

Evaluation of the psychometric properties of the Nighttime Symptoms of COPD Instrument

Michelle Mocarski¹

Erica Zaiser²

Dylan Trundell²

Barry J Make³

Asha Hareendran²

¹Forest Research Institute, Inc., an affiliate of Actavis, Inc., Jersey City, NJ, USA; ²Evidera, London, UK; ³National Jewish Health, Denver, CO, USA

Background: Nighttime symptoms can negatively impact the quality of life of patients with chronic obstructive pulmonary disease (COPD). The Nighttime Symptoms of COPD Instrument (NiSCI) was designed to measure the occurrence and severity of nighttime symptoms in patients with COPD, the impact of symptoms on nighttime awakenings, and rescue medication use. The objective of this study was to explore item reduction, inform scoring recommendations, and evaluate the psychometric properties of the NiSCI.

Methods: COPD patients participating in a Phase III clinical trial completed the NiSCI daily. Item analyses were conducted using weekly mean and single day scores. Descriptive statistics (including percentage of respondents at floor/ceiling and inter-item correlations), factor analyses, and Rasch model analyses were conducted to examine item performance and scoring. Test-retest reliability was assessed for the final instrument using the intraclass correlation coefficient (ICC). Correlations with assessments conducted during study visits were used to evaluate convergent and known-groups validity.

Results: Data from 1,663 COPD patients aged 40–93 years were analyzed. Item analyses supported the generation of four scores. A one-factor structure was confirmed with factor analysis and Rasch analysis for the symptom severity score. Test-retest reliability was confirmed for the six-item symptom severity (ICC, 0.85), number of nighttime awakenings (ICC, 0.82), and rescue medication (ICC, 0.68) scores. Convergent validity was supported by significant correlations between the NiSCI, St George's Respiratory Questionnaire, and Exacerbations of Chronic Obstructive Pulmonary Disease Tool-Respiratory Symptoms scores.

Conclusion: The results suggest that the NiSCI can be used to determine the severity of nighttime COPD symptoms, the number of nighttime awakenings due to COPD symptoms, and the nighttime use of rescue medication. The NiSCI is a reliable and valid instrument to evaluate these concepts in COPD patients in clinical trials and clinical practice. Scoring recommendations and steps for further research are discussed.

Keywords: nighttime symptoms, PRO, psychometric validation, COPD

Introduction

The evaluation of nighttime symptoms in chronic obstructive pulmonary disease (COPD) is important in understanding the patient's experience of COPD. Nighttime symptoms of COPD have been linked with disturbed sleep, frequent awakenings, nocturnal airflow obstruction, and nocturnal hypoxemia.¹ Nighttime dyspnea has also been found to be a significant predictor of poor prognosis in individuals with COPD.² There is evidence that COPD symptoms occurring at night can be particularly bothersome to patients, especially due to their impact on sleep.³ In an internet survey conducted with COPD patients, 25% said their symptoms were worse at nighttime. This was even more of an issue in severe COPD patients, with 34% of this population reporting that their symptoms were worse than usual at night.⁴ This evidence suggests

Correspondence: Asha Hareendran
Evidera, Metro Building, 6th Floor,
1 Butterwick, London W6 8DL, UK
Tel +44 20 8576 5007
Email asha.hareendran@evidera.com

that it would be beneficial to have an instrument that can specifically measure the presence of nighttime symptoms and the impact they have on patients with COPD.

The Global Initiative for Chronic Obstructive Lung Disease (GOLD) recommends evaluating symptoms, lung function, and exacerbations to assess COPD.⁵ Since the perception of the severity of symptoms is a subjective experience, symptom severity is ideally evaluated based on data collected using patient-reported outcome (PRO) instruments:

A measurement based on a report that comes directly from the patient about the status of a patient's health condition without amendment or interpretation of the patient's response by a clinician or anyone else.⁶

In diseases such as COPD, where symptoms can vary day to day, daily diaries are often the method of choice to capture the severity on each day with minimal recall bias.⁷ This also facilitates the computation of the average severity and variability over time, as well as symptom-free periods.

The Nighttime Symptoms of COPD Instrument (NiSCI) was developed to collect data for evaluating the treatment benefit of interventions that may reduce nighttime symptoms of COPD in clinical trials and in clinical practice.⁸ It was developed based on qualitative research with COPD patients according to the methods outlined in the US Food and Drug Administration's "Guidance for industry on patient-reported outcome measures: use in medical product development to support labeling claims".⁶ The instrument is designed as a self-completed electronic daily diary to measure the occurrence and severity of nighttime symptoms in patients with COPD, nocturnal awakening due to COPD symptoms, and nighttime rescue medication use. The conceptual framework for the instrument was developed based on concept elicitation from COPD patients (four focus groups; $n=27$) about their experiences of nighttime symptoms of COPD, with clinical validity supported by the literature, and information obtained through interviews with expert clinicians. Patient understanding of the instrument was tested via cognitive interviews with ten COPD patients. For full details on the development of the instrument, see Hareendran et al.⁸ Initial testing was conducted within a clinical trial for a new COPD maintenance medication.⁹

The objectives of this study were to reduce the final number of items, inform scoring, and examine the psychometric properties of the NiSCI. The aim was to ensure that the NiSCI is a valid and reliable instrument^{6,10–13} for collecting

data about patients' experience of COPD symptoms during the night.

Methods

Study design and population

The data used in this study were from a prospective Phase III, multicenter, multinational, randomized, parallel-group, active- and placebo-controlled clinical trial that had a 2-week run-in period followed by a 24-week treatment period. The trial was an efficacy and safety study that aimed to investigate the effects of a new fixed-dose combination bronchodilator for COPD maintenance treatment (aclidinium and formoterol) versus the monotherapy components.⁹ Participants aged 40 years or older with moderate-to-severe COPD and a smoking history of 10 pack-years or more (1 pack-year is the equivalent of smoking 20 cigarettes every day for 1 year) were randomized in the trial. Moderate-to-severe COPD was defined using the GOLD criteria: postalbuterol/salbutamol forced expiratory volume in 1 second (FEV_1) $\geq 30\%$ to $<80\%$ predicted and FEV_1 /forced vital capacity (FVC) $<70\%$ predicted. Patients were excluded from the trial if they did not maintain regular day/night waking/sleeping cycles (eg, night shift workers) in order to control for the impact of disturbed sleep not due to COPD. Only stable patients (who had not been hospitalized for an acute COPD exacerbation within 3 months prior to visit 1) were included.⁹

Measures

The e-diary included the NiSCI for completion in the morning to record any symptoms that occurred during the previous night and the Exacerbations of Chronic Obstructive Pulmonary Disease Tool-Respiratory Symptoms (E-RS)¹⁴ for completion at night to describe symptoms over that day. Thus, patients completed questions on an e-diary device twice a day.

