ORIGINAL ARTICLE

Measuring respiratory symptoms in clinical trials of COPD: reliability and validity of a daily diary

N K Leidy,[1] C C Sexton,[1] P W Jones,[2] S M Notte,[1] B U Monz,[3] L Nelsen,[4] M Goldman,[5] L T Murray,[1] S Sethi[6]

ABSTRACT

Background Although respiratory symptoms are characteristic features of COPD, there is no standardised method for quantifying their severity in stable disease.
Objective To evaluate the EXACT-Respiratory Symptom (E-RS) measure, a daily diary comprising 11 of the 14 items in the Exacerbations of Chronic Pulmonary Disease Tool (EXACT).
Methods Qualitative: patient focus group and interviews to address content validity. Quantitative: secondary data analyses to test reliability and validity.
Results Qualitative: n=84; mean (SD) age 65 (10) years, FEV_1 1.2(0.4) L; 44% male. Subject descriptions of their respiratory symptoms were consistent with E-RS content and structure. Quantitative: n=188; mean (SD) age 66 (10) years, FEV_1 1.2(0.5) L; 50% male. Factor analysis (FA) showed 3 subscales: RS-Breathlessness, RS-Cough & Sputum, and RS-Chest Symptoms; second-order FA supported a general factor and total score. Reliability (total and subscales): 0.88, 0.86, 0.73, 0.81; 2-day test-retest ICC: 0.90, 0.86, 0.87, 0.82, respectively. Validity: Total scores correlated significantly ($p < 0.0001$) with SGRQ Total (r=0.75), Symptoms (r=0.66), Activity (r=0.57), Impact (r=0.70) scores; subscale correlations were also significant (r=0.26, $p < 0.05$ (RS-Chest Symptoms with Activity) to r=0.69, $p < 0.0001$ (RS-Cough & Sputum with Symptoms). RS-Breathlessness correlated with rescue medication use (r=0.32, $p < 0.0001$), clinician-reported mMRC (r=0.33, $p < 0.0001$), and FEV_1% predicted (r=-0.17, $p < 0.05$). E-RS scores differentiated groups based on chronic bronchitis diagnosis ($p < 0.01$–0.001), smoking status ($p < 0.05$–0.001), and rescue medication use ($p < 0.05$–0.0001).
Conclusions Results suggest the RS-Total is a reliable and valid instrument for evaluating respiratory symptom severity in stable COPD. Further study of sensitivity to change is warranted.

INTRODUCTION

COPD is a progressive disease characterised by persistent airflow limitation with varying degrees of airway wall narrowing, inflammation and emphysema. Respiratory symptoms, including breathlessness, cough and sputum production, are characteristic features of the disease and have significant adverse effects on patient functioning and quality of life.[1–4] Although spirometric measures are useful for diagnosis and evaluating change in lung function, they do not capture symptom severity or variability; weak correlations between lung function and symptoms show they cannot be used

Key messages

What is the key question?
► Is there a standardised, reliable and valid diary to evaluate daily respiratory symptoms in clinical studies of stable COPD?

What is the bottom line?
► Yes, the E-RS.

Why read on?
► This paper presents methods and results of two studies showing that RS-Total and subscale scores (RS-Breathlessness, RS-Cough & Sputum, RS-Chest Symptoms) have content validity and are reliable and valid in stable patients with COPD.

as proxies for one another.[5–9] Because respiratory symptoms are patient-experienced, it is important to measure this outcome directly, through a patient-reported outcome (PRO) instrument.

To date, there is no standardised method for evaluating respiratory symptoms of stable COPD with a documented development programme consistent with good research practices[10 11] and the Food and Drug Administration (FDA) PRO Guidance.[12] Dyspnoea is the most frequently assessed symptom, often measured with the Baseline/Transition Dyspnea Index,[13] or modified Medical Research Council Dyspnea Scale (mMRC). Health-status measures with symptom components include the St George's Respiratory Questionnaire (SGRQ),[14] Chronic Respiratory Disease Questionnaire,[15] Clinical COPD Questionnaire,[16] and COPD Assessment Test.[17] These questionnaires are completed by subjects intermittently using varied recall periods (eg, since the last visit, past month, previous week, or now).

