CENTER FOR DRUG EVALUATION AND RESEARCH

Drug Development Tool Clinical Outcome Assessment Qualification Program

---

Clinical Outcome Assessment Staff Review

Drug Development Tool Number: COA DDT #017

DRUG DEVELOPMENT TOOL (DDT) QUALIFICATION
CLINICAL OUTCOME ASSESSMENT (COA) STAFF REVIEW

INSTRUMENT NAME	Evaluating Respiratory Symptoms (E-RS) in COPD
INSTRUMENT TYPE	Patient Reported Outcome Measure
DATE ACCEPTED INTO QUALIFICATION PROGRAM	March 3, 2011
COA STAFF REVIEWER	Paula Chakravarti, M.S., M.P.H.
SECONDARY REVIEWER	Ashley F. Slagle, M.S., Ph.D.
COA STAFF ASSOCIATE DIRECTOR (ACTING)	Elektra J. Papadopoulos, M.D., M.P.H
REVIEW COMPLETION DATE	December 22, 2015
INSTRUMENT DEVELOPER	Evidera (previously United BioSource Corporation)
ENDPOINT(S) CONCEPT(S)	Severity of respiratory symptoms of COPD
INTENDED POPULATION	Patients with stable COPD (including patients with chronic bronchitis)

Paula B. Chakravarti -A
Digitally signed by Paula B. Chakravarti -A
DN: c=US, o=U.S. Government, ou=HHS, ou=AHRQ, ou=People, 0.9.2342.19200300.100.1.1=2001275335, cn=Paula B. Chakravarti -A
Date: 2016.01.12 15:24:40 -05'00'

Ashley F. Slagle -S
Digitally signed by Ashley F. Slagle -S
DN: c=US, o=U.S. Government, ou=HHS, ou=FDA, ou=People, 0.9.2342.19200300.100.1.1=2001060836, cn=Ashley F. Slagle -S
Date: 2016.01.12 15:38:22 -05'00'

Elektra J. Papadopoulos -S
Digitally signed by Elektra J. Papadopoulos -S
DN: c=US, o=U.S. Government, ou=HHS, ou=FDA, ou=People, 0.9.2342.19200300.100.1.1=1300170743, cn=Elektra J. Papadopoulos -S
Date: 2016.01.12 16:14:22 -05'00'

COA DDT #017
Evaluating Respiratory Symptoms (E-RS) in COPD
COA Staff Review

TABLE OF CONTENTS

1 Executive Summary and Recommendations

This Clinical Outcome Assessment Qualification Review concludes that the total score of the patient-reported outcome (PRO) instrument, the Evaluating Respiratory Symptoms in Chronic Obstructive Pulmonary Disease (E-RS: COPD), is qualified for use in exploratory studies for the measurement of severity of respiratory symptoms in patients with stable chronic obstructive pulmonary disease (COPD) in the outpatient setting. The E-RS: COPD is derived from the EXACT (Exacerbations of Chronic Pulmonary Disease Tool), which previously received regulatory qualification for use in exploratory studies as a measure of the symptoms of acute bacterial exacerbation of chronic bronchitis in patients with chronic obstructive pulmonary disease.[1] This document provides a review of the submitted data to support the current qualification and provides suggestions for further study to support the potential future qualification of the E-RS: COPD for use as a key study endpoint in confirmatory trials.

The evidence of content validity, including evidence that the instrument items are relevant, understandable and complete in relation to the desired claims, was well-documented and was derived from qualitative research with patients in the targeted patient population with COPD. Additionally, a panel that included experts in pulmonary medicine, clinical research, and PRO instrument development and translation served as advisors throughout the development process to ensure the instrument was developed appropriately for use in multinational clinical trials.

With regard to the instrument's other psychometric properties, the submitter has also demonstrated evidence of reliability (test-retest and internal consistency) and of construct validity of the E-RS: COPD total score.

While preliminary evidence was provided for the E-RS: COPD's ability to detect change in the setting of clinical trials, we encourage the submitter to obtain additional data from clinical trials in the targeted patient population and context of use regarding the instrument's ability to detect change.

We also encourage the development of guidelines for the analysis and interpretation of the total score including thresholds for clinically meaningful within-patient score changes. In future clinical trials, we encourage use of multiple anchor measures including patient global assessments of respiratory symptom severity to guide interpretation of meaningful change in E-RS: COPD score.

1 http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM380961.pdf

The electronically administered E-RS: COPD was the instrument used as the basis for psychometric evaluation. Although a paper-pen version is available, we strongly recommend the use of the electronically administered instrument.

The submitter is also encouraged to consider and address the following in future work:

- The submitter should clarify when or how the E-RS: COPD will be analyzed for patients whose symptoms have worsened or are experiencing an exacerbation during the course of a clinical trial.

- It is not yet confirmed that there are sufficient response options to discriminate severity at the milder end of the symptom scale. Although concept elicitation using focus groups was conducted in truly stable patients, cognitive interviews were not conducted to evaluate the appropriateness of the items in stable patients. In addition, floor effects were noted in the item analysis from the psychometric study, meaning that a substantial proportion of the population of patients with stable COPD endorsed the mildest possible severity rating such that detection of symptom improvement would be more difficult to demonstrate. Therefore, the submitter is encouraged to investigate whether potential modification of the item content is needed, to ensure that the items are able to detect improvement in symptoms in the target population, especially for the chest and sputum domains. Additional details are provided in section 6.8 (Description of Instrument Scoring) of this review.

