Modified BODE indexes: Agreement between multidimensional prognostic systems based on oxygen uptake

José Luis Lopez-Campos
Pilar Cejudo
Eduardo Marquez
Francisco Ortega
Esther Quintana
Carmen Carmona
Emilia Barrot
Unidad Médico-Quirúrgica de Enfermedades Respiratorias, Hospitales Universitarios Virgen del Rocío, Seville, Spain

Aim: It has been recently shown that the original BODE index has a high degree of correlation with two modified BODE indexes using maximal oxygen uptake expressed either as mL/min/kg (mBODE) or as the percentage predicted (mBODE%). In this study we investigated the agreement between the two modified BODE indexes (mBODE and mBODE%) in patients with stable chronic obstructive pulmonary disease (COPD).

Methods: A total of 169 patients with stable COPD were enrolled in this cross-sectional study. Differences between the two mBODE indexes were assessed using kappa coefficients and Bland-Altman plots. One out of every three patients underwent the six-minute walking test to investigate the agreement with the original BODE index.

Results: Correlations between the two mBODE indexes with each other ($r = 0.96, P < 0.001$) and with the original BODE index (mBODE $r = 0.88, P < 0.001$; mBODE% $r = 0.93, P < 0.001$) were excellent. However, the two mBODE indexes were significantly different from each other (mBODE $5.27 \pm 2.3$ versus mBODE% $4.31 \pm 2.5; P < 0.001$). The kappa coefficients were significantly lower (entire study group $k = 0.5, P < 0.001$) for every GOLD stage. The mean difference between the two mBODE indexes was $0.8 \pm 0.6$ units. Differences with the original BODE were higher for the mBODE ($1.8 \pm 0.9$) than for the mBODE% ($0.6 \pm 0.8$).

Conclusions: The new mBODE indexes are highly correlated but significantly different from each other. The differences between the novel indexes deserve further scrutiny.

Keywords: BODE index, chronic obstructive pulmonary disease, exercise testing, multidimensional evaluation, oxygen uptake

Introduction

Chronic obstructive pulmonary disease (COPD) is a progressive and debilitating lung disease characterized by irreversible airflow limitation in the lungs, which is predicted to have a trend of steadily increasing morbidity and mortality in the near future. In addition to causing inflammatory and structural changes in the lung resulting in alterations of ventilatory mechanics and gas exchange, COPD may also have systemic manifestations, such as impairment in nutritional status and peripheral muscle dysfunction. In this context, a multidimensional grading system that assesses the respiratory and systemic expressions of COPD is important for categorization and outcome prediction.

Celli et al have recently introduced the use of the BODE index, a multidimensional 10-point scale assessing four prognostic factors in COPD patients, namely body mass index (B), degree of airflow obstruction (O), functional dyspnea (D), and exercise capacity (E) as assessed by the six-minute walk test (6MWT). The BODE index has
been extensively validated and shown to predict mortality and future hospitalizations. In addition, it has been found to be a good predictor of exacerbations and health-related quality of life.

However, several modifications of the original BODE index have been proposed to obtain novel indexes that may be more accurate and easier to use. In this regard, the 6MWT used in the original BODE index is a submaximal test, is dependent on individual motivation, and may not express the real physical capacity of the patient. Interestingly, correlations with the maximal oxygen uptake (VO$_{2\text{max}}$) are only moderate. In contrast, maximal cardiopulmonary exercise testing (CPET) is considered the gold standard to assess maximal exercise tolerance. In particular, VO$_{2\text{max}}$ is an objective measure of physical fitness and might serve as a physical training target, a predictor of survival rates, and a disability index for COPD patients. Of note, the original BODE index has been recently correlated with two modified BODE indexes replacing 6MWT with VO$_{2\text{max}}$ expressed either as the percentage of predicted values (mBODE%) or as absolute values in mL/min/kg (mBODE). Intriguingly, the modified BODE indexes have been recently linked to survival in COPD patients. In addition, a highly significant correlation between the original BODE and the modified indexes has been reported. It has been therefore suggested that the original BODE index is highly effective in the evaluation of COPD patients. Although valuable data have been presented – including the use of VO$_{2\text{max}}$ in the BODE index – an analysis of the agreement between the two modified BODE indexes has not yet been performed. Therefore, the question as to whether these indexes may have the same properties remains open. The purpose of this study was to analyze the agreement between the two modified BODE indexes in a group of patients with stable COPD. We also sought to investigate the relation between the modified BODE indexes and the original BODE in a subsample of patients.