Patients also completed the St George's Respiratory Questionnaire (SGRQ)¹⁵ and the Patient Global Impression of Change (PGIC) during study visits. Additionally, clinical trial personnel conducted spirometry tests and administered the Baseline Dyspnea Index (BDI) during study visits.¹⁶

NiSCI

The NiSCI was completed by patients each morning (7 am to 11 am) using an electronic daily diary (which included skip patterns). The NiSCI prompts patients to respond to items describing their experience from the time period between when they went to bed and the time they woke up and got out

of bed to start their day. The NiSCI was designed to measure three concepts of interest in patients with COPD:

1. Occurrence and severity of nighttime symptoms: six symptoms (“Did you experience any of the following last night: cough, wheezing, shortness of breath, tightness in the chest, chest congestion, and difficulty bringing up phlegm?”) and overall severity of symptoms (“Overall, how severe were your COPD symptoms last night?”)
2. The impact of these symptoms in terms of nighttime awakenings: (“Last night, did you wake up because of your COPD symptoms?”; “How many times did you wake up because of your COPD symptoms?”)
3. Rescue medication use: (“How many puffs of your rescue medication did you take last night?”)

If patients experienced a specific symptom in the previous night, they were asked to indicate the severity of the individual symptom (eg, “How severe was your cough?”) from mild to very severe on a four-point scale. Responses were coded from 1 to 4 (ie, mild=1, moderate=2, severe=3, very severe=4). If patients indicated that they had woken up due to COPD symptoms, they were asked to note the number of times that they had woken up due to these symptoms.

Respiratory symptoms (E-RS)

The E-RS was designed as a standardized respiratory symptom diary and utilizes eleven respiratory symptom items from the 14-item Exacerbations of Chronic Pulmonary Disease Tool.¹⁷ The E-RS has been demonstrated to be a valid and reliable measure of symptom severity in COPD patients.¹⁴ It yields a total score and three subscale scores, with higher scores indicating more severe symptoms. Item-level scores range from four to five points (0–3 or 0–4), which are summed to yield total and subscale scores. The respiratory symptoms (RS) total score is an aggregate of three domains: the RS-chest symptoms domain (derived sum of three items), the RS-cough and sputum domain (derived sum of three items), and the RS-breathlessness domain (derived sum of five items). The E-RS was administered daily as part of an e-diary during the clinical trial.

Health status (SGRQ)

The SGRQ, a validated measure of impaired health in diseases of chronic airflow limitation that has been widely used in clinical trials in COPD,^{15,18} contains 50 items divided into three subscales: symptoms, activity, and impacts. A score is calculated for each section and a total score is also calculated. In each case, the lowest possible value is 0 and the highest

is 100. Each item has an empirically derived weight, with the lowest equal to 0 and the highest equal to 100. Scores for each subsection are calculated by dividing the summed weights by the adjusted maximum weights for that component, and the results are then expressed as a percentage. Higher values correspond to greater impairment of quality of life. The SGRQ was administered on an electronic device during study visits: at baseline, week 4, week 12, and week 24 of the clinical trial.

PGIC

The PGIC assesses the patient’s perspective on how their nighttime COPD symptoms have changed since the start of the study. The PGIC has a seven-point response grade ranging from “very much worse” to “very much improved”. The PGIC provides a subjective summary index of patients’ perception of their degree of improvement (or worsening). The PGIC was administered on an electronic device during study visits: at baseline, day 4, week 4, week 12, week 18, and week 24 of the clinical trial.

Spirometry

The largest value of three technically satisfactory forced exhalation efforts measuring forced FEV₁ and FVC was used. FEV₁ measures the volume of air that can be forcibly expired in 1 second. FVC is the volume (in liters) of air that can be expired after full inspiration. These two pulmonary function tests are standardized measurements that are commonly used to assess patients with respiratory disorders.

Dyspnea (BDI)

Measurement of dyspnea was recorded at baseline using the BDI.¹⁹ The evaluation of dyspnea was performed by an independent interviewer experienced in taking histories of respiratory disease. The BDI can generate scores on three domains: functional impairment, magnitude of task, and magnitude of effort, as well as a total focal score.

Statistical analyses

A statistical analysis plan was created a priori for conducting the psychometric analyses. The intention was to conduct secondary analyses of blinded data pooled across all treatment groups. The analyses were conducted in two phases: the item analyses and scoring definitions phase followed by a phase testing the psychometric properties of the instrument. The PRO analytic sample was defined as subjects in the intention-to-treat population from the trial who had baseline

(before study medication) PRO assessments, including the NiSCI and clinical data required for analyses. Statistical analyses were performed using statistical analysis system software SAS version 9.1® (Cary, NC, USA). For analysis of the week score (a 7-day average), a minimum of four (of seven) diary entries during the 7-day period were required, following the approach of Junghard and colleagues.²⁰

Two random split-half samples were created based on data pooled across the treatment groups. The first split-half sample was used for item evaluation, item reduction, and development of scoring algorithms. Once the final version of the NiSCI was created and the scoring algorithms were determined, reliability and validity were examined using data from the second split-half sample.

Item analysis and psychometric validation were conducted using the average of the 7 days leading up to and including the day of randomization (referred to in this manuscript as baseline week) and a single day score 7 days before randomization (referred to as baseline day).

Item reduction and scoring algorithm development

Items were analyzed to determine if any items could be removed, if they were redundant, or if they had poor psychometric characteristics. Item descriptives and item-to-item correlations were conducted to examine the distributional characteristics of the individual NiSCI items. The following criteria were considered for flagging items as poorly performing: items that show a floor (minimum response >30% of patients) or ceiling (maximum response >30% of patients) effect; item-item correlations >0.80; factor loadings <0.3 or misfit to the Rasch model; a high negative (<-3.0) residual number, which suggests an overfitting item (meaning that the information provided by this item does not add any new information to the measurement); or a high positive (>3.0) residual number, which suggests that the item is underfitting (indicating that the item has a poor fit to the model and the response categories are underdiscriminating or not discriminating differences in severity). The following statistics and threshold values were used to evaluate the model fit: comparative fit index ≥ 0.90 ,²¹ root mean square error of approximately ≤ 0.08 , and root mean square residual number ≤ 0.05 . The results from the confirmatory factor analysis (CFA) and Rasch model analysis were used to inform item deletion or retention of items through an iterative process and to determine the scoring algorithm for the measure.