- We recommend that the scoring method continue to be evaluated. The total score computed as the unweighted sum of the items may be less sensitive to detect changes in patients on the low or high end of the scale as compared to a weighted total score that takes into account the different severity level assessed by each item. A weighted total score may not be needed if additional evidence confirming that the unweighted total score is reasonably sensitive in detecting changes at both ends of the scale.

- The E-RS: COPD does not include skip patterns, meaning responses were required for each of the items regardless of responses to previous items. As a result, a small number of responses were logically inconsistent with each other (e.g., patient endorsement of no breathlessness overall in one item while also endorsing some degree of breathlessness associated with activity in another item). See content validity section below for additional detail. While this is an uncommon occurrence, the submitter is encouraged to continue to evaluate the potential for difficulties in interpreting treatment benefit due to logically inconsistent responses among items in future clinical trials.

- The user manual does not provide adequate guidance on how to score and interpret the E-RS: COPD data in the presence of missing data. A clear description of how to handle item-level missing data should be included in the user manual, including instructions to the user on how or whether to score the instrument when missing data is present. In addition, information was not provided on the extent to which missing days of data would impact scores when they are aggregated as an average over multiple days and used as an endpoint. Therefore, additional information on how to establish endpoints in the setting of missing

data is needed and should be added to the user manual. Specifically, methods should be explored on how to compute the total score for a day or week if data for either individual items or individual days are missing.

Throughout the rest of this review document, the E-RS: COPD is referred to simply as the E-RS.

2 Introduction and Background

The E-RS is a PRO instrument designed to collect data to quantify the severity of respiratory symptoms in outpatients with stable COPD for use in clinical trials to evaluate treatment benefit. This instrument represents an unmet need in patient-focused outcome measurement in COPD trials and will provide valuable insights into the patient experience to complement existing outcome measures (e.g., pulmonary function tests).

Summary of Instrument Development and Validation:

The E-RS is derived from the EXACT instrument[2] which was developed to measure symptoms in the setting of an acute exacerbation of COPD. The submitters hypothesized that the respiratory symptoms were similar in nature between these two states (stable and acute), but become more intense/severe when patients transition from a stable to acute state. Therefore, the E-RS was conceptualized as the 11 respiratory items, a subset of the 14 items from EXACT, excluding the three systemic symptom items that are present only in the EXACT.

The submitter gathered information through literature, expert input and patient interviews to support the item content of the 11-item E-RS (for additional details see Section 6 of this review, Content Validity). In addition, quantitative research provides early evidence of other psychometric measurement properties (see additional details in Section 7).

3 Context of Use

3.1 Disease Definition

The E-RS Total Score measures respiratory symptoms of COPD in a stable (not exacerbating) state.

According to the FDA COPD Guidance[3], COPD refers to a "chronic progressive disease caused by chronic inflammation and destruction of the airways and lung parenchyma, and is

2. http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM380961.pdf
3 http://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm071575.pdf

usually associated with tobacco smoking or prolonged exposure to other noxious particles and gasses".

3.2 Target Population

The target patient population for the E-RS includes stable adults with chronic bronchitis and/or emphysema defined as free of exacerbation for the previous 60 days.

3.3 Clinical trial design

The E-RS is for use in randomized controlled superiority trials that will test the efficacy and safety of new treatments for COPD patients.

3.4 Labeling claims based on the COA

The E-RS total score is intended to ultimately support labeling claims related to change in overall respiratory symptoms of stable COPD. Since the E-RS has not yet been qualified for use in the context of a key study endpoint in confirmatory clinical trials, we recommend that drug developers discuss any labeling plans with the relevant CDER review division.

3.5 Limitations of use

The E-RS was developed as an e-diary to be self-administered daily via a hand-held device. Other modes of administration have not been assessed and therefore the respective score properties have also not been evaluated.

The E-RS is used to evaluate stable respiratory COPD symptoms in an outpatient setting. Its measurement properties have not been assessed in an acute care or hospital setting.

4 Instrument

A copy of the E-RS scoring and screenshots are appended to this report (Appendices A and B respectively). The 11 respiratory symptom items comprising the E-RS were selected from an existing measure, the 14-item EXACT, that was developed to assesses frequency, severity, and duration of acute exacerbations of COPD. The E-RS is not to be completed in isolation, but was developed to be scored following self-administration of the full 14-item EXACT daily diary, which is completed each evening just prior to bedtime. The E-RS total score is derived by summing item/question level scores across the 11-respiratory symptom items, and can range from 0 to 40. The E-RS also includes three separate domain scores.

Each domain score is calculated in a similar manner to the total score using the item/question level raw score. The respective questions and score ranges for the domains are as follows:

- Breathlessness domain: questions 7, 8, 9, 10, 11; (0 to 17)
- Cough and Sputum domain: questions 2, 3, 4; (0 to 11)
- Chest domain: questions 1, 5, 6; (0 to 12)

The E-RS is designed to be administered via electronic data capture.

Comment: Although the E-RS includes domain scores, this qualification is currently limited to the total score that is derived from all 11 respiratory symptom items.

