrence P Cahalin1,*

1Department of Physical Therapy, University of Miami Miller School of Medicine, Coral Gables, FL, USA;
2CAPES Foundation, Ministry of Education of Brazil, Brasilia, Brazil;
3Pulmonary Section, Miami Veterans Administration Medical Center, Miami, FL, USA;
4Division of Pulmonary, Allergy, Critical Care and Sleep Medicine, University of Miami Miller School of Medicine, Miami, FL, USA

*These authors contributed equally to this work

Purpose: The Test of Incremental Respiratory Endurance (TIRE) provides a comprehensive assessment of inspiratory muscle performance by measuring maximal inspiratory pressure (MIP) over time. The integration of MIP over inspiratory duration (ID) provides the sustained maximal inspiratory pressure (SMIP). Evidence on the reliability and validity of these measurements in COPD is not currently available. Therefore, we assessed the reliability, responsiveness and construct validity of the TIRE measures of inspiratory muscle performance in subjects with COPD.

Patients and methods: Test–retest reliability, known-groups and convergent validity assessments were implemented simultaneously in 81 male subjects with mild to very severe COPD. TIRE measures were obtained using the portable PrO2 device, following standard guidelines.

Results: All TIRE measures were found to be highly reliable, with SMIP demonstrating the strongest test–retest reliability with a nearly perfect intraclass correlation coefficient (ICC) of 0.99, while MIP and ID clustered closely together behind SMIP with ICC values of about 0.97. Our findings also demonstrated known-groups validity of all TIRE measures, with SMIP and ID yielding larger effect sizes when compared to MIP in distinguishing between subjects of different COPD status. Finally, our analyses confirmed convergent validity for both SMIP and ID, but not MIP.

Conclusion: The TIRE measures of MIP, SMIP and ID have excellent test–retest reliability and demonstrated known-groups validity in subjects with COPD. SMIP and ID also demonstrated evidence of moderate convergent validity and appear to be more stable measures in this patient population than the traditional MIP.

Keywords: psychometrics, respiratory muscles, chronic airflow obstruction

Introduction

Measures of respiratory muscle effort are considered to be useful parameters for the diagnosis, follow-up and prognosis of subjects with COPD.1–5 Disease progression is often associated with significant deconditioning of the respiratory musculature as evidenced by decreased inspiratory muscle performance.6 The inspiratory muscle dysfunction in COPD is attributed to myopathy related to systemic inflammation along with changes in chest wall geometry and diaphragm position.7 In addition, higher inspiratory load and ventilatory demand that result from hyperinflation further decrease the pressure-generating capacity of the inspiratory muscles.8 Clinical implications of a weak inspiratory musculature in COPD include poorer exercise and functional capacity, greater degrees of dyspnea and decreased health-related quality of life.3,8,9
The most traditional measure of inspiratory muscle performance in COPD is the maximal inspiratory pressure (MIP). MIP is commonly used as a diagnostic tool and outcome measure for trials of inspiratory muscle training in this population. Although previous studies have found MIP to be a reliable and valid measure of inspiratory muscle performance, MIP alone can only offer insight into inspiratory muscle strength, not encompassing other key elements of muscular performance such as work and endurance.

The Test of Incremental Respiratory Endurance (TIRE) provides a more comprehensive assessment of the inspiratory musculature by measuring muscle strength, endurance and work capacity in a single session. The TIRE includes measures of MIP, sustained maximal inspiratory pressure (SMIP) and inspiratory duration (ID) obtained during a maximal sustained inhalation, a distinct maneuver of this test. The TIRE MIP is defined as the highest pressure achieved during inspiration measured from residual volume (RV), while SMIP and ID reflect pressure-generating and endurance capacity from RV to total lung capacity. The TIRE inspiratory muscle performance measures are available in several devices such as the RT2 device (DeVilbiss Healthcare, Wollaston, UK), the TRAINAIR device (Project Electronics Ltd, Kent, UK) and the PrO2 device (Design Net, Smithfield, USA). Because of the potential usefulness of MIP, SMIP and ID in the evaluation of chronic lung diseases in general, it is important to establish the psychometric properties of these measures.