An alternative to periodic symptom assessment is a patient-completed diary, capturing symptoms each day. This approach is necessary for studies examining the temporal and dynamic nature of respiratory symptoms, including their relationship to other variables such as activity, stress, environmental conditions and rescue medication use. In clinical trials, diaries can be used to evaluate time to symptomatic improvement and/or magnitude and persistence of change over the treatment period. Regulatory agencies have expressed interest

in electronic daily symptom assessments in pharmaceutical trials, with their reduced recall bias and technology enabled compliance monitoring.[12] [18] One diary, the Breathlessness, Cough and Sputum Scale, has shown evidence of reliability, validity and responsiveness.[19] [20] However, it is limited to three items, and its development was not based on qualitative data from the target population, raising concerns about content validity.[10–12]

The Exacerbations of Chronic Pulmonary Disease Tool (EXACT)-Respiratory Symptoms (RS) (E-RS) was designed to meet the need for a standardised respiratory symptom diary with a development history consistent with good research practices and FDA PRO requirements. The E-RS uses 11 respiratory symptom items from the 14-item EXACT,[21–24] offering efficiencies for investigators and subjects by permitting two validated uses for one diary: (1) assessment of COPD exacerbations using the EXACT total score[21–24]; (2) quantification of respiratory symptoms in stable COPD using RS-Total and subscale scores.

This paper describes the methods and results of research to assess content validity, reliability and validity of the E-RS in patients with stable COPD.

MATERIALS AND METHODS
The research was conducted in two phases. Phase I addressed content validity through qualitative research methods. Phase II tested score reliability and validity using an existing dataset.[23] All data were gathered in accordance with the amended Declaration of Helsinki, with study protocols approved by an independent institutional review board (IRB) and all participants providing written informed consent prior to data collection (Essex IRB, ID# A2-3864, A2-3864B; Ethical Review Committee, ID# 472-05-09).

In each phase, study participants were recruited from pulmonary and primary care clinics across the USA using inclusion criteria similar to those used in pharmaceutical trials: >40 years of age; diagnosis of stable COPD; emphysema or chronic bronchitis; and ≥10 pack-year smoking history. Exclusion criteria: medical diagnosis of asthma without postbronchodilator airway obstruction; acute congestive heart failure or unstable angina, bronchiectasis, lung cancer, or tuberculosis; or treatment for respiratory infection or pneumonia within the past 60 days. Phase-specific inclusion/exclusion criteria and methods are described below. Following enrolment and consent, each site provided clinical information related to the participant's diagnosis, pulmonary function and clinician rating of disease severity.

Phase I: qualitative: content validity
A two-step process was used to assess and document the extent to which the items comprising the E-RS adequately and accurately reflect respiratory symptoms of COPD in a stable state. Methods are outlined below, with details provided in the online supplementary appendix.

Stage 1: secondary analyses of existing qualitative data
Qualitative analyses were performed on data gathered during the development of the EXACT,[21] that is, data from focus groups and interviews with patients with COPD and a history of a clinic visit or hospitalisation for exacerbation in the previous 6 months (n=63). The purpose of the original study was to characterise COPD exacerbations from the patient's perspective. The data included participant descriptions of the nature and severity of symptoms during a stable state presented as part of their characterisations of COPD and/or to facilitate descriptions of exacerbations, that is, relative to their stable state.

Stage 2: new focus groups
Additional focus groups were conducted in a new sample of clinically stable patients (n=21), that is, exacerbation-free for 12 months. The purpose of Stage 2 was to determine if there were new insights or information related to respiratory symptoms in stable COPD not discussed by patients with an exacerbation history and to ensure that saturation had been reached.

Each focus group was led by an experienced study team member using a semistructured discussion guide to elicit information on patient perspectives of respiratory symptoms. To characterise the sample, participants completed the mMRC Dyspnea Scale,[25] the SGRQ for COPD (SGRQ-C),[26] and a sociodemographic questionnaire.

Analyses
Descriptive statistics were used to characterise the sample. Atlas.ti 5.0 facilitated thematic analyses of the qualitative data. For each stage of analysis, symptomatic themes were summarised in a saturation grid and mapped to the respiratory symptom items in the EXACT to determine the extent to which this subset of items, named the E-RS, would adequately capture respiratory symptoms in stable disease. Patient understanding of instructions, items, and response categories was addressed using cognitive interviewing methodology during EXACT development.[21]

Phase II: quantitative: score reliability and validity
Secondary analyses were performed on a subset of data from a previously published prospective observational study used to develop the EXACT.[23] Specifically, data from the stable control sample (n=188) were used. These patients had no history of treatment for exacerbation in the preceding 60 days and were considered clinically stable on enrolment. Subjects were recruited through clinical sites in over 20 states across the USA.

Measures
All participants completed the SGRQ-C and sociodemographic forms during the enrolment clinic visit; site staff provided clinician ratings of patient dyspnoea (mMRC) and disease severity, and results of the most recent stable-state spirometry. Each subject completed an eDiary for the ensuing 7 days that included EXACT candidate items, rescue medication use and global ratings of change.[23]

Statistical analyses
Analyses were prespecified in a statistical analysis plan completed following Phase I. Because the intent was to evaluate the performance properties of the E-RS as a daily diary, analyses were conducted with data from Day 1, the same day clinic data were gathered, unless otherwise specified. Item-level analyses included measures of central tendency, floor and ceiling effects, item-total correlations, item response frequencies and interitem correlations. Confirmatory (CFA), exploratory (EFA), and second-order factor analyses were performed to evaluate the structure of the measure and develop the scoring algorithms.