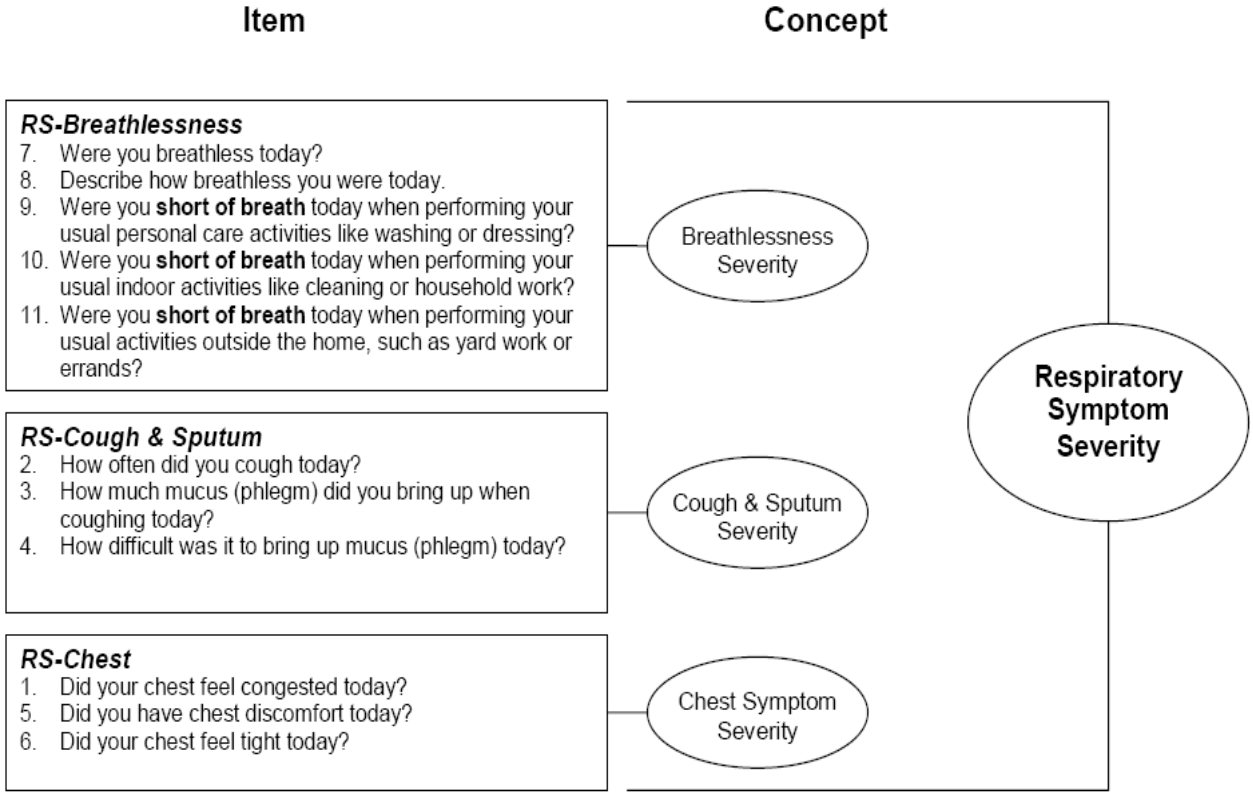
5 Conceptual Framework

The conceptual framework explicitly defines the concept(s) measured by the instrument in a diagram that presents a clear description of the relationships among items, domains (sub-concepts) and concepts measured as well as the scores produced by the instrument.

Review of the instrument development process indicated that the original items for EXACT were developed using literature review, expert input and qualitative research (with documentation of saturation from both a new sample of stable patients and a secondary analysis of data from EXACT development) in the target patient population. The E-RS total score was evaluated using a confirmatory factory analysis (CFA), and the three respiratory symptom severity domains included in the E-RS (breathlessness, combined cough with sputum, and chest symptoms) were identified through the use of an exploratory factor analysis (EFA) of the stable patient group in a quantitative study. See section 7.4 for additional details.

The conceptual framework is provided in Figure 1 below.

Figure 1 E-RS Conceptual Framework



6 Content Validity

Content validity is defined as evidence that the instrument measures the concept of interest including evidence from qualitative studies that the items and domains of an instrument are appropriate and comprehensive relative to its intended measurement concept, population, and use.

6.1 Overview and Process used to Establish Content Validity of the E-RS

The submitter employed the following steps in their development process:

a) Literature review

b) Concept elicitation and saturation achieved for EXACT (Phase I)

c) Phase II concept elicitation (new sample of stable patients)

d) Qualitative analysis

e) Item generation – conducted through EXACT

f) Cognitive Interviews

g) Item pool refinement

This section is focused on review of the following items: 1) the methods of concept elicitation; 2) the methods of qualitative analysis; 3) documentation of concept saturation; 4) confirmation of item relevance with the item mapping matrix; 5) description of scoring; 6) description of recall period; 7) review of respondent burden.

An overview of content validation as provided in the PRO dossier from the developer is shown in the following table (Table 1).

Table 1 Overview of Content Validation for E-RS

Development Activity	Purpose
Literature Review	• Identify the cardinal respiratory symptoms of COPD that patients typically experience • Explore the extent to which current tools measure and evaluate these symptoms in clinical studies of COPD
Qualitative Research Phase I: Secondary analysis of qualitative data previously collected during development of the EXACT Phase II: New focus groups in stable patients with COPD	• Assess patient descriptions of respiratory symptoms of COPD while in a stable state (not exacerbating) • Evaluate to what extent these symptoms are captured in items contained in the EXACT

The instrument developers used various sources to inform the content validity of the E-RS as a measure of respiratory symptoms of COPD. The following is a summary of the content validation activities.