Pilot studies indicate that SMIP and ID have the potential to be clinically relevant tools in COPD, even more so than the traditional quasistatic MIP. However, evidence on the reliability and validity of these measurements in COPD is not currently available. Therefore, the purpose of this study was to assess the test–retest reliability, responsiveness and construct validity of the MIP, SMIP and ID as measures of inspiratory muscle performance in subjects with obstructive lung disease. We expected all three TIRE measures to be reliable but expect SMIP and ID to demonstrate stronger validity as measures of inspiratory muscle performance in individuals with COPD.

Methods
Reliability and validity testing was performed using a total of 81 male subjects, mainly of white ethnicity, with mild to very severe airflow limitation between the ages of 56 and 87 years. The entire study involved different experiments implemented simultaneously on four convenience samples of the same population. Table 1 provides a demographic summary of the participants in each psychometric testing sample.

All subjects were recruited from the outpatient pulmonary clinics at the Miami Veterans Affairs Healthcare System. Criteria for inclusion were males over the age of 35 with a clinical and functional diagnosis of COPD according to the Global Initiative for Obstructive Lung Disease (GOLD) criteria, regardless of disease severity. COPD was contemplated if subject had a significant smoking history (>10 pack-years), fixed airflow obstruction post-bronchodilator, and compatible symptomatology. Subjects were also expected to be clinically stable at the time of data collection (ie, free from exacerbations for at least 2 months prior to study entry). Exclusion criteria included presence of other concomitant pulmonary diseases (eg, asthma demonstrated by reversibility of airflow obstruction to short-acting bronchodilator, pulmonary fibrosis, pulmonary malignancies or lung surgery), as well as psychiatric, cognitive, neurologic or neuromuscular disorders that may affect performance of TIRE measures. Written informed consent was obtained from all subjects. The Miami Veterans Affairs Medical Center Institutional Review Board approved all procedures.

Table 1 Demographic characteristics and inspiratory muscle performance of the psychometric testing samples

<table>
<thead>
<tr>
<th>Variables</th>
<th>Test-retest reliability (n=21)</th>
<th>Known-groups validity (n=81)</th>
<th>Convergent validity: 6MWT (n=47)</th>
<th>Convergent validity: mMRC (n=53)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>Mean = 68.7, SD = 6.2</td>
<td>Mean = 69.9, SD = 5.7</td>
<td>Mean = 70.3, SD = 5.5</td>
<td>Mean = 69.5, SD = 5.2</td>
</tr>
<tr>
<td>Height (cm)</td>
<td>Mean = 173.9, SD = 8.1</td>
<td>Mean = 175.5, SD = 7.7</td>
<td>Mean = 175.3, SD = 8.1</td>
<td>Mean = 176.1, SD = 8.3</td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>Mean = 83.7, SD = 26.2</td>
<td>Mean = 80.9, SD = 19.2</td>
<td>Mean = 79.0, SD = 19.4</td>
<td>Mean = 80.9, SD = 21.0</td>
</tr>
<tr>
<td>Post-BD FEV1 (% predicted)</td>
<td>Mean = 48.1, SD = 21.2</td>
<td>Mean = 46.5, SD = 16.9</td>
<td>Mean = 43.5, SD = 16.4</td>
<td>Mean = 47.3, SD = 16.7</td>
</tr>
<tr>
<td>MIP (cmH2O)</td>
<td>Mean = 78.1, SD = 25.4</td>
<td>Mean = 77.5, SD = 23</td>
<td>Mean = 74.8, SD = 17.9</td>
<td>Mean = 75.8, SD = 19.7</td>
</tr>
<tr>
<td>SMIP (PTU)</td>
<td>Mean = 387.3, SD = 160.7</td>
<td>Mean = 403.8, SD = 135.5</td>
<td>Mean = 392.9, SD = 115.2</td>
<td>Mean = 410.6, SD = 117.9</td>
</tr>
<tr>
<td>ID (seconds)</td>
<td>Mean = 9.6, SD = 3.7</td>
<td>Mean = 10, SD = 2.8</td>
<td>Mean = 43.5, SD = 2.4</td>
<td>Mean = 10.5, SD = 2.5</td>
</tr>
</tbody>
</table>