**Materials and methods**

This is an observational, cross-sectional, prospective study. All procedures were approved by the local Institution Review Board and written informed consent was obtained from all participants.

**Subjects**

A total of 169 COPD patients were included in the study. Study participants were recruited from our outpatient clinics during routine visits. Patients fulfilling the following inclusion criteria were invited to participate in the study: a) being an adult patient diagnosed as having COPD according to international guidelines and b) being clinically stable during the previous three months. Exclusion criteria were: a) recent exacerbations (in the previous three months) requiring systemic corticosteroids or antibiotics, and b) any contraindication or inability to perform any of the study tests (uncontrolled heart disease, acute pulmonary disease, and other conditions). Medications already prescribed were continued during the study.

Patients underwent a complete clinical and functional evaluation followed by a maximal CPET on a cycle ergometer. To be able to compare our findings with those achieved using the original BODE index, one out of every three patients performed the 6MWT according to international recommendations. Clinical and functional evaluation consisted of dyspnea scoring, nutritional evaluation, spirometric testing, and arterial blood-gas analysis.

**Measurements**

**Dyspnea**

Basal dyspnea was assessed using the modified Medical Research Council (MRC) scale. This is a five-point scale based on degrees of various physical activities that precipitate dyspnea.

**Nutritional status**

Nutritional status was assessed by calculating body mass index (BMI). BMI was obtained by dividing body weight in kilograms by height in square meters (kg/m$^2$). A BMI of less than 21.0 kg/m$^2$ was considered to be underweight.

**Pulmonary function tests**

Spirometric investigations were carried out on a pneumotachograph spirograph Masterlab (Erich Jaeger GMBH, Wuerzburg, Germany), following American Thoracic Society (ATS) recommendations. For the purpose of this study, values of forced expiratory volume in one second (FEV$_1$) were recorded. Blood-gas analyses were carried out with an ABL-700 analyser (Radiometer, Copenhagen, Denmark) using arterial blood samples taken from the radial or humeral artery at rest while the patient was breathing room air. Arterial oxygen partial pressure (PaO$_2$), carbon dioxide partial pressure (PaCO$_2$), and pH were recorded following the manufacturer’s instructions and in keeping with Spanish Society of Pulmonology and Thoracic Surgery recommendations.

**Maximal cardiopulmonary exercise test**

Maximal CPET was performed on a cycle ergometer (Collins Respiratory; ErgoMed, Seattle, WA, USA). Maximal exercise tolerance was measured by a symptom-limited graded
exercise test performed while the patient was breathing room air. Breathing was continuously monitored, and inhaled and exhaled gases were collected. We determined the VO\textsubscript{2peak} and carbon dioxide production, an indirect measure of anaerobic threshold. VO\textsubscript{2peak} was expressed as an absolute value or as a percentage of the predicted value (in mL/min/kg). An average of the last 20 seconds of each minute was used for the purpose of analysis. The respiratory pattern (maximal breathing capacity, tidal volume, respiratory rate) during exercise was carefully recorded. Electrocardiogram and pulse oxymetry monitoring were performed during exercise. Data for ventilation and gas analysis were instantly obtained and analyzed breath-by-breath by a computer connected with the cycle ergometer. Oxygen was measured by a zirconia cell and carbon dioxide by an infrared cell. Patients were monitored for three minutes at rest and then began to pedal without any resistance for one minute. The level of resistance was then increased by 10–15 watts every minute until the maximal exercise level was reached. After exercise, we assessed heart rate and blood pressure. Dyspnea, leg fatigue, and chest pain were assessed after exercise according to a modified Borg’s scale.\textsuperscript{21}

Six-minute walking test
The 6MWT was performed a few days before CPET following the current recommendations.\textsuperscript{21} Prior to the index test, a trial test was performed to avoid a training effect.