- A review of the published literature was performed to 1) identify the cardinal respiratory symptoms of COPD that patients typically experience; and 2) identify PRO tools that have been used to evaluate the respiratory symptoms of patients with COPD in clinical studies.

- The qualitative research to confirm the subset of respiratory symptom items to comprise the E-RS was conducted in two phases (see Table 1 for summary of qualitative research). Concept elicitation was conducted in stable COPD patients (Phase II) and a secondary analysis was conducted using qualitative data collected during the development of EXACT among patients who had not exacerbated in the previous 10 days (Phase I). Patients from both phases completed a sociodemographic questionnaire, the Modified Medical Research Council (MMRC) Dyspnea Scale, and the St. George's Respiratory Disease Questionnaire. Patients participating in Phase II of the qualitative research also completed the Cough Severity Diary (CSD) and the E-RS instrument. The inclusion criteria for both Phase I and Phase II are described in the intended population section (Section 6.2)

Phase I: Secondary Analysis of EXACT Data

The first phase of the qualitative research for E-RS involved a secondary analysis of the qualitative data collected during development of the EXACT specifically reflecting on the severity of patients' respiratory symptoms during their usual stable state. This secondary analysis included 63 patients, which is a subset of the original 83 patients used for the qualitative study to develop the EXACT. Of the 20 patients that were excluded from the secondary analysis, 8 were excluded since they had a recent exacerbation (within the 10 days of participation) and 12 patients were excluded because they had only participated in cognitive debriefing interviews.

The 63 patients had a COPD exacerbation within the past 6 months. The qualitative data was obtained through focus groups, 2:1 and 1:1 structured interviews with patients in 5 US locations (AZ, FL, MD, TX, MI). Due to the relatively high proportion of Caucasians in the first qualitative sample for Phase I, a second round of 1:1 structured interviews was conducted with a sample that included more racial and ethnic diversity, with a specific focus on COPD patients of African-American or Hispanic descent.

The transcripts from the focus groups and interviews were matched with the EXACT coding dictionary and additional terms were included if missing. The cognitive interviews were conducted using two modes of administration (paper and pencil as well as personal handheld electronic device).

Phase II: Focus group for stable COPD
In Phase II, there were 21 patients included in 3 focus groups and one 3:1 interview. These patients either had no history of a medically confirmed COPD exacerbation or no medically confirmed exacerbation in the 12 months prior to inclusion in the study. Recruitment occurred in three clinics from two states in the Mid-Atlantic region. Interviews were conducted at two sites in VA and one site in DE.

Comments:

Phase I patients included those with exacerbation within the previous 6 months, but not within the previous 10 days, therefore, the patients could have been still recovering from a recent exacerbation at the time of the interview. In contrast, patients participating in Phase II had no history of medically confirmed COPD exacerbation in the 12 months prior to the study. Therefore, the primary focus of this review are the respiratory symptoms elicited from the patients participating in Phase II as this is considered the more stable patient population.

The qualitative report indicates the cognitive debriefing occurred only during the EXACT development and did not occur in a separate sample of stable, non-exacerbating patients. One of the major goals of cognitive debriefing is to document that patients comprehend the instrument's instructions, questions and response options as intended (i.e., to avoid misunderstandings of the instrument instructions or items). We would not expect there to be any major differences in comprehension of the items according to whether participants were stable or had recently exacerbated.

Cognitive interviewing is also used to understand the thought process involved in selection from among a set of response options, i.e., the extent to which patients are able to match their internally generated answer to the response categories provided. It is possible that patients with stable COPD may be better assessed using a greater number of items or response options targeted to their severity level. Further refinement and targeting of the items as well as the response options in the stable population may be derived from

qualitative cognitive debriefing of the final PRO instrument and/or quantitative studies in the targeted population of patients with stable COPD.

The concepts provided by the patients during concept elicitation may not have been completely unprompted. The Respiratory Symptom and Exacerbation History Screening Form Source questions may have introduced the concepts to patients. However, based on the totality of the evidence from clinician input, patient input, literature review, and initial quantitative work there is sufficient evidence of content validity.

6.2 Patient Characteristics

A summary of both the Phase I and Phase II qualitative data for subject demographics and clinical characteristics is found in the table below. As described earlier, the data from Phase II is more representative of the stable COPD patient population and is the focus of the review.

Phase II key inclusion/exclusion criteria

Key Inclusion Criteria:

1. ≥ 40 years of age;
2. Smoking history of at least 10 pack/years;
3. Current medical diagnosis of COPD (including chronic bronchitis and/or emphysema).
    a. GOLD Stage I-IV as assessed within the past 12 months;
        i. GOLD-I indicates "mild COPD." This stage is characterized by mild airflow limitation and usually, but not always, chronic cough and sputum production. FEV1/FVC < 70%; FEV1 ≥ 80% predicted.
        ii. GOLD-II indicates "moderate COPD." This stage is characterized by worsening airflow and usually the progression of symptoms, with shortness of breath typically developing on exertion. FEV1/FVC < 70%; 50% ≤ FEV1 < 80% predicted.
        iii. GOLD-III indicates "severe COPD." This stage is characterized by further worsening of airflow limitation, increased shortness of breath, and repeated exacerbations. FEV1/FVC < 70%; 30% ≤ FEV1 < 50% predicted.
        iv. GOLD-IV indicates "very severe COPD." This stage is characterized by severe airflow limitation or the presence of respiratory failure or clinical signs of right heart failure. FEV1/FVC < 70%; FEV1 < 30% predicted or FEV1 < 50% predicted plus chronic respiratory failure.
4. Have no history of medically reported COPD exacerbation associated with an unexpected visit to the clinic, emergency department, or hospital and/or prescription medication to treat an exacerbation.