Internal consistency was assessed using Cronbach's α, with a target value of greater than 0.70, to make the instrument suitable for use in clinical trials.[27] Two-day test-retest reliability was evaluated for consecutive days (Days 1–2; 2–3; 3–4, etc.), with data from patients reporting no change in lung condition on the daily global assessment question, and Days 1 and 7 in all

subjects, with the latter assuming symptomatic stability across these two observations (no confirmatory global assessment). Intraclass correlation coefficients (ICC), paired t tests and effect sizes were computed.

Validity was examined by correlating E-RS scores with SGRQ total and domain scores, rescue medication use, clinician rating of dyspnoea on the mMRC, and forced expiratory volume in 1 s (FEV_1) % predicted. Known-groups validity was tested using Student t test, comparing E-RS scores of those with and without a medical diagnosis of chronic bronchitis (clinician-reported), current and former smokers (self-report), and those using no rescue medication versus three or more puffs (Day 1; self-report). Scores were also compared across clinician-rated disease severity, hypothesising weak relationships given the multidimensional nature of the clinician's assessment. To take this into account, analyses were performed with and without controlling for age, comorbidity status (≤ 1 or ≥ 2) and FEV_1.

Statistical analyses were performed using SAS statistical software V.9.1 (Cary, North Carolina, USA).

RESULTS
Samples
Sample demographic and clinical characteristics by phase and stage are shown in table 1.

Phase I: content validity
Qualitative analyses identified three categories of respiratory symptoms patients experience when stable: breathlessness, cough and sputum, and chest symptoms. Representative quotations for each symptom category and the interrelationship/co-occurrence of symptoms and the saturation grid are provided in online supplementary appendix tables S1 and S2.

Participants spoke of being 'breathless', 'short of breath', and having difficulty breathing, with severity levels that varied day to day. Cough and sputum were generally discussed together; most patients were unable to make it through the day without coughing. Sputum was characterised in terms of quantity and thickness. Most participants cited difficulty coughing up sputum or phlegm. Chest symptoms included congestion, tightness and discomfort, which also varied day to day. Participants often presented their respiratory symptom descriptions as inter-related experiences, describing how they related to and affected one another.

Content of the final instrument, in the context of the 14-item EXACT, is shown in online supplementary appendix table S3.

Phase II: reliability and validity
Item and factor analyses and scoring algorithm
Participants used the full range of response options, with no missing data and minimal floor and ceiling effects. The CFA

Table 1 Sample demographic and clinical characteristics by phase

| Characteristic | Phase I: content validity | | Phase II: score reliability & validity |
	Stage 1: secondary analysis of EXACT qualitative data* (n=63)	Stage 2: new focus groups† (n=21)	(n=188)
Age mean (SD)	65.2 (10.4)	67.8 (8.7)	65.8 (9.7)
Male (%)	28 (44.4)	14 (66.7)	91 (48.4)
Smoking status, n (%)			
Current	20 (31.7)	9 (42.9)	58 (30.9)
Former	42 (66.7)	12 (57.1)	130 (69.1)
Duration of disease in years, mean (SD)	7.9 (6.2)	8.9 (7.3)	7.5 (7.0)
FEV_1	1.16 (0.4)	1.3 (0.5)	1.3 (0.6)
FEV_1% predicted, mean (SD)	45.8 (16.1)	47.9 (16.6)	51.2 (19.6)
mMRC dyspnoea score, n (%)			
0 None	1 (1.6)	5 (23.8)	15 (8.0)
1 Hurrying uphill	8 (12.7)	8 (38.1)	62 (33.0)
2 Walks slower on level	18 (28.6)	3 (14.3)	56 (29.8)
3 Stops after walking	27 (42.9)	4 (19.0)	49 (26.1)
4 Too breathless	6 (9.5)	1 (4.8)	6 (3.2)
Missing	3 (4.8)	0 (0.0)	0 (0.0)
SGRQ total‡	56.3 (18.6)	48.0 (16.0)	50.8 (19.9)
Symptoms	64.3 (23.3)	62.0 (23.1)	60.8 (21.2)
Activity	72.6 (19.7)	61.0 (21.0)	69.7 (23.5)
Impact	45.7 (21.3)	35.8 (16.4)	36.8 (22.3)
Patient-reported comorbidity, n (%)§			
Hypertension	21 (33.3)	10 (47.6)	73 (38.8)
Diabetes	12 (19.0)	4 (19.0)	24 (12.8)
Heart disease	10 (15.9)	1 (4.8)	35 (18.6)
Chronic sinusitis	5 (7.9)	2 (9.5)	13 (6.9)
Other	16 (25.4)	4 (19.0)	12 (6.4)

*Three focus groups, twenty-five 1:1 interviews, two 2:1 interviews among stable COPD patients.
†Four new focus groups.
‡Phase I, Stage 1: n=54–61; Phase II: n=181–188.
§Not mutually exclusive.
EXACT, Exacerbations of Chronic Pulmonary Disease Tool; FEV_1, forced expiratory volume in 1 s; GOLD, Global Initiative on Obstructive Lung Disease; mMRC, modified Medical Research Council; SGRQ, St. George's Respiratory Questionnaire.