Modified BODE Indexes
The two mBODE indexes were calculated using BMI, FEV\textsubscript{1}, and MRC as in the original index.\textsuperscript{7} However, the use of 6MWT was replaced with VO\textsubscript{2peak} values, expressed either in mL/min/kg (mBODE) or as the percentage of predicted (mBODE%). Each variable was rated on a four-point (0–3) scale, the only exception being BMI, which was coded as a dichotomous variable (0 or 1). mBODE indexes were calculated as the sum score ranging from 1 to 10. The scale was then divided into quartiles, with the first quartile corresponding to scores of 0–2 points, the second quartile to scores of 3–4 points, the third quartile to scores of 5–6 points, and the fourth quartile to scores of 7–10 points.\textsuperscript{7} Although several different stratifications may be theoretically used, VO\textsubscript{2peak} values were stratified as proposed by Cardoso et al.,\textsuperscript{17} to make results comparable. The values of VO\textsubscript{2peak} in mL/min/kg were therefore stratified according to previous methodology\textsuperscript{26} as follows: score 0 $\geq$ 25 mL/min/kg; score 1 = 20–25 mL/min/kg; score 2 = 15–20 mL/min/kg; score 3 $\leq$ 15 mL/min/kg. Similarly, values of VO\%\textsubscript{2peak} predicted were stratified\textsuperscript{27} as follows: score 0 $\geq$ 70%\textsubscript{pred}; score 1 = 60%–69%\textsubscript{pred}; score 2 = 40%–59%\textsubscript{pred}; score 3 $\leq$ 40%\textsubscript{pred}.

Statistical analysis
Statistical analysis was performed using the Statistical Package for Social Sciences (version 15.0; SPSS Inc., Chicago, IL, USA). Qualitative variables are expressed as absolute and relative frequencies. Quantitative variables are expressed as means and standard deviations. For descriptive purposes, variables are presented for the entire study cohort and according to GOLD stages.\textsuperscript{19} GOLD stages I and II were unified within one group since only eight patients had Stage I disease. Comparisons between qualitative variables were performed using the $\chi^2$ test or the Fisher’s exact test.

Correlations were analyzed by calculating the Pearson’s correlation coefficient (r). Comparisons of quantitative variables were performed using the Student’s t test. Differences across different GOLD stages were assessed using analysis of variance (ANOVA). Agreement between the mBODE and mBODE% was studied using the kappa coefficient (k) according to the quartiles of the two mBODE indexes. The degree of agreement was also investigated using the Bland–Altman plot (difference versus average). In this method, the mean differences between measures are reported as means $\pm$ 2 standard deviations in parentheses.\textsuperscript{28} The alpha error was set at 0.05.

Results
Comparison between modified BODE indexes
The clinical and functional characteristics of the 169 participants are shown in Table 1. Correlations between the two mBODE indexes and the corresponding VO\textsubscript{2peak} value were good (mBODE index: r = −0.68, P < 0.001; mBODE% index: r = −0.75, P < 0.001). The correlation between the two mBODE indexes was excellent (r = 0.96, P < 0.001, Figure 1). However, the two mBODE indexes were significantly different from each other (Table 1). These differences were evident for the entire cohort (P < 0.001) and throughout all GOLD stages (P < 0.001 for all stages, Table 1).

Kappa coefficients were significantly low in the entire study cohort (k = 0.5, P < 0.001), but improved in patients with advanced GOLD stages (GOLD I–II: k = 0.49, P < 0.001; GOLD III: k = 0.3, P < 0.001; GOLD IV: k = 0.69, P < 0.001).