OR

No medically reported exacerbation for **at least** the past 12 months. (For these participants, the screening form requested the site record of the date of the last medically reported exacerbation)

Key Exclusion Criteria:

1. Patients with current diagnosis of asthma and those with symptomatic heart failure were excluded (i.e., New York Heart Association class 2-4).

2. Also excluded were patients with a current diagnosis or recent history (past 60 days) of clinically relevant bronchiectasis, lung cancer, tuberculosis, and respiratory infection/pneumonia; or having been prescribed medication to treat any of these conditions within the past 60 days.

Phase I key inclusion/exclusion criteria are not described here but may be found in the EXACT review[4].

Patient demographics, clinical characteristics and screening data from the E-RS focus groups are found below in Tables 2 and 3.

4 http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM380961.pdf

Table 2 Phase I and Phase II Qualitative Sample Demographic Characteristics

Characteristics	Phase I: Secondary Analysis of EXACT Data[1] (N=63)	Phase II:FGs[2] in Stable COPD Patients (N=21)
Age mean (SD) [range]	65.2 (10.37) [40.9-87.4]	67.8 (8.7) [50.0-83.0]
Male (%)	28 (44.4%)	14 (66.7%)
Racial Background n (%)[3]		
White	48 (76.2%)	19 (90.5%)
Black or African American	12 (19.0%)	2 (9.5%)
Hispanic[4]	11 (17.5%)	1 (4.8%)
American Indian or Alaska Native	1 (1.6%)	0 (0.0%)
Other[5]	2 (3.2%)	0 (0.0%)
Domestic Situation n (%)[6]		
Alone	26 (41.3%)	2 (9.5%)
With Partner/Spouse/ Family	36 (57.1%)	16 (76.2%)
Widow	0 (0.0%)	2 (9.5%)
Divorced/separated	0 (0.0%)	1 (4.8%)
Employment Status n (%)		
Employed, full-time	6 (9.5%)	1 (4.8%)
Employed, part-time	7 (11.1%)	0 (0.0%)
Homemaker	3 (4.8%)	0 (0.0%)
Unemployed	7 (11.1%)	2 (9.5%)
Retired	27 (42.9%)	18 (85.7%)
Disabled	13 (20.6%)	0 (0.0%)
Education n (%)[7]		

Characteristics	Phase I: Secondary Analysis of EXACT Data[1] (N=63)	Phase II:FGs[2] in Stable COPD Patients (N=21)
< High School	10 (15.9%)	0 (0.0%)
High School	25 (39.7%)	6 (28.6%)
Some college	0 (0.0%)	6 (28.6%)
College	20 (31.7%)	6 (28.6%)
Graduate School	6 (9.5%)	3 (14.3%)
Other[8]	2 (3.2%)	0 (0.0%)
Patient Reported Comorbidity n (%)[3]		
Hypertension	21 (33.3%)	10 (47.6%)
Diabetes	12 (19.0%)	4 (19.0%)
Heart disease	10 (15.9%)	1 (4.8%)
Arthritis[9]	2 (3.2%)	9 (42.9%)
Chronic sinusitis	5 (7.9%)	2 (9.5%)
Depression	0 (0.0%)	2 (9.5%)
Angina	0 (0.0%)	1 (4.8%)
Asthma[10]	4 (6.3%)	0 (0.0%)
Other[11]	16 (25.4%)	4 (19.0%)
Smoking Status n (%)		
Current	20 (31.7%)	9 (42.9%)
Former	42 (66.7%)	12 (57.1%)

[1] Included 25 one-on-one interviews, three focus groups, and two two-on-one interviews.
[2] FG=Focus Groups, included four independent sessions with 3-7 patients each.
[3] Not mutually exclusive
[4] Participants were asked to mark Hispanic or not-Hispanic in a separate ethnicity question.
[5] Two participants marked Other: Hispanic; Spanish-Latina for Race.
[6] Missing data from one Phase I participant
[7] Categories in Phase I form: less than high school, high school, college, graduate school; categories in Phase II form: high school, some college, college, postgraduate degree
[8] Three participants checked Other: Trade School, RN
[9] Categories varied based on phase; Arthritis was not a category in Phase I but reported as 'Other'
[10] Asthma was reported as "Other", but is presented here separately (Phase I only)
[11] Other Phase I comorbid diseases included: Bronchitis (n=1), Bipolar (n=1), Chronic bronchitis (n=2), CHF, Osteoarthritis (n=2), Gout (n=2), Tremors, Epilepsy, Glaucoma, Hypoglycemia, High cholesterol, Hyperthyroidism (n=1), IR HT Beat, Kidney problems, Lung Cancer , Obese, Pacemaker valve, Stenosis, Colitis, Crohn's, and Sleep apnea; Other Phase II comorbid diseases included: high cholesterol, acid reflux.