Table 2 Exploratory factor analysis: promax factor loading*

Item number	Factor 3	Factor 2	Factor 1
RS-Breathlessness			
Q7—Breathless today	**0.777**	0.040	0.056
Q8—Describe how breathless	**0.747**	0.122	−0.170
Q9—Short of breath—personal care	**0.749**	0.103	−0.053
Q10—Short of breath—indoor	**0.750**	−0.079	0.081
Q11—Short of breath—outdoor	**0.644**	−0.154	0.196
RS-Cough & Sputum			
Q2—How often cough	0.017	**0.717**	0.156
Q3—How much mucus when cough	0.004	**0.850**	−0.082
Q4—Difficulty bringing up mucus	0.071	**0.447**	0.284
RS-Chest Symptoms			
Q1—Chest feel congested	0.006	0.290	**0.575**
Q5—Have chest discomfort	−0.029	−0.043	**0.876**
Q6—Chest feels tight	0.113	0.069	**0.622**

n=188.
In bold: Factor 3: RS-Breathlessness; Factor 2: RS-Cough & Sputum, Factor 1; RS-Chest Symptoms.
For each item, response scale scores range from 5 to 6, with higher scores reflecting greater severity.
*χ^2=44.24 (DF=25), p=0.010; RMSEA=0.064, RMSR=0.027.
RMSEA, root mean square error of approximation; RMSR, root mean square residual; RS, respiratory symptoms.

confirmatory fit index (CFI) was 0.75, less than the 0.95 prespecified as a good fit for unidimensionality, thereby precluding the use of Rasch analysis for the total score. EFA showed a three-factor solution (table 2), indicating that three respiratory symptom subscales comprise the measure. The second-order factor model fit the data very well (CFI=0.96), with standardised coefficients between the items and respiratory symptom factors, and between the respiratory symptom factors and the general factor, all greater than 0.60 (range 0.68 to 0.85 and 0.65 to 0.95, respectively) (see figure 1). The 0.94 correlation between general factor and RS-Total scores provided further support for an empirical general factor governing the three E-RS factors.

The scoring algorithms for the E-RS yield a total score and three subscale scores, with higher scores on these ordinal-level scales indicating more severe symptoms. Item-level scores range from 4 to 5 points (0 to 3 or 0 to 4), which are summed to yield total and subscale scores. E-RS scores are calculated for each day the diary is completed, and may be aggregated or summarised in a manner consistent with the study purpose and design.

Descriptive statistics, floor and ceiling effects, total and subscale intercorrelations, and sample patient-level figures for RS-Total scores over 7 days are provided in online supplementary appendix tables S4a and b and appendix figure S1.

Reliability
Reliability parameters are presented in table 3; reproducibility estimates for Days 1 to 2 are shown, with the remaining consecutive-day analyses appearing in online supplementary appendix table S5.

Validity
Correlations between E-RS scores and alternative assessments of respiratory symptom severity (concurrent validity), related constructs (convergent validity), and weakly related constructs (divergent validity) are provided in table 4.

RS-Total and subscale scores (RS-Breathlessness, RS-Cough & Sputum, and RS-Chest Symptoms) differentiated those with and without a diagnosis of chronic bronchitis, current and ex-smokers, and rescue medication free versus ≥ three puffs (table 5). The univariate models of clinician-rated disease severity were not significant, but the multivariate models for RS-Total (p < 0.05), RS-Breathlessness (p < 0.05), and RS-Cough & Sputum scores (p < 0.01) were significant (see online supplementary appendix table S6).