Bland–Altman plots are presented in Figure 2. In the entire study cohort, the mean difference between the two mBODE indexes was 0.95 (−0.46–2.37) units. The mean differences according to GOLD stages were as follows: GOLD I–II: 1.1 (−0.4–2.63); GOLD III: 1 (−0.35–2.35), and GOLD IV: 0.7 (−0.5–1.98) units.
Table 1 Clinical and functional characteristics of the 169 study participants

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Cohort (n = 169)</th>
<th>GOLD I–II (n = 63)</th>
<th>GOLD III (n = 53)</th>
<th>GOLD IV (n = 53)</th>
<th>P value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Males</td>
<td>165 (97.6%)</td>
<td>62 (98.4%)</td>
<td>50 (94.3%)</td>
<td>53 (100%)</td>
<td>NS</td>
</tr>
<tr>
<td>Age (years)</td>
<td>64 ± 7</td>
<td>63 ± 6</td>
<td>64 ± 7</td>
<td>64 ± 7</td>
<td>NS</td>
</tr>
<tr>
<td>Dyspnea (MRC)</td>
<td>2.12 ± 1.2</td>
<td>1.52 ± 1.1</td>
<td>2.3 ± 0.9</td>
<td>2.6 ± 1.2</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>28.6 ± 5.09</td>
<td>28.9 ± 4.6</td>
<td>28.9 ± 4.6</td>
<td>27.8 ± 5.9</td>
<td>NS</td>
</tr>
<tr>
<td>FVC (%)</td>
<td>81.6 ± 17.9</td>
<td>94.9 ± 14.9</td>
<td>77.1 ± 13.2</td>
<td>70.07 ± 14.8</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>FEV₁ (%)</td>
<td>46.2 ± 17.4</td>
<td>64.1 ± 12</td>
<td>40.5 ± 5.9</td>
<td>30.5 ± 10.3</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>FEV₁/FVC (%)</td>
<td>44.4 ± 11.02</td>
<td>53.6 ± 7.8</td>
<td>41.7 ± 7.7</td>
<td>35.7 ± 8.5</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>pH</td>
<td>7.40 ± 0.03</td>
<td>7.41 ± 0.03</td>
<td>7.39 ± 0.02</td>
<td>7.40 ± 0.03</td>
<td>NS</td>
</tr>
<tr>
<td>PaCO₂ (mmHg)</td>
<td>42.5 ± 5.9</td>
<td>39.6 ± 3.6</td>
<td>42.3 ± 4.7</td>
<td>45.8 ± 7.4</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>PaO₂ (mmHg)</td>
<td>70.5 ± 10.4</td>
<td>75.3 ± 6.6</td>
<td>74.1 ± 7.9</td>
<td>61.7 ± 10.4</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>CPET power (watts)</td>
<td>63.7 ± 32.4</td>
<td>86.6 ± 32.4</td>
<td>59.4 ± 23.1</td>
<td>40.7 ± 20.4</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>VO₂max (mL/min/kg)</td>
<td>16 ± 4.9</td>
<td>18.7 ± 4.5</td>
<td>15.2 ± 4.3</td>
<td>13.5 ± 4.3</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>VO₂max (%)</td>
<td>61.8 ± 19.1</td>
<td>72.9 ± 18.1</td>
<td>58.9 ± 17.1</td>
<td>51.4 ± 15.1</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>mBODE</td>
<td>5.27 ± 2.3</td>
<td>3.1 ± 1.6</td>
<td>5.9 ± 1.4</td>
<td>7.5 ± 1.7</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>mBODE%</td>
<td>4.31 ± 2.5</td>
<td>2.02 ± 1.8</td>
<td>4.9 ± 1.5</td>
<td>6.4 ± 1.9</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

*Calculated by ANOVA with the exception of gender analysis (χ² test). Data are expressed as means and standard deviations.

Abbreviations: NS, not significant; MRC, Medical Research Council scale; BMI, body mass index; FVC, forced vital capacity; FEV₁, forced expiratory volume in one second; PaCO₂, arterial carbon dioxide partial pressure; PaO₂, arterial oxygen partial pressure; CPET, cardiopulmonary exercise test; VO₂max, maximal oxygen uptake.