An item analyses meeting, which included participation of the PRO tool developers, statisticians, and a COPD clinical expert, was conducted following an initial item evaluation, factor analysis, and Rasch model analysis. Attendees at the meeting included four of the authors of this paper, as representatives of the PRO tool development team (MM, AH, EZ) and a COPD clinical expert (BM). In addition, two expert statisticians attended to provide statistical advice and interpretation, and another clinician involved in drug development attended to provide clinical advice on the conceptual importance of items. The meeting involved iterations of item analyses and the CFA and Rasch model analyses to ensure that a mixed method, including information from psychometric, PRO tool development, and clinical relevance perspectives, was considered for evaluating the content validity of the instrument and for determining its scoring.

Validity and reliability testing was conducted for the scores on the final instrument (scored using the algorithm finalized at the item analysis meeting). Internal consistency was assessed using Cronbach's coefficient alpha, with >0.70 considered to indicate good reliability.²² Test-retest reliability was primarily assessed by an intraclass correlation coefficient (ICC) within a prespecified subset of stable patients (those who selected "no change" in nighttime symptoms on the PGIC between the baseline week data and average scores from the week leading up to and including study visit 4, termed visit 4 data). It was expected that there should be no significant differences in NiSCI scores when there is no change in the concept of interest (in this case, nighttime symptoms). An ICC ≥ 0.7 would indicate good test-retest reliability, scores between 0.4 and 0.7 indicate moderate reliability, and scores <0.4 indicate low test-retest reliability.^{23,24}

The construct validity (ie, the extent to which a scale actually measures what it is hypothesized to measure) of the NiSCI was tested by examining its correlation with other indicators of similar/related constructs²⁴ (SGRQ total for health status and symptom subscale scores, E-RS total score for RS, and trough morning predose FEV₁ value for lung function). All relationships were assessed using the Spearman rank order correlation coefficients: scores indicative of greater incidence and impact of nighttime symptoms were expected to be associated with worse health status and more severe symptoms. Moderate correlations were expected (± 0.50). Lower correlations were expected with lung function, as COPD symptoms have been shown to have a poor relationship with such physiological parameters.²⁵ It was

anticipated that construct validity would be supported when the NiSCI scores are substantially correlated (>0.40) with items or scales measuring similar concepts.²⁶

Known-groups validity (ie, the extent to which scores from an instrument are distinguishable from groups of subjects that differ by a clinically relevant marker or other indicator) was examined by exploring the relationship of NiSCI scores to clinical measures of disease status:^{27,28} GOLD spirometry severity stage (I–IV), SGRQ total score (\leq sample median versus $>$ sample median), SGRQ symptoms score (\leq sample median versus $>$ sample median), and E-RS total score (\leq sample median versus $>$ sample median). Analysis of variance models were used to assess the significance of the differences in mean scores on the NiSCI at baseline week for each of the groups described.

Results

Sample description

Sociodemographic characteristics

A total of 1,663 participants were enrolled in this study, with a mean age of 63.9 ± 8.9 years (range: 40–93 years); 46.8% ($n=778$) of the sample were female. Participants had a mean predose FEV₁ of 1.38 ± 0.52 L and FVC of 2.77 ± 0.85 L, with each of the means increasing postdose (FEV₁: 1.49 ± 0.53 L; FVC: 2.97 ± 0.88 L). The mean percentage predicted FEV₁ at baseline was $49.0\% \pm 14.0\%$. Most participants were categorized as GOLD stage II ($n=946$; 56.9%) or III ($n=696$; 41.9), with only four participants classified as GOLD stage I (0.24%) and 12 as GOLD stage IV (0.72%). Participants had a mean BDI of 6.4 ± 2.2 , and the most common COPD-related concomitant medication was a short-acting β_2 -adrenergic agonist ($n=940$; 56.5%), followed by inhaled corticosteroids ($n=549$; 33.0%) and systemic corticosteroids ($n=243$; 14.6%). Participants had an average SGRQ total score of 46.5, an SGRQ symptom score of 65.1, and a 9.6 on the E-RS. The key patient demographics and clinical characteristics at day 1/screening are presented in Table 1. Further details of the sample are provided in the published clinical trial (NCT01437397).⁶

Patients' experience of COPD symptoms at nighttime

During the night before the baseline day, 771 (50.3%) participants experienced coughing, 609 (39.7%) experienced wheezing, 702 (45.8%) experienced shortness of breath, 427 (27.8%) experienced tightness in their chest, 471 (30.7%) experienced chest congestion, and 321 (20.9%) experienced difficulty

Table 1 Sociodemographic and clinical characteristics

Characteristics	Overall sample day 1 ^a ($n=1,663$)
Age, mean (SD)	63.9 (8.9)
Sex female, n (%)	778 (46.8)
Race/ethnicity, n (%)	
American Indian/Alaskan native	7 (0.4)
Asian	6 (0.4)
Black or African American	95 (5.7)
Native Hawaiian/Pacific Islander	2 (0.1)
White	1,550 (93.2)
Other	3 (0.2)
FEV ₁ , mean (SD), L	1.4 (0.5)
FEV ₁ , % predicted, mean (SD)	49.0 (14.0)
FVC, mean (SD), L	2.77 (0.85)
SGRQ, ^b n	1,648
Total score, mean (SD)	46.5 (18.5)
Symptoms domain score, mean (SD)	65.1 (21.4)
Activity domain score, mean (SD)	60.6 (22.3)
Impacts domain score, mean (SD)	32.7 (19.4)
E-RS total, ^c mean (SD; n)	9.6 (6.30; 1,658)
GOLD stage, n (%)	
I	4 (0.2)
II	946 (56.9)
III	696 (41.9)
IV	12 (0.7)

Notes: ^aFor patients who were missing the NiSCI diary entry on day 1 (day of randomization), data from the closest day before day 1 were used. ^bSGRQ scores range from 0 to 100; higher score=more severe health status. ^cThe E-RS total score is an aggregate of three domains identified on the E-RS. Scores range from 0 to 40; higher score=more severe respiratory symptoms.

Abbreviations: SD, standard deviation; FEV₁, forced expiratory volume in 1 second; FVC, forced vital capacity; SGRQ, St George's Respiratory Questionnaire; E-RS, Exacerbations of Chronic Obstructive Pulmonary Disease Tool-Respiratory Symptoms; GOLD, Global Initiative for Chronic Obstructive Lung Disease; NiSCI, Nighttime Symptoms of COPD Instrument.

bringing up phlegm. On baseline day, 311 (20.3%) participants had awoken due to COPD symptoms the previous night. On average, participants took 1.50 ± 1.80 puffs of rescue medication the previous night. The mean scores from the symptom severity items during baseline week are presented in Table 2.