DISCUSSION
Although respiratory symptoms play a key role in the diagnosis, assessment and management of patients with COPD and symptom relief is an important target of therapy, there is no standardised, reliable and valid daily diary for evaluating this
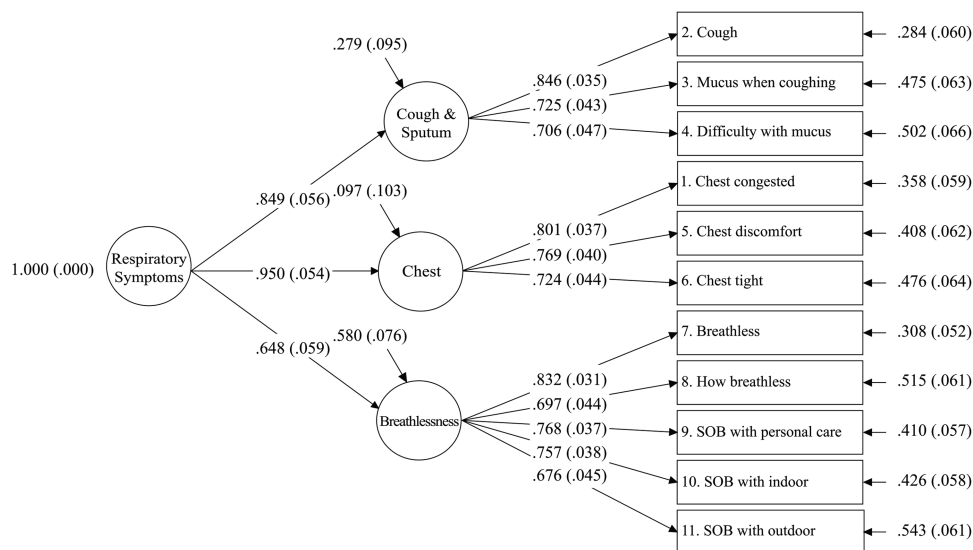

Figure 1 Higher-Order Factor Model for the E-RS. Confirmatory Fit Index (CFI)=0.958. Root Mean Square Error of Approximation (RMSEA)=0.073 (90% CI 0.050 to 0.096). Standardised Room Mean Square Residual (SRMR)=0.043. The E-RS is a derivative instrument, using the 11 respiratory symptom items from the 14-item EXACT. SOB, shortness of breath.

Table 3 E-RS reliability: internal consistency and test-retest reproducibility

| E-RS* scale (# items) | Internal consistency (n=188) | Reproducibility (Day 1 and Day 2) (n=76)† | | | | | |
	Cronbach's α	ICC‡	Day 1 mean (SD)	Day 2 mean (SD)	Difference (SD)	p Value§	Effect size
RS-Total Score (11)	0.88	0.90	9.62 (6.24)	9.66 (6.36)	−0.04 (2.88)	0.91	−0.01
RS-Breathlessness (5)	0.86	0.86	5.07 (3.65)	4.87 (3.63)	0.20 (1.95)	0.38	0.05
RS-Cough & Sputum (3)	0.73	0.87	2.82 (2.30)	2.88 (2.18)	−0.07 (1.15)	0.62	−0.03
RS-Chest Symptoms (3)	0.81	0.82	1.74 (1.86)	1.91 (2.07)	−0.17 (1.18)	0.21	−0.09
		Reproducibility (Day 1 and Day 7) (n=171)¶					
RS-Total Score		0.73	11.9 (6.77)	12.1 (7.22)	−0.15 (5.14)	0.70	−0.02
RS-Breathlessness		0.71	5.95 (3.75)	6.09 (4.13)	−0.15 (3.00)	0.53	−0.04
RS-Cough & Sputum		0.69	3.51 (2.28)	3.42 (1.96)	0.09 (1.67)	0.49	0.04
RS-Chest Symptoms		0.62	2.45 (2.11)	2.54 (2.27)	−0.09 (1.92)	0.52	−0.04

*Ranges: RS-Total 0 to 40; RS-Breathlessness 0 to 17; RS-Cough & Sputum 0 to 11; RS-Chest Symptoms 0 to 12; higher scores=more severe symptoms.
†Intraclass correlation coefficient.
‡Participants reporting 'no change' between the two consecutive days.
§p value from paired t test.
¶Participants completing the diary at Day 1 and Day 7.
E-RS, Exacerbations of Chronic Pulmonary Disease Tool-Respiratory Symptoms; ICC, intraclass correlation coefficient.

outcome in natural history studies and clinical trials. This paper presents the first evidence of the validity and reliability of the E-RS to meet this need.

Phase I of this work addressed content validity.[10] [28] Participant descriptions of their breathlessness, cough and sputum were consistent with the literature[3] [5] [8] [9] [29] and the content and structure of the E-RS. Of particular note were descriptions of chest symptoms (chest congestion, discomfort and tightness), a symptom set not measured with existing questionnaires.[14–17] [19]

Participant descriptions of symptom variability suggest an unstable component to 'stable' COPD, consistent with findings reported by Kessler et al[29] in severe patients. Of the symptomatic patients participating in this pan-European observational study (70% of the 2441), most (63%) experienced symptom variability, with over half indicating variance throughout the week (54%) or across seasons (60%). The most variable symptoms were breathlessness and chest tightness; variability in breathlessness was associated with a history of two or more exacerbations the prior year and greater adverse impact on daily activity.[29] These results suggest respiratory symptoms in COPD may not be as stable as previously believed,[30] and that further research on day-to-day variability is needed.