Comparison of mBODE indexes with original BODE

In order to compare the mBODE indexes with the original BODE index, a subgroup of 54 patients (Table 2) performed the 6MWT (mean distance 517 ± 118 metres). There was a high correlation between original BODE and both the mBODE (r = 0.88; P < 0.001) and mBODE% (r = 0.93; P < 0.001) indexes (Figure 3). Kappa coefficients were low for both mBODE-BODE (k = 0.12; P < 0.001) and mBODE%-BODE (k = 0.38; P < 0.001). The agreement between the indexes was investigated using the Bland–Altman analysis. Results showed a mean difference in BODE points of 1.8 (0.03–3.77) for mBODE and 0.6 (−1.1–2.32) for mBODE% (Figure 4).

Discussion

The present study demonstrates that the new modified BODE indexes are significantly different from each other with a mean difference of around 1 BODE point. These differences were less pronounced in the advanced stages of the disease, in which they were characterized by better agreement. Moreover, the agreement with the original BODE index was overall weak, and significantly lower for the mBODE than for the mBODE%.

The original BODE index is a simple multidimensional grading system which is superior to FEV₁ alone for prediction of mortality and hospitalization rates among COPD patients. Since its first publication in 2004, several attempts have recently been made to improve the original BODE index, either by replacing or adding new components. In our study, two novel mBODE indexes were constructed by replacing 6MWT with VO₂peak expressed either as the
percentage of predicted values (mBODE%) or as the absolute values in mL/min/kg (mBODE). Although the novel indexes thus far have not been studied in detail, preliminary evidence has indicated that the mBODE% and the original BODE index show similar prognostic properties.18

The identification of novel diagnostic tools is an important clinical goal. Novel diagnostic tools or multicomponent measures are frequently being developed to improve prognostic models and to identify novel therapeutic targets for patients with COPD. In this context, novel assessment tools need to be validated prospectively against pre-existing methodologies. Two previous studies have investigated and validated the novel mBODE indexes.17,18 To this end, a feasible approach is simple linear regression analysis, as previously done by Cardoso et al.17 These authors have previously demonstrated a highly significant correlation between the original BODE index and the two mBODE indexes. Interestingly, in our study we were able to replicate these findings. However, it should be noted that the correlation between the original BODE index and VO_{2\text{max}} was only moderate.17 It is well known that good correlation does not necessarily imply good agreement and vice versa.28 Therefore, the agreement between both indexes should be carefully evaluated by means of sound statistical analysis. The second report18 was a validation study of mBODE% suggesting that this instrument has a good ability to predict mortality in a similar manner as the original BODE. However, the authors found a difference of 1.4 points between the mBODE% and the original BODE. The study

Figure 2 Bland-Altman plots exploring the differences between the two modified BODE indexes in A) the entire study cohort, as well as in B) patients with GOLD I–II, C) GOLD III, and D) GOLD IV. Mean differences: entire cohort: 0.95 (-0.46–2.37); GOLD I–II: 1.1 (-0.41–2.63), GOLD III: 1 (-0.35–2.35), and GOLD IV: 0.7 (-0.5 –1.98) units.
did not focus on the degree of agreement between the two measures. The present report expands previous studies on the mBODE indexes by analyzing the degree of agreement between each other using the kappa and Bland-Altman plots. Future longitudinal studies are needed to investigate whether the two new mBODE indexes differ in their ability to predict mortality or exacerbations in patients with COPD. Although this issue has been partly addressed in the past, our study was not of prospective design and is therefore not suitable to confirm previous findings.

In our study, the two mBODE indexes showed a low degree of agreement, despite a high degree of correlation. Interestingly, kappa coefficients were low for the entire study cohort and in all GOLD stages, improving in advanced stages. Hence, a Bland–Altman plot of difference versus mean values showed a difference of about 1 point among the two mBODE indexes, which have proven to be a clinically relevant prognostic difference.