Abbreviations: 6MWT, 6-minute walk test; ID, inspiratory duration; MIP, maximal inspiratory pressure; mMRC, modified Medical Research Council dyspnea scale; Post-BD FEV1, post-bronchodilator forced expiratory volume in the first second; PTU, pressure time units; SMIP, sustained maximal inspiratory pressure.
Test of incremental respiratory endurance

The TIRE was implemented through the PrO2 device, which links to a desktop or mobile computer using wireless technology and provides the user with a graphic representation of their inspiratory effort throughout all of inspiration with real-time biofeedback. This device has a fixed leak through a 2-mm-diameter opening that prevents glottal closure during maximal inspiration. The TIRE maneuver was performed with the subjects seated in a chair and wearing nose clips following published American Thoracic Society (ATS) standards on respiratory muscle testing.

Subjects were instructed to perform a maximal and sustained inspiratory effort following a full expiration from RV. They were encouraged to inspire deeply generating as much pressure as possible within 1–2 seconds of inspiration and continue to inspire maximally for as long as possible. A single TIRE session consists of three to five consecutive trials with rest intervals of 60 seconds between efforts. The greatest SMIP was used to identify the best of the repeated trials in each session. This trial was then considered for study purposes and provided the TIRE measures. Visual feedback and strong encouragement were given throughout testing. MIP was recorded in centimeters of water (cmH2O) and SMIP documented in pressure time units, representing the area under the curve generated from the start to the end of inspiration. The ID was recorded in seconds and characterized the total duration of inspiration during each sustained maximal inspiratory effort through the PrO2 device. Figure 1 presents an example of a maximal and sustained inspiratory effort obtained through the PrO2 device.

Reliability testing procedures

Test–retest was used to determine the reliability of MIP, SMIP and ID measurements in COPD. Inspiratory muscle performance measures were obtained from each subject twice, with time between test and retest sessions ranging from 30 to 60 minutes. We chose to perform two measurements on the same day (ie, before and after the subjects’ physician visit) to investigate test–retest reliability of the TIRE measures without the influence of COPD-related changes over time.

Validity testing procedures

Our validation strategy included known-groups and convergent validity assessments. Known-groups validity examined whether subjects with different levels of airflow limitation could be distinguished based on TIRE measures of inspiratory muscle performance. Spirometry was performed using the Vmax® Encore System (CareFusion, Yorba Linda, CA, USA) following ATS standards to quantify airflow limitation, providing post-bronchodilator measures of forced expiratory volume in 1 second (FEV1), expressed as the percentage of predicted values. For this experiment, participants were assigned to one of two groups based on their pulmonary function. Subjects with mild-to-moderate airflow limitation (ie, FEV1/forced vital capacity [FVC]<0.7; FEV1≥50% predicted) were included in the GOLD I–II group, while individuals with severe-to-very severe obstruction (ie, FEV1/FVC<0.07; FEV1≤49% predicted) were included in the GOLD III–IV group.

Convergent validity examined the degree to which inspiratory muscle performance measures obtained using...
the TIRE were associated with specific COPD-related outcomes, such as dyspnea and functional exercise capacity. The severity of dyspnea in relation to various physical tasks was assessed through the modified Medical Research Council (mMRRC) dyspnea scale, which ranges from zero to four, with higher scores indicating a greater degree of breathlessness. Functional exercise capacity was demonstrated by the greatest distance ambulated in meters during the 6-minute walk test (6MWT), which was performed as recommended by ATS guidelines.

Statistical analyses
Statistical analyses were performed with IBM SPSS Statistics 24 (Armonk, New York, USA 10504-1722) and included descriptive statistics and Shapiro–Wilk normality tests. Statistical significance was set at an alpha-level of 0.05.

Test–retest reliability and responsiveness
Intraclass correlation coefficients (ICC) were calculated with 95% CIs for all measures of inspiratory muscle performance using a two-way mixed-effects model for single measurements and absolute agreement. An ICC generally implies poor stability for values <0.40, fair stability for values between 0.40 and 0.59, good stability for values between 0.60 and 0.74 and an excellent agreement for values >0.74. The standard error of measurement, which represents the amount of variability that can be attributed to measurement error, was calculated as SD × \( \sqrt{1-ICC} \). Bland–Altman plots were also used to determine if any systematic differences across the range of values occurred between the two testing sessions.