Table 3 Phase I and Phase II Qualitative Sample Clinical Characteristics

	Phase I: Secondary Analysis of EXACT Data (N=63)[1]	Phase II: FGs[2] in Stable COPD Patients (N=21)
Duration of disease in years, mean (SD), [range]	7.9 (6.2) [1.0-30.0]	8.9 (7.3) [0.5-30.0]
Days between exacerbation and participation mean (SD), [range][3]	85.4 (50.87) [11.0-238.0]	n/a
History of medically confirmed exacerbation[4]		
Never	n/a	10 (47.6%)
> 12 months	n/a	11 (52.4%)
GOLD STATUS n (%)		
GOLD-I	2 (3.2%)	0 (0.0%)
GOLD-II	24 (38.1%)	11 (52.4%)
GOLD-III	24 (38.1%)	8 (38.1%)
GOLD-IV	13 (20.6%)	2 (9.5%)
Spirometry Results[2] mean (SD) [range]		
FEV_1	1.16 (0.44) [0.4-2.9]	1.3 (0.5) [0.5-2.5]
FEV_1/FVC	0.5 (0.13) [0.2-0.8]	0.5 (0.1) [0.2-0.7]
FEV_1 % Predicted	45.8 (16.08) [10.0-79.0]	47.9 54.0 [14.0-78.0]

[1] Included 25 one-on-one interviews, three focus groups, and two two-on-one interviews.
[2] FG=Focus Groups, included four independent sessions with 3-7 patients each.
[3] Phase I only
[4] Based on clinical chart; Phase II only

Phase I: Of 63 patients from the EXACT qualitative data subset that participated, there were 4 patients who also had asthma. The mean number of days between their last exacerbation and participation in the study was 85.4 (SD 50.9 days) days with a range of 11-238 days. For 48 patients (76%) their GOLD status was evenly distributed between GOLD-II and III. The mean FEV-1% predicted was 45.8% (SD 16.1%).

Phase II: Of the 21 patients from the stable COPD data, 19 (90.5%) were White. There were 10 patients (47.6%) that had never had a medically confirmed exacerbation and 11 patients (52.4%) who have not had a medically confirmed exacerbation in over 12 months. The GOLD status for the majority spanned between GOLD II (11, 52.4%) and GOLD III (8, 38.1%) with a mean FEV1 % predicted of 47.9% (SD 54%).

Comments:

The Phase II study excluded those with a current diagnosis of NYHA class 2-4, which is typical of clinical trials for stable COPD; therefore, the Phase II study population was representative of the intended population for this instrument. In addition, consistent with clinical trials for stable COPD patients, Phase II excluded those with a current or recent history as defined as past 60 days of lung cancer, TB, respiratory infection/pneumonia or prescribed any medication for any of these conditions in the past 60 days.

In Phase I, there is limited clarity on whether the patients were truly in a stable state for the concept elicitation and also unclear how many of them were cognitively interviewed which are concerns. In addition, the inclusion of asthmatics and those with heart disease potentially confounds the identified symptoms that are exclusively due to COPD. Therefore, the data for Phase I is being viewed exclusively as supportive in nature. In Phase II, which consisted of concept elicitation only, there were no patients recruited that had less than a high school diploma, and racial diversity was limited.

We note that wheezing was identified as a symptom by 11 of 21 patients in Phase II, but was not included in the E-RS assessment. Patients inconsistently described the symptom of wheezing and the instrument developer indicates there is overlap with wheezing and symptoms that are included in the assessment (e.g., chest congestion, chest tightness, and shortness of breath). We do not view the omission of wheezing as a critical flaw.

The development and validation of the E-RS (and EXACT) were conducted in the outpatient COPD patient population. There are no data to support the use of the E-RS in clinical studies in hospitalized patients. It is not expected that E-RS would be used in a hospital setting, because it is intended for use in stable COPD patients.

While not deemed to be a critical flaw, additional information (e.g., confirmed comorbidities, oxygen status, and current medication use) may have provided additional context for the interpretation of the qualitative data. First, comorbidities are indicated in the tables as being patient-reported; however, it is unclear if these comorbidities were confirmed with the medical record. Some of the patients indicated that they were on oxygen (based on the transcript), but that was not captured in the data tables, so it is not clear how oxygen use (or no use) might have impacted qualitative research findings. In addition, no listing of current medications was provided by the submitter, so it is unclear how medication status may have impacted the study findings.

6.3 Summary of qualitative analysis methods

The following methods were used to develop the item pool as described in the qualitative study report.

In Phase I, the E-RS project team members used the EXACT coding dictionary as the basis for the E-RS coding dictionary and added symptoms that were not already captured from their analysis of the EXACT transcripts focusing on those symptoms that could be used to describe their normal or stable (non-exacerbating) state. The acute experience was only described in order to provide a context to when the patient is not stable. There was independent coding of the first two transcripts by 2 coders which was followed by a comparison and reconciliation among the independent coders then a single coder completed the remaining transcript(s) using the reconciled version has a reference.

In Phase II, the E-RS coding dictionary that was developed in Phase I served as the basis to be updated using the data from the stable COPD patients. There was independent coding of the first three transcripts by 2 coders which was followed by a comparison and reconciliation then a single coder completed the remaining transcript.