Item analyses for item reduction

Most item-to-item correlations were significant. High correlations were found between the overall severity of COPD symptom item and the severity of shortness of breath item at both baseline day (0.79) and baseline week (0.80). The predetermined cutoff for identifying potential item redundancy was a correlation of 0.80, a threshold reached at baseline week for these items. Excluding the correlation between these items, the correlations between NiSCI items ranged from 0.14 to 0.72 (baseline day) and from 0.11 to 0.73 (baseline week). The high item-to-item correlation potentially demonstrates

Table 2 Patients' experience of chronic obstructive lung disease symptoms at nighttime evaluated using the NiSCI: baseline week^a

NiSCI items	n	Mean (SD)	Median	Range
Cough severity ^b	1,630	0.77 (0.76)	0.57	0–3.8
Wheezing severity ^b	1,630	0.61 (0.74)	0.29	0–3.8
Shortness of breath severity ^b	1,630	0.76 (0.81)	0.50	0–3.8
Tightness in chest severity ^b	1,627	0.43 (0.65)	0	0–3.8
Chest congestion severity ^b	1,627	0.47 (0.70)	0	0–3.8
Difficulty with bringing up phlegm severity ^b	1,630	0.41 (0.76)	0	0–4.0
Overall nighttime symptoms severity ^b	1,630	1.03 (0.70)	1	0–3.8
Number of nighttime awakenings ^c	243	2.03 (1.06)	1.8	1–≥10
Nighttime puffs of rescue medication	1,630	1.39 (1.49)	1	0–9.4

Notes: ^aBaseline week: 7 days up to and including randomization day. ^b0=no symptoms; 1=mild; 2=moderate; 3=severe; and 4=very severe. ^cNighttime awakening is censored at ten times. Only those who answered "yes" to a question on if they woke up due to nighttime symptoms were included.

Abbreviations: NiSCI, Nighttime Symptoms of COPD Instrument; SD, standard deviation.

the importance of shortness of breath as a driver of symptom severity, and hence it was not considered for deletion.

The overall symptom severity item was retained as a summary of the patients' overall experience of COPD symptoms, which may include factors other than the six symptoms specifically noted in the NiSCI. When a Rasch model analysis using the overall symptom severity item and the six specific symptom items was conducted, the overall symptom severity item did not fit the model, as indicated by chi-square probability <0.001 (fit residual: -3.63). This is likely due to redundancy between the six symptom items and the overall symptom severity item. A second Rasch model analysis was conducted using just the six symptom items.

While most items were identified as having correctly ordered response categories, one item ("how severe was the difficulty with bringing up phlegm?") was identified as having incorrectly ordered response categories (suggesting that the "mild" response option is redundant), indicted by the threshold parameters. All of the items were identified as fitting the model (chi-square probability >0.001) and fell within the acceptable -3.0 to 3.0 fit residual range. The overall model-fit chi-square was 81.42 ($P<0.002$). The mis-ordering of the item threshold parameters for the item "how severe was the difficulty with bringing up phlegm?" was likely caused by a small percentage of subjects who chose the response option "mild", relative to the large percentage who indicated they did not experience difficulty bringing up phlegm at all (5.6% vs 79.0%, respectively). Given the small percentage of respondents choosing "mild", it is not advisable to drop the response option based on the Rasch results, especially considering its inclusion in the other items (where consistency between items in their response options will likely reduce patient burden). A summary of the item analyses for decision making can be found in Table 3.

Determining scoring

The results of the Rasch analysis and CFA suggested that the overall symptom severity item should be scored separately from the other symptom severity items. Therefore, it was proposed that two scores would be generated to describe patients' experiences of severity of symptoms at night: 1) a six-item symptom severity score including the six NiSCI COPD symptom-specific severity items and 2) an overall symptom severity score based on the NiSCI overall COPD symptom severity item. The revised conceptual framework reflecting these changes is shown in Figure 1. A decision was made that the NiSCI items on the number of nighttime awakenings and rescue medication use would be examined separately as individual items, as they conceptually define the impact of symptoms as in the original conceptual framework. These four provisional NiSCI scores – six-item symptom severity score, overall symptom severity score, rescue medication score, and number of nighttime awakenings score – were tested for their psychometric properties.

Psychometric properties

Reliability

Internal consistency reliability (assessed using Cronbach's alpha) was 0.85 for the six-item symptom severity score, indicating good internal consistency, as the individual items are highly related to each other and to the scale as a whole, without evidence of redundancy. Furthermore, Cronbach's alpha values decreased slightly when individual constituent items were deleted (range: 0.821–0.847).

Good test-retest reliability (ICC: >0.7) was demonstrated for all NiSCI scores (ICC range: 0.82–0.85) except the rescue medication score (ICC: 0.68), which showed moderate reliability. There was strong agreement between the ICC and the concordance correlation coefficient for all scores.

Table 3 Summary of NiSCI item analyses for decision making

n=831	Floor effect % (baseline day)	Floor effect % (baseline week)	Inter-item correlation (baseline day)	Inter-item correlation (baseline week)	CFA loading (baseline day)	CFA loading (baseline week)	IRT fit	Decision to reject or accept
NiSCI item 1a How many times did you wake up because of your COPD symptoms?	NA	NA	≤0.369	≤0.306	NA	NA	NA	Accept
NiSCI item 2ai How severe was your cough?	49.7%*	29.5%	≤0.269	≤0.727	0.690	0.758	Yes	Accept
NiSCI item 2bi How severe was your wheezing?	60.3%*	41.4%*	≤0.601	≤0.692	0.695	0.759	Yes	Accept
NiSCI item 2ci How severe was your shortness of breath?	54.2%*	32.1%*	≤0.719	≤0.801*	0.708	0.783	Yes	Accept
NiSCI item 2di How severe was the tightness in your chest?	72.0%*	51.0%*	≤0.559	≤0.634	0.729	0.775	Yes	Accept
NiSCI item 2ei How severe was your chest congestion?	69.2%*	51.4%*	≤0.541	≤0.625	0.781	0.839	Yes	Accept
NiSCI item 2fi How severe was the difficulty with bringing up phlegm?	79.1%*	61.3%*	≤0.519	≤0.622	0.674	0.692	Yes	Accept
NiSCI item 3 Overall, how severe were your COPD symptoms last night?	27.6%	13.0%	≤0.719	≤0.801	NA	NA	No	Accept but score separately
NiSCI item 7 How many puffs of your rescue medication did you take last night?	NA	NA	≤0.489	≤0.533	NA	NA	NA	Accept

Notes: *Indicates values that are above the threshold criteria that were set a priori to flag items for potential problems (ie, floor effect minimum response >30% and ceiling effect maximum response >30%).