Known-groups construct validity
Based on the assumption that a valid measure of inspiratory muscle performance should distinguish between subjects with COPD classified as GOLD I–II versus GOLD III–IV, we performed one-tailed independent-samples \( t \)-tests to compare the MIP, SMIP and ID values of subjects in the above groups. Hedge’s \( g \) was also calculated, providing a measure of effect size weighted according to the relative size of each sample for the independent-samples \( t \)-test analysis.

Convergent validity
Bivariate correlation (Pearson’s) analyses were used to examine the assumption that valid measures of inspiratory muscle performance should demonstrate a moderate relationship to COPD-related factors such as dyspnea and functional exercise capacity.

Results

Test–retest reliability
We assessed the test–retest reliability of MIP, SMIP and ID with a convenience sample of 21 male subjects with mild-to-very severe COPD between the ages of 56 and 86 years. Descriptive information of the test–retest study participants is shown in Table 1.

All TIRE measurements were found to be highly reliable in subjects with COPD with ICC scores ranging from 0.975 and 0.994. The results for test–retest reliability of MIP, SMIP and ID in the study sample are summarized in Table 2, with the associated Bland–Altman plots illustrating the agreement between the TIRE testing sessions in Figure 2.

Known-groups validity
Twenty-nine subjects were included in the GOLD I–II group, while 52 individuals were grouped as GOLD III–IV. Descriptive characteristics of the known-groups study participants are summarized in Table 1. Shapiro–Wilk normality tests found all measures of pulmonary function and inspiratory muscle performance to be normally distributed.

One-tailed independent-samples \( t \)-tests found significant differences in all TIRE inspiratory muscle performance measures obtained by subjects with mild-to-moderate obstruction when compared to those with greater severity. Table 3 shows the greater level of statistical significance and effect sizes found in SMIP and ID versus MIP.

Convergent validity
Descriptive statistics of the convergent validity study samples are provided in Table 1. Forty-seven individuals with mild-to-very severe COPD underwent TIRE and 6MWT measurements. The mean±SD 6MWT distance achieved by the subjects was 284.3±115.1 meters. Whereas no significant relationship was found between the 6MWT and the commonly used measure of inspiratory muscle strength MIP, both SMIP \((r=0.399, p=0.006)\) and ID \((0.413, p=0.004)\)
demonstrated moderate, positive correlations that were statistically significant with the 6MWT distance, as shown in Figure 3.

Fifty-three individuals with mild-to-very severe COPD underwent TIRE and dyspnea measurements. The mean±SD mMRC dyspnea scores of the subjects were 1.74±1.2. No significant association was found between dyspnea and MIP. However, both SMIP (r=-0.322, p=0.019) and ID (r=-0.320, p=0.019) demonstrated moderate, statistically significant inverse correlations with dyspnea severity, as shown in Figure 4.

### Discussion

To the best of our knowledge, this is the first study to examine the psychometric properties of MIP, SMIP and ID in individuals with obstructive lung disease using the PrO2 device. Poor inspiratory muscle strength and endurance were noted in the entire sample based upon previous work demonstrating the TIRE normative reference values.

All three measures of interest were found to be highly reliable in subjects with COPD. SMIP demonstrated the strongest test–retest reliability with a nearly perfect ICC of 0.994, while MIP and ID clustered closely together behind SMIP with ICC values of about 0.97, which is also considered excellent reliability. Given the changing clinical presentation of subjects with COPD, participants were assessed within approximately an hour between testing sessions, which appeared to be an adequate time frame for subjects to recover, providing the above results. The Bland–Altman plots (Figure 2) showed no systematic bias across the range of values. Of all test–retest measures, only one ID value fell outside the limits of agreement when considering the mean difference between the scores of the two testing sessions. Nevertheless, our findings highlight the excellent agreement between test–retest MIP, SMIP and ID values obtained through the TIRE PrO2 device in the study population.