It is important to note that the E-RS is administered in the evening prior to bedtime with respondents rating their symptoms as they reflect back on the day. This method is efficient and less burdensome than twice-daily assessments. However, this approach may be less precise for those interested in characterising and tracking nighttime or morning symptoms, specifically.[29] [31] Studies could be performed to evaluate the added precision of administering the E-RS twice-daily, with the corresponding adjustment in recall period, or using a separate morning diary for this purpose.

Participants described breathlessness, cough, sputum and chest congestion as co-occurring and interacting, suggesting a respiratory symptom complex in COPD that can be captured through a total score, representing the overall severity of this symptom complex, and subscale scores capturing the three types of respiratory symptoms. This was supported quantitatively in the second-order factor structure and strong interscale correlations, internal consistency reliability levels and validity metrics. This measurement structure permits step-down hypothesis testing, with respiratory symptom severity overall tested first, followed by tests for breathlessness, cough and sputum, and chest symptoms.

Internal consistency estimates for RS-Total, RS-Breathlessness, and RS-Chest Symptom scores exceeded the conservative 0.80 standard, indicating a high degree of precision with low measurement error. Score reproducibility over two consecutive days in patients reporting no change was very high. The lower estimates over a six-day interval (Day 1 and Day 7) are consistent with the known symptom variability day to day in stable patients. This finding is pertinent given regulatory authority interest in the use of daily symptom assessments[12] and the need for further study of symptom temporal-severity dynamics.

Table 4 Validity: correlation with related constructs and alternative measures

| Measure | n* | E-RS scale† | | | |
		Total score	Breathlessness	Cough & Sputum	Chest Symptoms
SGRQ total	181	**0.75*****	**0.69*****	0.58***	0.52***
Symptoms	188	**0.66*****	0.44***	**0.69*****	**0.60*****
Activity	184	0.57***	**0.68*****	0.31***	0.26***
Impact	181	**0.70*****	0.61***	0.55***	0.54***
mMRC	188	0.33***	**0.38*****	0.24**	0.16*
Rescue medication use	188	0.32***	**0.34*****	0.26**	0.17*
FEV₁% predicted	188	−0.10	**−0.17***	−0.05	0.01

†Sample sizes are a function of missing item-level data.
‡Day 1, Spearman's rank-order correlation.
Bold coefficients identify related constructs with strongest expected value.
*p<0.05; **p<0.001; ***p<0.0001.
E-RS, Exacerbations of Chronic Pulmonary Disease Tool-Respiratory Symptoms; FEV₁, forced expiratory volume in 1 s; mMRC, modified Medical Research Council; SGRQ, St George's Respiratory Questionnaire.

Table 5 Known-groups validity: E-RS Mean (SD) scores by chronic bronchitis diagnosis, smoking status, and rescue medication use

E-RS scores*	Chronic bronchitis			Smoking status			Rescue medication use†		
	Yes (n=51)	No (n=137)	p Value‡	Current smoker (n=58)	Ex-smoker (n=130)	p Value‡	3 or more puffs (n=58)	No (n=94)	p Value‡
RS-Total	15.0 (6.2)	10.8 (6.6)	0.0001	14.8 (6.8)	10.6 (6.3)	<0.0001	14.7 (6.1)	9.9 (6.3)	<0.0001
RS-Breathlessness	7.3 (3.4)	5.4 (3.7)	0.0018	6.9 (3.7)	5.5 (3.6)	0.0183	7.6 (3.4)	4.7 (3.4)	<0.0001
RS-Cough & Sputum	4.5 (2.1)	3.1 (2.2)	0.0002	4.7 (2.0)	3.0 (2.2)	<0.0001	4.3 (2.0)	3.0 (2.3)	0.0009
RS-Chest Symptoms	3.2 (2.3)	2.2 (2.0)	0.0059	3.2 (2.4)	2.1 (1.9)	0.0012	2.9 (2.1)	2.2 (2.1)	0.0453

*RS-Total and RS-domain score ranges are as follows with the higher values indicating greater severity of respiratory symptoms: RS-Total scores range from 0 to 40; RS-Breathlessness scores range from 0 to 17; RS-Cough & Sputum scores range from 0 to 11; and RS-Chest Symptoms scores range from 0 to 12.
†Study Day 1.
‡Student t test.
E-RS, Exacerbations of Chronic Pulmonary Disease Tool-Respiratory Symptoms.

The magnitude and pattern of correlations between health status (SGRQ), dyspnoea (mMRC), rescue medication use, and $FEV_1\%$ and E-RS subscales were consistent with score validity. The chronic bronchitic phenotype and current smokers were more symptomatic,[32–34] while subjects who reported no rescue medication use on Day 1 also reported significantly less severe symptoms, with the strongest effect observed in the RS-Breathlessness scale. Although one might speculate that the relationship between symptom severity and the clinician's global assessment of COPD severity would be stronger, the modest relationship had improved power when confounding factors were controlled. This supports the differential and complementary roles played by direct symptom assessment from the patient and the clinician's integrated assessment of the patient's COPD, with the latter including a clinical appraisal of symptoms, spirometry, physical assessment, exacerbation history, treatment history, comorbidity and general health, among other factors.