Abbreviations: NiSCI, Nighttime Symptoms of COPD Instrument; CFA, confirmatory factor analysis; IRT, item response theory; COPD, chronic obstructive pulmonary disease; NA, not applicable.

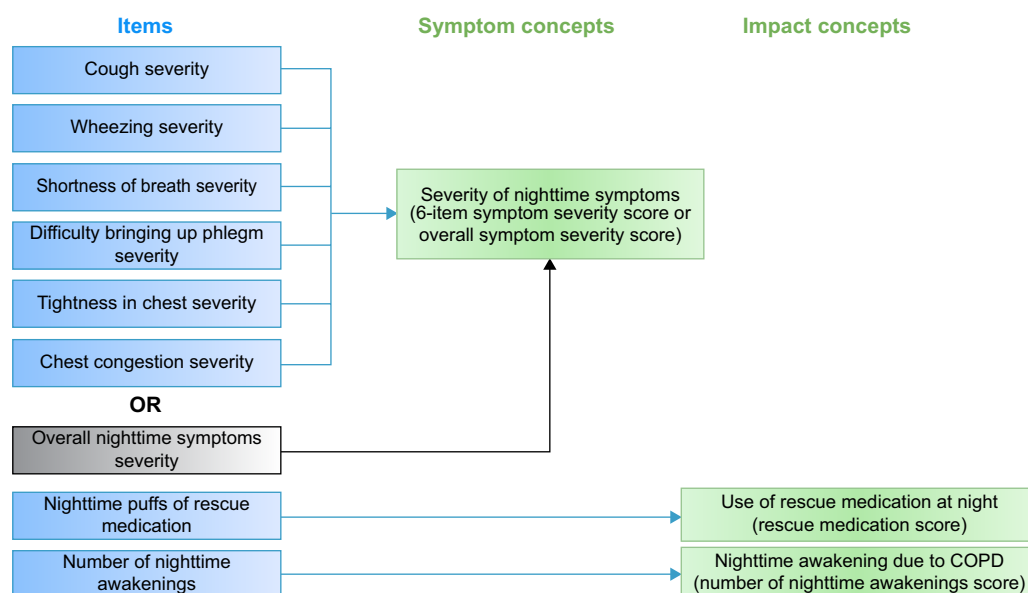


Figure 1 Final conceptual framework.

Abbreviation: COPD, chronic obstructive pulmonary disease.

Validity

Results of the analyses of convergent validity are presented in Table 4. Moderate-to-strong correlations (>0.4) were demonstrated for the NiSCI six-item symptom severity and overall symptom severity scores with the SGRQ (symptom and total) scores and E-RS total score. A stronger correlation was identified between the NiSCI six-item symptom severity score and the SGRQ symptom domain score (0.66) than with the other SGRQ domain scores (0.55–0.58). Significant correlations were also found between all the NiSCI scores and the E-RS total score (0.35–0.50). No significant correlations were found between the NiSCI scores and trough FEV_1 .

Known-groups validity

All NiSCI scores differentiated between known groups of SGRQ total and symptom domain scores (Figures 2 and 3, respectively) and between the E-RS total score groups

(Figure 4). Due to the small sample sizes of patients in GOLD stage I ($n=4$) and GOLD stage IV ($n=12$), the only meaningful pairwise comparison between GOLD stages was between GOLD stages II and III. The overall symptom severity and rescue medication scores differed in these two groups ($P<0.001$; Figure 5).

Discussion

Since COPD symptoms at night can have a significant impact on patients' quality of life, especially when symptoms disrupt sleep, nighttime COPD symptoms are important to evaluate in both clinical practice and clinical trials.¹ However, there is no standardized, reliable, and valid outcome measure that can be used to evaluate nighttime symptoms of COPD. While there is evidence about the content validity of the NiSCI,⁸ this paper presents the first evidence to support the scoring of the instrument and its psychometric properties.

Table 4 Correlation^a of NiSCI scores with SGRQ, E-RS total, and FEV_1 ^b at baseline week

	Six-item symptom severity score ^c	Overall symptoms severity score ^d	Number of nighttime awakenings score	Rescue medication score
SGRQ total score	0.58***	0.57***	0.45***	0.35***
SGRQ symptoms score	0.66***	0.60***	0.45***	0.39***
SGRQ impacts score	0.55***	0.53***	0.45***	0.35***
E-RS total score	0.76***	0.73***	0.50***	0.36***
FEV_1 (trough) ^b	−0.01	−0.04	0.01	−0.06

Notes: ^aSpearman rank order correlation coefficients: *** $P<0.0001$. ^bMorning predose value. ^cAverage score of six symptom severity items (cough, wheezing, shortness of breath, tightness in your chest, chest congestion, and difficulty bringing up phlegm). ^dSingle item on overall nighttime symptoms severity.

Abbreviations: NiSCI, Nighttime Symptoms of COPD Instrument; SGRQ, St George's Respiratory Questionnaire; E-RS, Exacerbations of Chronic Obstructive Pulmonary Disease Tool-Respiratory Symptoms; FEV_1 , forced expiratory volume in 1 second.

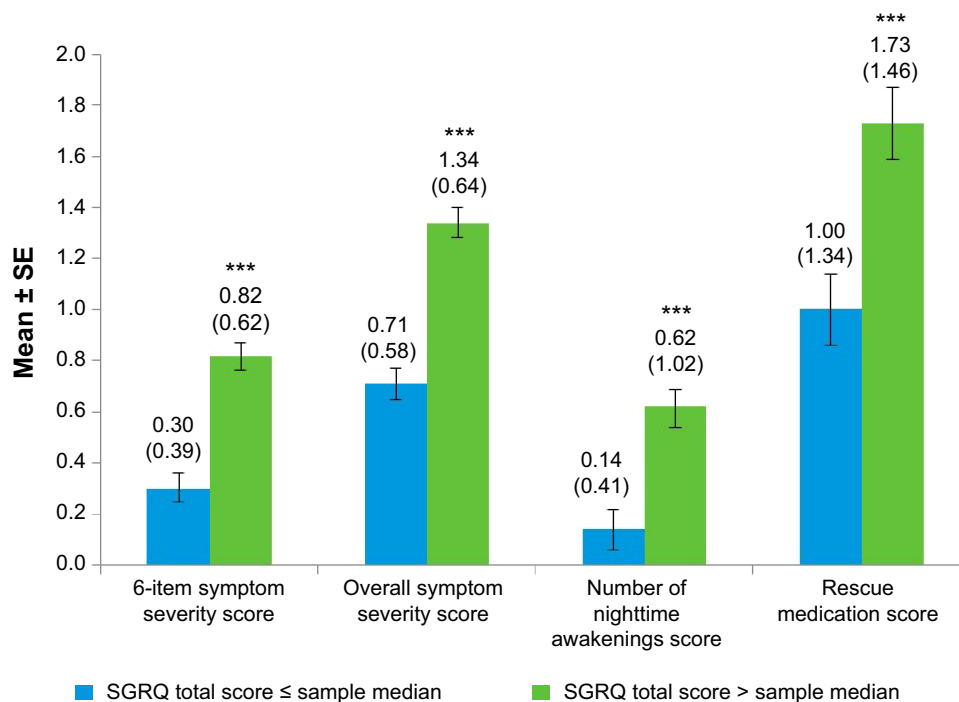


Figure 2 NiSCI scores by SGRQ total score at baseline week.