When investigating the known-groups validity of the TIRE measures of inspiratory muscle performance in COPD, we observed that both inspiratory muscle strength (ie, MIP) and endurance (ie, SMIP and ID) were significantly reduced in individuals with more severe airway obstruction. This is consistent with findings from several previous studies suggesting inspiratory muscle strength as reflected by MIP to be considerably lower in subjects with more advanced COPD. Kim and associates found MIP to be decreased with increasing disease severity in 130 subjects, stating that differences are even more significant when subjects are divided into mild-to-moderate and severe-to-very severe groups, which is similar to our methodologic approach.

The SMIP, represented as the area under the curve in the pressure–time graph generated during the TIRE maneuver (Figure 1), and its differences in accordance with COPD stages have not been previously examined. In our known-groups analysis, SMIP yielded a more statistically significant known-groups difference and larger effect size...
when compared to MIP. This finding highlights the greater value SMIP appears to have in detecting differences among COPD severity groups over traditional measures of inspiratory muscle performance. In view of this, SMIP may then provide another surrogate of COPD severity that can be assessed. Further investigation of the diagnostic utility of SMIP in COPD is warranted.

Like SMIP, the ID demonstrated superior known-groups validity when compared to MIP. Our findings suggest that subjects with less airflow limitation were able to achieve a longer total duration of inspiration during each sustained maximal inspiratory effort through the TIRE PrO2 device. There is no previous literature examining differences in either SMIP or ID across COPD stages, but our analyses suggest that these are likely better markers of inspiratory muscle performance with greater discriminatory abilities than MIP. Data dealing with respiratory muscle endurance in this population are limited, thus identifying the novelty and importance of the above findings. Further investigation of the TIRE measures of inspiratory muscle endurance as diagnostic tools in COPD appears plausible.

Our analyses also confirmed convergent validity for both SMIP and ID, but not for MIP. We observed no significant association between MIP assessed through the TIRE method and the COPD-related outcomes used in this study. However, MIP values tended to be higher in subjects able to achieve
longer distances in the 6MWT and lower in subjects reporting more breathlessness, in agreement with findings from previous studies.3,4,30,31

SMIP and ID demonstrated evidence of significant moderate convergent validity, suggesting that subjects with greater single-breath work capacity and able to achieve a longer total duration of inspiration in the TIRE displayed greater performance in the 6MWT. Greater SMIP and ID values were also significantly related to less dyspnea in our study sample. TIRE measures of inspiratory muscle endurance and their relationship with clinically relevant COPD outcomes have not been previously examined, but our findings are consistent with preliminary data obtained by the current authors.16 Specific TIRE measures, SMIP and ID, appear to correlate better than MIP with important COPD outcomes.

Limitations

Our study is subject to several limitations. First, the study population was composed only of males recruited from a single medical center, of whom 61.7% were white, which may limit the generalization of the results. Further investigation in a larger sample of both males and females of different ethnicities with both stable and unstable COPD through the TIRE is likely to provide a further comprehensive understanding of the psychometric properties of the PrO2 device in assessing inspiratory muscle performance in obstructive lung disease. Also, the convergent validity analyses were cross-sectional and, therefore, could not assess the temporal relationships among the inspiratory muscle performance measures and COPD-related outcomes. Finally, the findings of this study are valid in subjects with clinically stable COPD only and may not be generalizable to subjects experiencing exacerbations.

Conclusion

TIRE measurements using the device described in the present study are highly reliable and valid in assessing the inspiratory muscle strength and endurance of subjects with stable COPD. MIP, SMIP and ID have excellent test–retest reliability and demonstrated known-groups validity, all significantly distinguishing between subjects with different levels of airflow limitation. TIRE measures of inspiratory muscle endurance also confirmed convergent validity. There is no previous literature examining the relationship between either SMIP or ID and COPD-related outcomes, but our findings suggest that SMIP and ID are more closely associated with functional exercise capacity and dyspnea when compared to the traditional MIP.

Disclosure

The authors report no conflicts of interest in this work.

References

17. Formiga MF, Campos MA, Balestrini K, Bagley V, Luna L, Cahalin LP. The test of incremental respiratory endurance is significantly related to pulmonary function in COPD. Eur Respir J. 2017;50(Suppl 61);PA2484.