Development of a Barthel Index based on dyspnea for patients with respiratory diseases

Background: As Barthel Index (BI) quantifies motor impairment but not breathlessness, the use of only this index could underestimate disability in chronic respiratory disease (CRD). To our knowledge, no study evaluates both motor and respiratory disability in CRD during activities of daily living (ADLs) simultaneously and with a unique tool. The objective of this study was to propose for patients with CRD an additional tool for dyspnea assessment during ADLs based on BI items named Barthel Index dyspnea.

Methods: Comprehensibility, reliability, internal consistency, validity, responsiveness, and ability to differentiate between disease groups were assessed on 219 subjects through an observational study performed in an in-hospital rehabilitation setting.

Results: Good comprehensibility, high reliability (intrarater interclass correlation coefficient was 0.93 [95% confidence interval 0.892–0.964] and test–retest intraclass correlation coefficient was 0.99 [95% confidence interval 0.983–0.994]), good internal consistency (Cronbach’s alpha 0.89), strong concurrent validity with 6 minute walking distance (Pearson r=−0.538, P<0.001) and Medical Research Council (Spearman r S=0.70, P<0.001), good responsiveness after rehabilitation (P<0.001), and good appropriateness of the index were found evidencing patients with different dyspnea severity. Divergent validity showed weak correlation (Pearson r=−0.38) comparing Barthel Index dyspnea and BI.

Conclusion: The BI based on dyspnea perception proved to be reliable, sensitive, and adequate as a tool for measuring the level of dyspnea perceived in performing basic daily living activities. A unique instrument simultaneously administered may provide a global assessment of disability during ADLs incorporating both motor and respiratory aspects.

Keywords: activities of daily living, psychometric tests, chronic obstructive pulmonary disease, pulmonary rehabilitation

Introduction

Dyspnea affects quality of life, exercise tolerance, and mortality in various disease conditions. Relief from dyspnea during activities of daily living (ADLs) represents the major goal of respiratory rehabilitation and its quantification through specific instruments (scales) is essential to define disability level and postrehabilitation improvement. Approximately 40% of patients with chronic obstructive pulmonary disease (COPD) report a degree of disability and 68% lose at least one relevant function in daily life.

In routine clinical practice, various scales measuring ADLs are used to evaluate patients’ motor and functional autonomy/disability, the most widely used being the Barthel Index (BI). This index was developed for chronic patients and long-term hospital patients with neurological diseases examining their performance...
before and after treatment and predicting time needed for motor rehabilitation and amount of nursing aid required. Unfortunately, the BI is not supported as a useful measure of ADLs in respiratory patients. Furthermore, by not taking breathlessness into account, BI equivalents, which quantify motor disability, could underestimate real disability in chronic respiratory disease (CRD).

Up to date there is no study describing the use of an instrument for assessing routinely both motor and respiratory disability. A unique tool could enlighten on these two different outcomes improving prescription of more tailored rehabilitation programs.

The aim of the present study was to describe the development and validation of an additional tool for dyspnea assessment for patients with CRD during ADLs based on the BI items named Barthel Index dyspnea (BI-d). The BI-d is expected to improve the assessment of chronic respiratory patients when administered simultaneously with BI.

Materials and methods
The study was approved by the Technical and Scientific Committee of the Salvatore Maugeri Foundation (protocol registered: 2013-04; CEC registered: CEC 1078). Patients gave their written informed consent for use of their clinical data for scientific purposes. The present study was carried out in two phases.

Phase 1: development of the BI-d
The development process had two stages.

Stage 1
A team of experts (three lung specialists, five physiotherapists, and one psychologist) discussed the content of the BI-d and the instructions for patients in order to capture their self-evaluation of dyspnea perception during the same ADL items included in the BI.19 In detail, experts decided to approach the original BI item on incontinence focusing mainly on dyspnea induced during the act of urinating or defecating instead of the item pertaining incontinence per se, being bladder and bowel control not nearly as central to pulmonary patients. A focus group of ten patients was arranged and patients were asked to answer the questions (performed in Italian) described in the Supplementary materials. The BI-d required someone to interview each patient. In summary, differences between BI and BI-d were related to the patient’s evaluation of the dyspnea symptoms during ADLs. In the BI case, operator signed the level of disability according to clinical conditions during ADL; while in the BI-d case, patients gave a judgment to their dyspnea between 0 = no sign of dyspnea and 4 = extreme severe level of dyspnea, such as to preclude or reduce specific requested activity. The total BI-d score ranges from 0 (no dyspnea) to 100 (maximum level of dyspnea) according to the original BI grading score. In this pilot phase, patients were also asked to discuss on the comprehensibility of the items describing them as “clear” or “unclear”.