Although the dataset did not permit an evaluation of sensitivity to change, results of known-groups analyses offer preliminary insight into score interpretation. For example, the difference in RS-Total scores between symptomatic and less symptomatic patients (table 5) was approximately 4 on the 40-point scale. Effect sizes were very large (>0.60), indicating that 4 points may be substantially greater than a 'minimally clinically important difference' (MCID). Applying a commonly used distribution-based method for estimating the MCID (0.5 SD of the sample mean; online supplementary appendix table S4a), estimates would be: RS-Total: 3.35; RS-Breathlessness: 1.85; RS-Cough & Sputum: 1.15; RS-Chest Symptoms: 1.05. Until anchor-based estimates across multiple samples are available, these values should be considered preliminary since they are probably higher than the true MCID.

Although respiratory symptoms may be assessed periodically as part of existing COPD-specific health status questionnaires, there are several potential advantages of daily assessments. First, daily scores reduce recall bias and provide a prospective, daily accounting of symptom severity. Second, this approach yields information on day-to-day variability. Third, daily data offer analytical flexibility, with methods for evaluation over time determined by the study purpose. Advantages of the E-RS specifically include its embedded position within the EXACT, providing data on exacerbations and respiratory symptoms simultaneously, with no additional patient burden. The EXACT was developed as an eDiary, with short, easy-to-read questions and recommendations for formatting and reminders to enhance compliance. Compliance rates were high (94%) in this short-term validation study, and have exceeded 88% across several 3–6 month clinical trials,[24] suggesting eDiaries are feasible in

this patient population. Widespread smart phone access may facilitate future testing and use of eDiaries in natural history studies or clinical practice.[35]

An important question is the suitability of the E-RS for international studies. During the development of the parent instrument, international content and translation experts served on the advisory panel. To date, the EXACT has been translated into more than 50 languages with cognitive interviews conducted with COPD patients in the target countries to assure cultural and linguistic equivalence. Although the EXACT has performed well in international trials,[24] an evaluation of E-RS score reliability, validity and responsiveness in international settings is warranted.

CONCLUSIONS

The E-RS was designed to assess daily respiratory symptoms in clinical studies of COPD. Results suggest the instrument is content valid with quantitative evidence of score reliability and validity. Further research on the performance properties of E-RS scores in new samples and its sensitivity to change, including MCID estimation and development of responder definitions, are warranted.