6.4. Qualitative findings

Representative quotes from the qualitative research are as follows for Phase II. Phase 1 findings were not inconsistent with phase 2 findings, though as expected in a population with more recent exacerbations and the likelihood of more infections, the patients in Phase I also described severity of symptoms associated with exacerbations (e.g., colored sputum leading them to seek medical care).

Breathlessness:

"A simple thing like emptying the dishwasher makes me short of breath because I'm bent over and back-or even doing dishes at night. That makes me short of breath and I've got to stop and catch my breath."

"But if I go up a short-it don't have to be long stairs, just a couple stairs and I get shortness of breath. And if I do too much of anything I get shortness of breath."

Cough:

"…somebody talked to me on the phone the other day and asked me about coughing, and I mentioned earlier, I wasn't aware that I coughed as much as I did, and it's a dry cough, it seems…But it must be annoying to other people. I just didn't realize how much I coughed, a self-awareness, but somebody made me realize that I must be-it's worse than what I had expected."

"Coughing, here, there, and yonder is not a particular problem, but about once or twice a day I have a cough that if I'm in the house and there's anybody in there, they will come in and say, are you all right? You want me to call the doctor? You want me to call 911? Cause I cough really, really hard like once or twice a day and so, anyway, after a while it stops and everything's okay, but if you're in a crowd of people oh, my goodness, they would probably be calling 911. And there's times when I feel like I'm going to choke in the middle of the coughing and it takes a bit of throat clearing and this, that, and the other to get past that point. But, anyway, once I'm through with that it's all right."

Sputum:

"When I have those hard coughs like he's talking about I don't always bring up phlegm, but most of the time I'll bring up some. If you bring up that phlegm then it seems like it's all over…It seems like it's getting better."

"I think the production of sputum is a bigger problem I have…If I could get-do away with it and live-…I would have a perfect world."

Chest symptoms:

"I kind of describe it as if you were in some kind of a container and you weren't able to draw air in, what you had around you would restrict it. And you just-no matter what you did, you couldn't get enough air in or the air was so heavy or so thick that you couldn't get it into your lungs. It's kind of hard to describe but it's - it's like you're in an enclosed area and you can't draw air in."

"It sounds like anxiety, for me, and it sounds like it triggers it, the tightness. … the anxiety-once I catch myself breathing heavy, and then I start getting anxious about things, and I think it might trigger the tightness and the other, and I don't think it's like psychosomatic. I think that-they're very real to me, but usually when I-and if I sit down and try to relax and start trying to breathe a little bit easier, it goes away."

"I'm the one that put that chest heaviness there to begin with… It does not happen often…Like it's not an everyday symptom. But I did think about when activities or-what creates these problems. And I listen to-the first thing that creates my problem is just activity. Then the second thing that creates it is position. You know, when you're bending over doing something. But in my

case it's at night. And the third thing is allergies. Allergy creates a lot of problem. And the fourth thing is anxiety. For me that's the sequence of events."

The following table (Table 4) demonstrates the alignment between items in the E-RS and corresponding patient quotes.

Table 4

E-RS Item Mapping Document

Breathlessness		
Item	**Supporting information**	**Relevant Quotes**
Item 7: Were you breathless today? Item 8: Describe how breathless you were today.	• 19 participants mentioned breathlessness or shortness of breath in Phase II FGs	*"…anything that involves the expenditure of energy causes us to get short of breath" (101)* *"The heat and also cold, if it's real, real cold and the air is blowing in my face, a lot of times it will take my breath away from me." (302)* *"Just flat running out of breath, choking up." (207)* *"Well, the shortness of breath I think is the clue to the whole thing, the key to the whole thing." (305)*
	• 17 participants mentioned breathlessness s or shortness of breath in Phase I secondary analysis	*"I'm always short of breath" (FG #308)* *" I just can't motivate. On bad days, I get up, and I can't walk across the hall without losing my breath." (FG #203)* *"Sometimes I'll have trouble breathing and can't get my breath. I do use oxygen at night" (FG #204)* *"It's just an overall feeling of not getting enough oxygen…Not being able to just take a deep breathe, usually it's semi." (EXACT FG #504)* *"I just started having a lot more coughing and more difficulty breathing doing everyday things that I wouldn't normally have a problem with." (EXACT FG #303)*
Item 9: Were you short of breath today when performing your usual personal care activities like washing or dressing?	• 3 participants mentioned shortness of breath with personal care activities in Phase II FGs	*"…taking a shower makes me short…I get short of breath" (102)* *"I cannot go into the shower before – without taking medication because once I get in the shower I'm breathing like a horse"(203)* *"bending over" (202)… "tying your shoes" (203)*
	• 2 participants mentioned shortness of breath with personal care activities in Phase I secondary analysis	*"…It takes so much longer to take a shower. I take a shower with oxygen, because that's exercise. It takes longer to put makeup on, if I'm going to. Everything is harder. Dressing in the morning is harder. It's more difficult with this than it ever was." (FG #202)* *"It's between moderate, more noticeable doing light activities, and between noticeable when washing or dressing I can notice it. I do get short of breath at that time all the time, regardless of how good I feel." (CD #302)*
Item 10: Were you short of breath today when performing your usual indoor activities like cleaning or household work?	• 5 participants mentioned shortness of breath with household tasks in Phase II FGs	*"A simple thing like emptying the dishwasher makes me short of breath because I'm bent over and back-or even doing dishes at night. That makes me short of breath and I've got to stop and catch my breath. " (102)* ***Any other kind of just daily life activities that are affected by shortness of breath?*** *"Scrubbing my floor." (202)*
	• 4 participants	*"It's kind of shallow breathing. I really can't do too much housework. It really*