Stage 2
In the second stage, the patient-perceived comprehensibility was tested on a sample of 50 patients. Patients were asked to rank their perceived comprehensibility of the questionnaire on a 4-point Likert scale, between 0 and 3 (0 = very difficult, 1 = difficult, 2 = with low difficulties, and 3 = no difficulties). Time for completing the BI-d was also assessed. Reliability of BI-d was assessed by means of interrater reliability (two raters) and of test–retest reliability between admission and 48 hours from admission maintaining the same interviewer. Reliability measures were based on the same sample of 50 patients and assessed by the intraclass correlation coefficient (ICC) and related 95% confidence interval (95% CI).

Phase 2: the observational study
In the second phase, the assessment of metric properties of the BI-d scale was carried out on an additional sample of 169 patients and internal consistency, validity, and responsiveness of BI-d were evaluated.

Study subjects
In the period between June 2013 and September 2014, 169 consecutive patients (Phase 2) were enrolled. The inclusion criteria were as follows: patients with any CRD (mainly COPD, emphysema, and chest wall diseases) admitted to any of three rehabilitation wards and with arterial blood gases (ABG) stability. In-hospital rehabilitation program was proposed after hospitalization or by general practitioner/pulmonologist due to progressive worsening of disability. Patients were excluded if any of the following were present: presence of respiratory exacerbation with clinical instability (fever, persistent purulent sputum, cough, oxygen saturation <89% under oxygen supply, respiratory rate at rest >25 acts/min, or heart rate >110 beats/ min), oncological or degenerative disease, neuromuscular degenerative diseases, and altered cognitive status.

Data collection and assessment
The patients enrolled in Phase 2, within 48 hours from admission, underwent the following assessments: anthropometrics evaluation (ie, age, sex, body mass index);
Cumulative Illness Rating Scale (CIRS1 and CIRS2);^{20} forced expiratory volume in 1 second (FEV$_1$)% predicted; forced vital capacity (FVC% predicted); residual volume; ABG analysis (pH, arterial partial pressure of oxygen (PaO$_2$), arterial partial pressure of carbon dioxide (PaCO$_2$), inspiratory fraction of oxygen (FiO$_2$), PaO$_2$/FiO$_2$ to normalize the ABG with different FiO$_2$); ongoing pharmacological therapy; exercise tolerance measured by the 6-minute walk test; the Medical Research Council (MRC) dyspnea scale;^{21} BI to assess motor disability using the Shah’s weight version;^{19} and the BI-d scale asking how the perceived subjective dyspnea was item by item.

Before discharge (within the last 24 hours of hospital stay), the patients underwent the following assessments again: ABG analysis, 6-minute walk test, MRC dyspnea scale, BI, and BI-d. The scales were always administered face to face by an expert interviewer (physiotherapist). For each patient, the number of pulmonary rehabilitation sessions was recorded.

Statistical methods

Descriptive statistics were performed reporting means and standard deviations [mean (SD)] for quantitative variables, frequencies, and percentages for qualitative variables. In the case of clearly non-Gaussian distributions of quantitative variables, medians, and interquartile ranges (IR) (expressed as difference between the 75th and 25th quartile) were reported. Internal consistency was measured through Cronbach’s alpha.

Both concurrent and divergent validity were assessed using Pearson correlation coefficient for analyses on quantitative variables and Spearman correlation in the case of ordinal variables. Responsiveness was investigated studying differences between pre- and postpulmonary rehabilitation. For this purpose a Kruskal–Wallis test was applied. In order to test the ability of the scale to detect subgroups of patients, the population was divided into two specific dyspnea subgroups of patients, with chronic respiratory failure (CRF) and without CRF. CRF was defined as a patient presenting PaO$_2$/FiO$_2 < 290$ (PaO$_2$/FiO$_2$ has been chosen to normalize the available ABG values under different FiO$_2$ prescribed by doctors) and differences between the two subgroups were investigated by means of a Mann–Whitney U-test. The data were analyzed by using Statistical Package for Social Science (SPSS, version 19; IBM Corporation, Armonk, NY, USA) software.