Notes: Values above bars indicate the mean (standard deviation) for each. Error bars represent standard error of the mean. Significance level for comparison: *** $P < 0.0001$; sample median 47.48.

Abbreviations: NiSCI, Nighttime Symptoms of COPD Instrument; SGRQ, St George's Respiratory Questionnaire; SE, standard error.

The results of this research confirm that the nine items in the NiSCI can be used to collect data about nighttime symptoms of COPD and to generate four scores: 1) the six-item symptom severity score, 2) the overall symptom severity score, 3) the rescue medication score, and 4) the number of nighttime awakenings score. CFA and Rasch model analysis confirmed the one-factor structure of the six-item symptom severity score based on the six COPD symptom severity items. The score exhibited good internal consistency and test-retest reliability. Test-retest reliability was also confirmed for the three other NiSCI scores. It is recommended that the four scores should be used separately, as there is no empirical evidence to support a total score.

Convergent validity was confirmed, with evidence of predicted correlation between the four NiSCI scores and the scores on other PRO instruments measuring similar concepts (ie, SGRQ total and symptom scores and E-RS total scores). As expected, none of the scores on the NiSCI was correlated with FEV_1 . Previous research has also shown that symptoms and lung function parameters are often not correlated in COPD.²⁵ This further highlights the importance of collecting data using PRO instruments (in addition to pulmonary function outcomes) in order to understand COPD patients' experiences, for monitoring patients, and for evaluating the outcomes of interventions.

All four NiSCI scores differentiated between groups that differed in their SGRQ total ($P < 0.0001$) and symptom scores ($P < 0.0001$) and E-RS total scores ($P < 0.0001$ for all except for rescue medication item, $P < 0.001$). Differences between known groups of patients defined as GOLD stage II or GOLD stage III could be observed on all four scores and were significant for the single-item overall symptom severity score ($P < 0.001$) and the single-item rescue medication score ($P < 0.001$). However, differences between other GOLD stages could not be examined, due to the small samples for these stages in the clinical trial data. It would be beneficial in future research to examine the differences between other GOLD severity stages.

The NiSCI also provides a simple measure of COPD symptom severity with the single item about overall symptom severity. Similar overall assessment of severity of symptoms has been used in both COPD and asthma; for example, use of short-acting beta-agonists is commonly assessed in COPD trials, and both number of puffs of albuterol and number of symptom-free days have been reported as outcomes.²³ The American Thoracic Society (ATS) statement on endpoints for evaluating asthma control recommends that for evaluating the number of "symptom-free days" in asthma, a general question about symptoms (rather than several questions about individual symptoms) is most appropriate.²⁴ Symptom-free

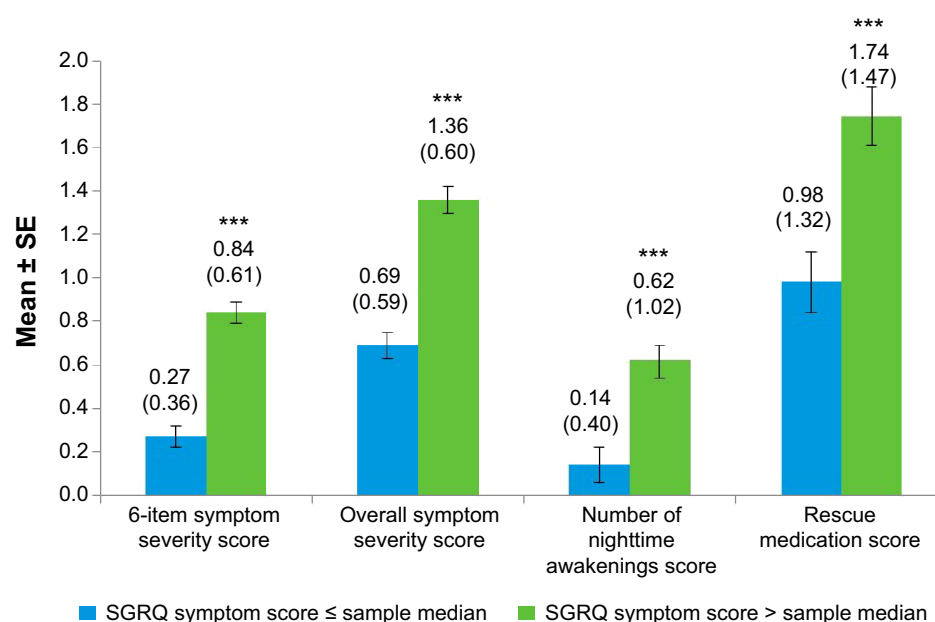


Figure 3 NiSCI scores by SGRQ symptom score at baseline week.

Notes: Values above bars indicate the mean (standard deviation) for each. Error bars represent standard error of the mean. Significance level for comparison: *** $P < 0.0001$; sample median 69.46.

Abbreviations: NiSCI, Nighttime Symptoms of COPD Instrument; SGRQ, St George's Respiratory Questionnaire; SE, standard error.

days may be an important end point for trials evaluating interventions for milder forms of COPD.