REFERENCES

1. Calverley PM, Georgopopoulos D. Chronic obstructive pulmonary disease: symptoms and signs. *Eur Respir Mon* 2006;38:7–23.
2. Food and Drug Administration. Draft guidance for industry on chronic obstructive pulmonary disease: developing drugs for treatment; availability. *Fed Regist* 2007;72:63618.
3. Global Initiative for Chronic Obstructive Lung Disease. Global Strategy for the Diagnosis, Management, and Prevention of COPD, 2011. [cited 2013 Aug]. http://www.goldcopd.org/uploads/users/files/GOLD_Report_2011_Feb21.pdf
4. Viegi G, Pistelli F, Sherrill DL, *et al*. Definition, epidemiology and natural history of COPD. *Eur Respir J* 2007;30:993–1013.
5. Kuyucu T, Guclu SZ, Saylan B, *et al*. A cross-sectional observational study to investigate daily symptom variability, effects of symptom on morning activities and therapeutic expectations of patients and physicians in COPD-SUNRISE study. *Tuberkuloz ve toraks* 2011;59:328–39.
6. Schlecht NF, Schwartzman K, Bourbeau J. Dyspnea as clinical indicator in patients with chronic obstructive pulmonary disease. *Chron Respir Dis* 2005;2:183–91.
7. Tsiligianni I, Kocks J, Tzanakis N, *et al*. Factors that influence disease-specific quality of life or health status in patients with COPD: a review and meta-analysis of Pearson correlations. *Prim Care Respir J* 2011;20:257–68.
8. Voll-Aanerud M, Eagan TM, Plana E, *et al*. Respiratory symptoms in adults are related to impaired quality of life, regardless of asthma and COPD: results from the European community respiratory health survey. *Health Qual Life Outcomes* 2010;8:107.
9. Wijnhoven HA, Kriegsman DM, Hesselink AE, *et al*. Determinants of different dimensions of disease severity in asthma and COPD: pulmonary function and health-related quality of life. *Chest* 2001;119:1034–42.
10. Patrick DL, Burke LB, Gwaltney CJ, *et al*. Content validity—establishing and reporting the evidence in newly developed patient-reported outcomes (PRO) instruments for medical product evaluation: ISPOR PRO good research practices task force report: part 1–eliciting concepts for a new PRO instrument. *Value Health* 2011;14:967–77.
11. Patrick DL, Burke LB, Gwaltney CJ, *et al*. Content validity—establishing and reporting the evidence in newly developed patient-reported outcomes (PRO) instruments for medical product evaluation: ISPOR PRO Good Research Practices Task Force report: part 2–assessing respondent understanding. *Value Health* 2011;14:978–88.
12. Food and Drug Administration. Guidance for industry on patient-reported outcome measures: use in medical product development to support labeling claims. *Fed Regist* 2009;74:65132–3.
13. Mahler DA, Wells CK. Evaluation of clinical methods for rating dyspnea. *Chest* 1988;93:580–6.
14. Jones PW, Quirk FH, Baveystock CM, *et al*. A self-complete measure of health status for chronic airflow limitation. The St. George's Respiratory Questionnaire. *Am Rev Respir Dis* 1992;145:1321–7.
15. Guyatt GH, Berman LB, Townsend M, *et al*. A measure of quality of life for clinical trials in chronic lung disease. *Thorax* 1987;42:773–8.
16. van der Molen T, Willemse BW, Schokker S, *et al*. Development, validity and responsiveness of the Clinical COPD Questionnaire. *Health Qual Life Outcomes* 2003;1:13.
17. Jones PW, Harding G, Berry P, *et al*. Development and first validation of the COPD Assessment Test. *Eur Respir J* 2009;34:648–54.
18. 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. London: European Medicines Agency, 2012. [cited 2013 December 19]. http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2012/08/WC500130880.pdf
19. Leidy NK, Rennard SI, Schmier J, *et al*. The breathlessness, cough, and sputum scale: the development of empirically based guidelines for interpretation. *Chest* 2003;124:2182–91.
20. Leidy NK, Schmier JK, Jones MK, *et al*. Evaluating symptoms in chronic obstructive pulmonary disease: validation of the Breathlessness, Cough and Sputum Scale. *Respir Med* 2003;97(Suppl A):S59–70.
21. 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;13:965–75.
22. Jones PW, Chen WH, Wilcox TK, *et al*. Characterizing and quantifying the symptomatic features of COPD exacerbations. *Chest* 2011;139:1388–94.
23. Leidy NK, Wilcox TK, Jones PW, *et al*. Standardizing measurement of chronic obstructive pulmonary disease exacerbations. Reliability and validity of a patient-reported diary. *Am J Respir Crit Care Med* 2011;183:323–9.
24. Leidy NK, Murray LT, Jones PW, *et al*. Performance of the EXAcerbations of Chronic Pulmonary Disease Tool (EXACT) in three randomized controlled trials of COPD. *Ann Am Thorac Soc* Published Online First: 16 Jan 2014.
25. Darbee JC, Ohtake PJ. Outcome measures in cardiopulmonary physical therapy: Medical Research Council (MRC) dyspnea scale. *Cardiopul Phys Ther J* 2006;17:29–37.
26. Meguro M, Barley EA, Spencer S, *et al*. Development and Validation of an Improved, COPD-Specific Version of the St. George Respiratory Questionnaire. *Chest* 2007;132:456–63.
27. Hays RD, Anderson RT, Revicki DA. Assessing reliability and validity of measurement in clinical trials. In: Staquet MJ, Hays RD, eds. *Quality of Life Assessment in Clinical Trials: Methods and Practice*. New York: Oxford University Press, 1998:169–82.
28. Rothman M, Burke L, Erickson P, *et al*. Use of existing patient-reported outcome (PRO) instruments and their modification: the ISPOR Good Research Practices for Evaluating and Documenting Content Validity for the Use of Existing Instruments and Their Modification PRO Task Force Report. *Value Health* 2009;12:1075–83.
29. Kessler R, Partridge MR, Miravitlles M, *et al*. Symptom variability in patients with severe COPD: a pan-European cross-sectional study. *Eur Respir J* 2011;37:264–72.
30. Lopez-Campos JL, Calero C, Quintana-Gallego E. Symptom variability in COPD: a narrative review. *Int J Chron Obstruct Pulmon Dis* 2013;8:231–8.
31. 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;25:2043–8.
32. Etter JF. Short-term change in self-reported COPD symptoms after smoking cessation in an internet sample. *Eur Respir J* 2010;35:1249–55.
33. Sumner H, Woodcock A, Kolsum U, *et al*. Predictors of objective cough frequency in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2013;187:943–9.
34. de Oca MM, Halbert RJ, Lopez MV, *et al*. The chronic bronchitis phenotype in subjects with and without COPD: the PLATINO study. *Eur Respir J* 2012;40:28–36.
35. Coons SJ, Gwaltney CJ, Hays RD, *et al*. Recommendations on evidence needed to support measurement equivalence between electronic and paper-based patient-reported outcome (PRO) measures: ISPOR ePRO Good Research Practices Task Force report. *Value Health* 2009;12:419–29.