Results

Phase 1: stage 1

A focus group was conducted in a group of ten patients (five COPD, five with restrictive disease – one with interstitial lung disease and four with chest wall disease – mean age, 65.8 (6.2) years; four were female). With the aim to achieve consensus for each item included in BI-d, patients were asked to describe it as “clear” or “unclear”. All patients evaluated the scale as easily understandable. Final version of BI-d is reported in the Table S1.

Phase 1: stage 2

A pilot test on 50 patients (16 female, 34 male), mean age, 66.1 (8.4) years; FEV$_1$% predicted, 60% (19); FVC% predicted, 80% (27); FEV$_1$/FVC, 52% (26) was conducted. Patients who fulfilled the inclusion criteria were selected for the preliminary study to assess the amount of time taken and questionnaire comprehensibility for the patients. The average time to administer the questionnaire was 163 (50) seconds.

Patient-perceived comprehensibility

The median rank was 3 (no difficulties) with a range 0–3 showing that patients considered the scale easy or very easy.

Reliability

Both interrater and test–retest reliability reached high levels, considering that the interrater ICC was 0.93 (95% CI 0.892–0.964) and test–retest ICC was 0.99 (95% CI 0.983–0.994).

Phase 2

Patient characteristics

The demographic and clinical characteristics of the whole study group (n=169) are presented in Table 1.

Metric properties of the BI-d

Internal consistency

The internal consistency measured with Cronbach’s alpha was 0.89.

Concurrent validity

Relationships between BI-d and both 6 minute walking distance, as measure of effort tolerance, and MRC dyspnea scale, as measure of dyspnea, were assessed. Strong correlation between meters at 6-minute walking distance and the BI-d scale (Pearson r=-0.538, P<0.001) (Figure 1) as well as between MRC and BI-d (Spearman r$_s$=-0.70, P<0.001) were observed.

Figure 2 shows a box plot of the relationship between BI-d and MRC grading score. Increasing levels of BI-d were observed across increasing values of MRC from 0 to 4. A significant difference (P<0.01) was detected in post hoc analyses between BI-d at score 4 of MRC and each of the other subgroups.
Divergent validity

Relationship between motor disability and dyspnea showed a low correlation (Pearson $r=-0.38$) between BI-d and BI.

Responsiveness

The BI-d demonstrated a significant change over time in the level of perceived dyspnea across the whole population, from a median of 21 (IR: 28.5) to a median of 9.5 (IR: 15.25), $P<0.001$.

Patients underwent 20 (4) days of individually tailored rehabilitation activities, including lower and upper limbs endurance training, respiratory muscle training, and peripheral muscle strength exercises. After rehabilitation, a significant improvement ($P<0.001$) was also seen for all BI-d items.

Conversely, BI total score remained unchanged between admission and discharge (median value 98 and IR: 5).

Discussion

The presence of the dyspnea symptom can influence the execution of basic daily activities in patients affected by CRDs.
Patients with COPD were found to have breathlessness during normal physical activities, such as washing, drying, and dressing.\textsuperscript{18,22–24}

In clinical practice, the motor autonomy and ability of chronic patients are often evaluated through the well-known modified BI.\textsuperscript{19} However, the burden of motor disability on ADLs in patients with chronic pulmonary diseases is different from that seen in patients with neuromotor difficulties, because it is heavily influenced by the presence of the dyspnea symptom. The symptom’s presence can influence the execution of basic daily activities, making them discontinuous, or extremely slow.

To the best of our knowledge, no study is available on the adaption of a standardized ADL scale, such as BI, to assess the impact of breathlessness on performance of each ADL with the possibility to measure, simultaneously and with a unique instrument, both motor and respiratory disability.