The qualitative research used to support the development of the NiSCI indicated that all six symptoms measured in the instrument were experienced by COPD patients. We suggest

that the six-item symptom severity and overall symptom severity scores together provide the most comprehensive picture of COPD patients' symptoms at nighttime and advise that, when possible, the two scores should both be used. If there are limitations to the study, the authors would

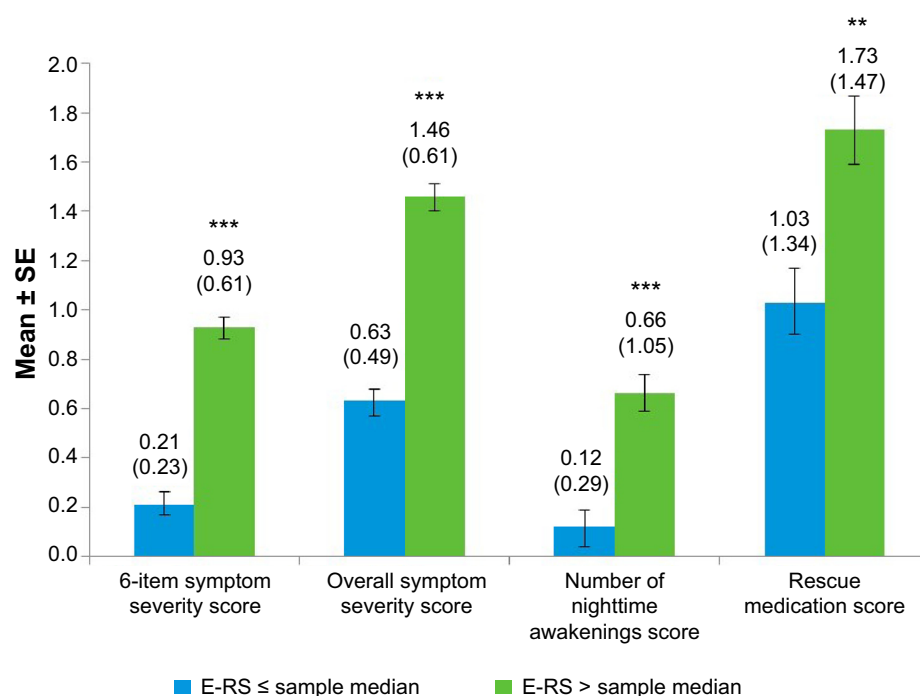


Figure 4 NiSCI scores by E-RS score at baseline week.

Notes: Values above bars indicate the mean (standard deviation) for each. Error bars represent standard error of the mean. Significance level for comparison: ** $P < 0.001$; *** $P < 0.0001$; sample median 11.

Abbreviations: NiSCI, Nighttime Symptoms of COPD Instrument; E-RS, Exacerbations of Chronic Obstructive Pulmonary Disease Tool-Respiratory Symptoms; SE, standard error.

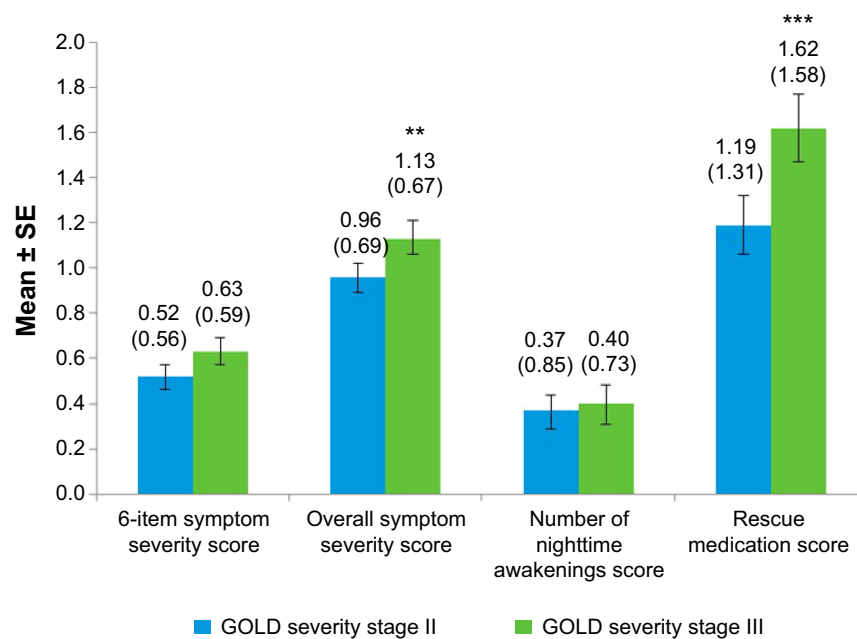


Figure 5 NiSCI scores by GOLD severity stages II and III at baseline week.

Notes: Values above bars indicate the mean (standard deviation) for each. Error bars represent standard error of the mean. Significance level for comparison: ** $P < 0.001$; *** $P < 0.0001$.

Abbreviations: NiSCI, Nighttime Symptoms of COPD Instrument; GOLD, Global Initiative for Chronic Obstructive Lung Disease; SE, standard error.

recommend using the six-item severity score, as it would provide a more precise measurement of change in the most relevant nighttime symptoms of COPD. If needed, due to study constraints (eg, patient burden), the overall score could be used on its own to measure symptom severity or to calculate symptom-free days. While there is evidence that the overall symptom severity score is robust, the properties of this single-item score were evaluated in the context of its application along with the six individual symptom severity questions in the daily diary. Further testing of this single item score on its own is recommended to confirm these findings.

Study limitation and directions for future research

This validation work was limited to a sample population taken from a clinical trial with a largely homogenous population, particularly in regards to race, as the majority (93%) of patients were white. One limitation to this study is that it did not include a more representative sample of the population. As a result, the validity of the NiSCI outside of this population is still uncertain and must be examined further.

In addition, the sample population was primarily made up of patients classified as GOLD stage II or III COPD. Sample sizes of patients classified as GOLD stage I or IV COPD were not large enough to be meaningfully analyzed. Therefore, future validation work with the NiSCI should aim to evaluate the measure's appropriateness among other

populations with varying degrees of COPD severity according to the GOLD criteria.

Previous research has suggested that there is a weak relationship between FEV_1 and COPD symptoms. The results of this study are in line with those findings. Nevertheless, because the sample is somewhat homogenous, it is possible among patients with very severe (versus mild) COPD that the relationship between FEV_1 and nighttime symptoms may differ. However, the sample was not large enough to test this question.

Further information is required on the responsiveness of the NiSCI, and minimally important differences should be established for each of the four scores. This additional information is particularly important for determining whether or not to include the NiSCI as an end point in an interventional trial.

Conclusion

The results of the assessment of the measurement properties suggest that the NiSCI is a reliable and valid instrument to evaluate nighttime COPD symptoms. The NiSCI can help provide data about patients' experience of nighttime symptoms of COPD, which other instruments do not currently measure. While the NiSCI was designed primarily to collect data to support labeling claims in clinical trials, the instrument may also be useful in clinical practice and other research studies, and can help in making decisions about treatment options. Using the instrument during clinical consultations

could help clinicians and patients make decisions, for example, about the timing of maintenance medication for treating nighttime symptoms. Additional work is currently being conducted to define responders and to evaluate the responsiveness of the NiSCI to interventions for COPD.

Acknowledgments

The authors would like to thank the AUGMENT COPD study investigators. Editorial assistance and technical writing were provided by Debika Chatterjea PhD and Mary Clare Kane PhD of Prescott Medical Communications Group (Chicago, IL, USA), funded by Forest Research Institute (Jersey City, NJ, USA), a subsidiary of Actavis plc.