Two findings of our study merit consideration. Firstly, patients in advanced stages of COPD had a higher degree of agreement. It is widely recognized that patients with advanced COPD suffer from severe impairment of functional capacity. In addition, the limitation in exercise capacity shows a lower variability in advanced stages characterized by a progressive reduction in the standard deviations of exercise variables (see Table 1). Secondly, the degree of agreement with the original BODE was different for the two mBODE indexes. Interestingly, the VO$_2$peak is the only variable that differs between the two parameters. VO$_2$peak is currently considered to be the best index of aerobic capacity and the gold standard for cardiorespiratory fitness. It represents the maximal achievable level of oxidative metabolism in large muscle groups. VO$_2$peak is generally expressed as oxygen uptake in absolute units. However, VO$_2$peak can be normalized for body weight to compare VO$_2$peak among subjects with different anthropometric measures. Of note, VO$_2$peak may be expressed in normalized and absolute units, and the two different modalities may express complementary facets

### Table 2 Clinical and functional characteristics of participants performing the six-minute walking test

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Entire cohort (n = 54)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Males</td>
<td>51 (94.4%)</td>
</tr>
<tr>
<td>Age (years)</td>
<td>63 ± 7</td>
</tr>
<tr>
<td>Dyspnea (MRC)</td>
<td>1.09 ± 0.7</td>
</tr>
<tr>
<td>BMI (kg/m$^2$)</td>
<td>28.8 ± 5.4</td>
</tr>
<tr>
<td>FVC (%)</td>
<td>90.7 ± 18.3</td>
</tr>
<tr>
<td>FEV$_1$ (%)</td>
<td>58.4 ± 19</td>
</tr>
<tr>
<td>FEV$_1$/FVC (%)</td>
<td>49.9 ± 10.8</td>
</tr>
<tr>
<td>pH</td>
<td>7.42 ± 0.03</td>
</tr>
<tr>
<td>PaCO$_2$</td>
<td>40.6 ± 5.5</td>
</tr>
<tr>
<td>PaO$_2$</td>
<td>71 ± 9.7</td>
</tr>
<tr>
<td>CPET power (watts)</td>
<td>84.5 ± 36</td>
</tr>
<tr>
<td>VO$_2$max (mL/min/kg)</td>
<td>173 ± 5.5</td>
</tr>
<tr>
<td>VO$_2$max (%)</td>
<td>71.4 ± 20.9</td>
</tr>
<tr>
<td>Six-minute walking test (metres)</td>
<td>517 ± 118</td>
</tr>
</tbody>
</table>

Data are expressed as means and standard deviations.

Abbreviations: MRC, Medical Research Council scale; BMI, body mass index; FVC, forced vital capacity; FEV$_1$, forced expiratory volume in one second; PaCO$_2$, arterial carbon dioxide partial pressure; PaO$_2$, arterial oxygen partial pressure; CPET, cardiopulmonary exercise test; VO$_2$max, maximal oxygen uptake.

**Figure 3** Scattergram and regression line showing a highly significant correlation between the original BODE and A) mBODE ($r = 0.88; P < 0.001$), as well as B) the mBODE% ($r = 0.93; P < 0.001$).
of aerobic capacity. It is possible that the use of different thresholds to stratify VO$_{2peak}$ values could yield different results regarding correlation and agreement. This may also explain the differences in agreement observed between the original BODE and the mBODE% compared with the mBODE. However, in this study we used the same thresholds proposed by Cardoso et al$^{17}$ to ensure comparability. Future studies should investigate the optimal cut-off values for VO$_{2peak}$.

One difference between the present study and that of Cardoso et al$^{17}$ relies in the use of a cycle ergometer for VO$_{2peak}$ stratification instead of a treadmill. It is widely recognized that both methods have some differences. Maximal oxygen uptake is reported to be, on average, 5%–10% higher on a treadmill than on a cycle ergometer. Since some abnormalities may occur only at the highest metabolic demand, the exact determination of VO$_{2max}$ in clinical practice is critical.$^{14}$ It should be noted, however, that both methods have been validated for COPD patients, with no clear difference between the two forms of exercise.$^{16}$

Conclusions

In summary, our study suggests that the novel mBODE indexes are highly correlated, but significantly different from each other (mean difference of approximately 1 BODE point) as well as from the original BODE index. Taken together, our data reinforce the idea that the original BODE index is useful in clinical characterization of COPD patients. Nonetheless, the differences between the two novel indexes deserve further scrutiny.

Acknowledgements

We are grateful to Enzo Emanuele for expert editorial assistance.

Disclosures

The authors report no conflict of interest in this work.

References