The MRC scale\textsuperscript{21} assesses dyspnea by asking which activities, ranging from vigorous exercise to minimal ADLs, are limited by dyspnea. The MRC evaluates the impact of dyspnea on many of the activities in the BI, albeit with a greater emphasis on mobility. However, scales such as the MRC are indirect; they do not actually ask the patient how much dyspnea they experience assessing the degree to which breathlessness limits mobility.

Following a previous experience,\textsuperscript{25} we considered it useful to maintain the same items as in the modified BI’s score because keeping the same items for the subjective analysis of dyspnea made the BI-d’s comprehension easier and permitted collection of important information useful for a respiratory rehabilitation program. The scale’s acceptability reported by the patients was good, as well as the time spent to administer the scale. Therefore, the results show that the BI-d scale is a reliable and efficient instrument for investigating the level of dyspnea during daily activities in patients affected by CRD.

In order to test the effectiveness of the new score, we hypothesized a good agreement between BI-d total score and MRC measure (considered the gold standard). In fact, we found a good relation between the two scales, even if there was not a complete correspondence with the symptoms studied by the MRC. However, this fact was predictable, as the BI-d scale examines eleven daily activities, while the MRC scale considers just the walking activity. A major drawback of the MRC scale is the lack of a scale point for patients who experience dyspnea during different ADLs. The MRC scale is a strong tool for patient severity stratification according to breathless level; however, due to the restricted amplitude of score that may produce floor and ceiling effects, it is somehow difficult to point out minimal but clinically significant patient improvements after treatments (drugs and rehabilitation). On the contrary, we believe that a multi-items score focusing on different basic ADLs and dyspnea may better describe the real disability and modifications following rehabilitation programs.

The internal consistency, namely the degree to which the survey’s items are connected, and their ability to measure the same concept, was demonstrated by the good Cronbach alpha value showing the intrinsic homogeneity of the new instrument. In addition, the BI-d scale shows a good reliability, both interoperator and at the test–retest assessment. This demonstrates the generalizability of the instrument to different clinical and rehabilitative settings. The BI-d scale is able to capture changes following the rehabilitation program, showing that it could become a new outcome marker.
The BI-d shows a good internal consistency, as well as a good validity. Furthermore, the scale demonstrated a good ability to separate significantly, patients with CRF from those without CRF.

The good correlation between the BI, effort tolerance, and the BI-d scale shows that the general motor autonomy reduction in daily life is only partially able to describe the dyspnea disability during the same daily activities.

As study limitations, we did not create a new customized score based on ADL conditions found as the worst for dyspnea, but we decided to use “a priori” for all the items present in the BI score. The study includes mainly patients with COPD (80%) and thus the results are not well generalized into all respiratory patient populations; this fact reduces the applicability of the results for all respiratory conditions. The sample of patients with respiratory disease involved in this study was in a stable state, while patients with acute problems could present completely different results.

American Thoracic Society statement on dyspnea recommends categorizing measures as pertaining to domains of sensory-perceptual experience (what breathing feels like), affective distress, or symptom impact or burden. A BI-d measure such as presented in this study attempts to quantify the extent to which functional abilities are limited by dyspnea; therefore, it should be classified as a measure of dyspnea impact or burden.

We believe that this BI adaptation may be considered a complementary tool to the existing instruments for assessing some aspect of dyspnea in the respiratory population; the combined administration of the BI and the BI-d scale, with a unique and simple instrument administered simultaneously, could be useful in defining multifactor disability (both motor and dyspnea impact-related) and the corresponding components to include in different rehabilitation programs.

Future studies, 1) should include oxygen output measurement in order to assess which ADL is more involved and if it is associated to increased dyspnea, 2) should define the BI-d’s applicability in a large population, in different diseases prone to dyspnea during ADLs, with or without comorbidities, and in different clinical settings.

**Conclusion**

The BI-d is proved to be reliable, sensitive, and adequate as a tool for measuring the level of dyspnea perceived in performing basic daily living activities. Moreover, it is easy and quick to administer and it is correlated to the MRC scale. Based on this fact, this BI modification for dyspnea would aspire to add important information to patient’s level of motor disability, in patients who have CRD. Further studies using BI-d in respiratory rehabilitation programs are mandatory.