Author contributions

All authors were involved in the conception, design, analysis, and interpretation of the PRO-specific psychometric analyses of this study, as well as the creation and critical review of the manuscript. All authors provided approval of the final manuscript and take full responsibility for the interpretation and integrity of the data.

Disclosure

This study was conducted by Evidera (London, UK) on behalf of the Forest Research Institute (FRI) (Jersey City, NJ, USA), a subsidiary of Actavis plc, who funded this work. MM was an employee of FRI at the time the study was conducted but is currently employed by Novo Nordisk. EZ, DT, and AH, as employees of Evidera, served as paid consultants to FRI during the conduct of this study and the development of this manuscript. BM has served on medical advisory boards and speakers bureaus and has consulted on a PRO for Forest Laboratories. BM has received travel expenses from FRI for abstract presentations at national and international meetings.

References

- Bhullar S, Phillips B. Sleep in COPD patients. *COPD*. 2005;2(3):355–361.
- Lange P, Marott JL, Vestbo J, Nordestgaard BG. Prevalence of nighttime dyspnoea in COPD and its implications for prognosis. *Eur Respir J*. 2014;43(6):1590–1598.
- Agusti A, Hedner J, Marin JM, Barbe F, Cazzola M, Rennard S. Nighttime symptoms: a forgotten dimension of COPD. *Eur Respir Rev*. 2011;20(121):183–194.
- Partridge MR, Karlsson N, Small IR. Patient insight into the impact of chronic obstructive pulmonary disease in the morning: an internet survey. *Curr Med Res Opin*. 2009;25(8):2043–2048.
- European Medicines Agency Respiratory Drafting Group. Guideline on clinical investigation of medicinal products in the treatment of chronic obstructive pulmonary disease (COPD). EMA/CHMP/483572/2012. Available from: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2012/08/WC500130880.pdf. Accessed December 8, 2014.
- US Food and Drug Administration. Guidance for industry on patient-reported outcome measures: use in medical product development to support labeling claims. *Fed Reg*. 2009;74:65132–65133.
- Stull DE, Leidy NK, Parasuraman B, Chassany O. Optimal recall periods for patient-reported outcomes: challenges and potential solutions. *Curr Med Res Opin*. 2009;25(4):929–942.
- Hareendran A, Palsgrove AC, Mocarski M, et al. The development of a patient-reported outcome measure for assessing nighttime symptoms of chronic obstructive pulmonary disease. *Health Qual Life Outcomes*. 2013;11:104.
- D'Urzo A, Rennard S, Kerwin E, Mergel V, Leselbaum A, Caracta C. Efficacy and safety of fixed-dose combinations of acclidinium bromide/formoterol fumarate: the 24-week, randomized, placebo-controlled AUGMENT COPD study. *Respir Res*. 2014;15(1):123.
- Fayers P, Hays R. *Assessing Quality of Life in Clinical Trials*. New York: Oxford University Press; 2005.
- Streiner DL NG. *Health Measurement Scales: A Practical Guide to their Development and Use*. 4th ed. New York: Oxford University Press; 2008.
- European Medicines Agency (EMA) Committee for Medicinal Products for Human Use (CHMP). Reflection paper on the regulatory guidance for the use of health-related quality of life (HRQL) measures in the evaluation of medicinal products. Available from: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003637.pdf. Accessed December 8, 2014.
- Mahler DA. *Dyspnea: Mechanisms, Measurement, and Management*. 3rd ed. Boca Raton, FL: CRC Press; 2014.
- Leidy NK, Sexton CC, Jones PW, et al. Measuring respiratory symptoms in clinical trials of COPD: reliability and validity of a daily diary. *Thorax*. 2014;69(5):424–430.
- Jones PW, Quirk FH, Baveystock CM, Littlejohns P. A self-complete measure of health status for chronic airflow limitation: the St George's Respiratory Questionnaire. *Amer Rev Respir Dis*. 1992;145(6):1321–1327.
- Mahler DA, Wells CK. Evaluation of clinical methods for rating dyspnea. *Chest*. 1988;93(3):580–586.
- Leidy NK, Wilcox TK, Jones PW, et al. Development of the EXacerbations of Chronic Obstructive Pulmonary Disease Tool (EXACT): a patient-reported outcome (PRO) measure. *Value Health*. 2010;13(8):965–975.
- Jones PW, Quirk FH, Baveystock CM. The St George's Respiratory Questionnaire. *Respir Med*. 1991;85 Suppl. B:25–31; discussion 33–27.
- Mahler DA, Weinberg DH, Wells CK, Feinstein AR. The measurement of dyspnea. Contents, interobserver agreement, and physiologic correlates of two new clinical indexes. *Chest*. 1984;85(6):751–758.
- Junghard O, Lauritsen K, Talley N, Wiklund I. Validation of seven graded diary cards for severity of dyspeptic symptoms in patients with non ulcer dyspepsia. *Eur J Surg*. 1998;164(Suppl 583):106–111.
- Bentler PM. Comparative fit indexes in structural models. *Psychol Bul*. 1990;107(2):238–246.
- LJC. Coefficient alpha and the internal structure of tests. *Psychometrika*. 1951;16:297–334.
- Nunnally JC BI. *Psychometric Theory*. 3rd ed. New York: McGraw-Hill; 1994.
- Reddell HK, Taylor DR, Bateman ED, et al. An official American Thoracic Society/European Respiratory Society Statement: asthma control and exacerbations: standardizing endpoints for clinical asthma trials and clinical practice. *Am J Respir Crit Care Med*. 2009;180(1):59–99.
- Nishimura K, Izumi T, Tsukino M, Oga T. Dyspnea is a better predictor of 5-year survival than airway obstruction in patients with COPD. *Chest*. 2002;121(5):1434–1440.
- Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates; 1988.
- Stewart AL, Hays RD, Ware JE. *Methods of Validating MOS Health Measures*. Durham, NC: Duke University Press; 1992.
- Hays RD, Revicki D. *Reliability and Validity (Including Responsiveness)*. 2nd ed. New York, NY: Oxford University Press; 2005.

International Journal of COPD**Publish your work in this journal**

The International Journal of COPD is an international, peer-reviewed journal of therapeutics and pharmacology focusing on concise rapid reporting of clinical studies and reviews in COPD. Special focus is given to the pathophysiological processes underlying the disease, intervention programs, patient focused education, and self management protocols.

Submit your manuscript here: <http://www.dovepress.com/international-journal-of-chronic-obstructive-pulmonary-disease-journal>

This journal is indexed on PubMed Central, MedLine and CAS. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Dovepress