**Acknowledgments**

The authors thank Laura Comini (Scientific Direction S. Maugeri Foundation, IRCCS, Institute of Lumezzane, Brescia) for her constructive criticism in preparing the paper, Rosemary Allpress for the English revision of the manuscript, Carla Simonelli and Maria Laura Bertussi (Respiratory Rehabilitation Division, S. Maugeri Foundation, IRCCS, Institute of Lumezzane, Brescia) and Vito Spinelli (Respiratory Division, S. Maugeri Foundation, IRCCS, Institute of Cassano delle Murge, Bari) for their important clinical contribution.

**Author contributions**

All authors made substantial contributions to conception and design, acquisition of data, or analysis and interpretation of data; took part in drafting the article or revising it critically for important intellectual content; gave final approval of the version to be published; and agree to be accountable for all aspects of the work.

**Disclosure**

The authors report no conflicts of interest in this work.

**References**


Supplementary materials
Guidelines for administration of the Barthel Dyspnea Scale

The following phrases simulate the interview between operator and patient:

Dear Mr/Ms XXX, now, I will submit a questionnaire investigating your current level of dyspnea (within the last 2 days) during the same activities of daily life (ADLs) that I just described, for your limitations of the musculoskeletal system.

You will give a judgment to your dyspnea as follows: 0= no sign of dyspnea during execution of ADLs; 1= slight dyspnea that doesn’t prevent or slow down execution of ADLs; 2= moderate dyspnea, which can slow down ADLs; 3= severe dyspnea, which can greatly slow down ADLs; 4= extremely severe level of dyspnea, such as to preclude or reduce that activity. Now we begin the questionnaire analyzing the individual activities of daily life for which you will give me your score of your dyspnea.

The patient was holding a summary of the five options of severity of dyspnea, but he had no feedback about the content of the questions posed by the operator.

The operator read the following questions to the patient and then filled out the proper score from their responses:

1. “What is the degree of dyspnea during the course of your grooming-personal hygiene (washing your face, hair, teeth, shaving)?”
2. “What is the degree of dyspnea when bathing?”
3. “What is the degree of dyspnea when feeding-eating?”
4. “How is the degree of dyspnea when using the toilet (get up and sit by the water, undress and get dressed to go to the toilet)?”
5. “What is the degree of dyspnea when going up or down a flight of stairs?”
6. “What is the degree of dyspnea when dressing, including socks and shoes?”
7. “What is the degree of dyspnea when you must urinate?”
8. “What is the degree of dyspnea when you must defecate?”
9. “What is the degree of dyspnea when you walk at your own pace (more than 50 m)?”
9b. “What is the degree of dyspnea in case you need to use the wheelchair to get around for more than 50 m?”
10. “What is the degree of dyspnea when you are performing transfers, such as from bed to chair or vice versa?”

Table S1 BI-d

<table>
<thead>
<tr>
<th>Items</th>
<th>Assessment Options</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
</tr>
<tr>
<td>Grooming</td>
<td>0</td>
</tr>
<tr>
<td>Bathing</td>
<td>0</td>
</tr>
<tr>
<td>Feeding</td>
<td>0</td>
</tr>
<tr>
<td>Toilet use</td>
<td>0</td>
</tr>
<tr>
<td>Stairs</td>
<td>0</td>
</tr>
<tr>
<td>Dressing</td>
<td>0</td>
</tr>
<tr>
<td>Bowels</td>
<td>0</td>
</tr>
<tr>
<td>Bladder</td>
<td>0</td>
</tr>
<tr>
<td>Mobility</td>
<td>0</td>
</tr>
<tr>
<td>Wheelchair*</td>
<td>0</td>
</tr>
<tr>
<td>Transfers (bed to chair and back)*</td>
<td>0</td>
</tr>
</tbody>
</table>

**Notes:** *Not to be filled in if the patient is able to walk. Answer options: 0= no sign of dyspnea during execution of ADLs; 1= slight dyspnea that does not prevent or slow down execution of ADLs; 2= moderate dyspnea, which can slow down ADLs; 3= severe dyspnea, which can greatly slow down ADLs; 4= extremely severe level of dyspnea, such as to preclude or reduce that activity.

**Abbreviations:** ADLs, activities of daily living; BI-d, Barthel Index-dyspnea.