	mentioned shortness of breath with household tasks in Phase I secondary analysis	*knocks me out. Even walking, like going shopping." (EXACT FG #403)* *"…when I managed to get out of bed, I might get out of breath just trying to do some normal things like empty the dishwasher or something like that." (EXACT FG #502)*
Item 11: Were you short of breath today when performing your usual activities outside the home such as yard work or errands?	• 6 participants mentioned shortness of breath during outdoor activities in Phase II FGs	*"Well, besides coughing, of course, the shortness of breath and not being able to do sports and some of the things I used to be able to do or yard work." (304)* *"I mean I the cut grass. I get about halfway through. I have to breathe hard but I don't have to stop." (103)* *"I can't work as long in the outdoors gardening and doing whatever has to be done with lawns and plants and whatever. And I've got to rest more-I've got to stop more frequently." (205)*
	• 4 participants mentioned shortness of breath during outdoor activities in Phase I secondary analysis	*"Well, it makes it difficult to breathe. Some of the breathing that you do actually stings your nasal passages, and you find yourself getting throaty as I am today. When the smog goes away, the throatiness will also go away. But it's very uncomfortable. It restricts everything that you do-going to the church, going to the store, visiting friends. I no longer just go visiting." (EXACT 2:1 #101)* *"Oh yes, when I try to function. I couldn't take my mother or my girlfriend to the grocery store. I couldn't bend over to get the groceries out of the trunk. I can't go from here to outside, and I have to stop." (EXACT FG #201)* *"No, it's bad all of the time, except doing the computer. Thank goodness I can order food on the phone now for my grocery shopping." (EXACT 2:1 #702)*
Cough and Sputum		
Item	**Supporting information**	**Relevant quotes**
Item 2: How often did you cough today?	• 13 participants mentioned cough in the Phase II FG's	*"I cough. I cough all the time… with me it's almost a daily thing that I cough. Maybe not so many times a day but I'm going to cough sometime today if I haven't already." (104)* *"…but I cough all day long on and off…" (306)* *"Usually every day I have to cough." (402)* *"I do cough in the morning and sometimes during the day, it all depends what I'm doing and from exerting myself a little bit. Sometimes it causes me to cough, but most of the coughing I have is in the morning first thing…" (302)*
	• 16 participants mentioned cough in the Phase I secondary analysis	*"Yes, I cough. I cough every day. Every morning I get up, I cough." (EXACT FG #307)* *"Well, it's a very deep and a very stringent cough. Once you've started, it's very difficult to stop. It seems to challenge, probably because of the airway condition that I have with this esophagus. It seems to not be able to stop it. Once you start coughing, you cough, and the cough generates a bigger cough..." (EXACT 2:1 #101)*

		"[I cough] A lot. In the middle of a conversation, I am talking and I just go to coughing. I can't stop and it's real dry." (EXACT FG #304) *"Just coughing, but a dry cough. There hasn't been anything, no mucous or anything coming up." (EXACT FG #403)*
Item 3: How much mucus (phlegm) did you bring up when coughing today?	• 13 participants mentioned coughing up sputum/phlegm/ mucus in the in the Phase II FG's	*"Yeah, I'd like to quit coughing up phlegm…sometimes you get to coughing and the next thing you know you've got to hack it up and coughing up a bunch of phlegm and stuff and that's embarrassing, not only just coughing, but it's embarrassing, too, if you're out in public and around people" (306)* *"It's a constant thing. It's just more or less. And less is good and more is bad" (206).* *"I just have a lot of phlegm, especially when you get out of bed in the morning. You have to clear your lungs out." (402)* *"it's all clear, thick mucus that I cough up and it's not all that pleasant to describe" (304)*
	• 14 participants mentioned coughing up sputum/phlegm/ mucus in the secondary analysis	*"No, no. You know sometimes I'll spit up some-like I've got a lot of phlegm now. I spit up a lot of it now. Sometimes it's a green color and sometimes it's white." (EXACT CD #101)* *"Now, I'm coughing to clear my lungs. From the pulmonary rehab classes that I had, I'm supposed to encourage that. I'm not supposed to fight it back, because we do have all sorts of crap in our lungs from forever…" (EXACT FG #202)* *"All of that phlegm builds up in your lungs overnight. As soon as you get up and start stirring around, you got to get that out." (EXACT FG #302)* *"It's very viscous, and it sticks to everything. Trying to clean your mouth off, it's very difficult to get it free. It will kind of be a pulling event going on in your mouth, and that's the part I hate the most. I don't mind any of the other things." (EXACT 2:1 #101)*
Item 4: How difficult was it to bring up mucus (phlegm) today?	• 12 participants mentioned difficulty bringin up sputum in the Phase II FG's	*"Well, you know you need to cough up some mucus but it just don't seem to want to come up." (104)* ***Do you always cough that phlegm up when you cough?*** *"If I'm lucky.* ***If you're lucky. So, say more about that so I understand.*** *Well, sometimes it's very difficult to get it to come up. So, when that happens it makes it more difficult to breathe." (206)* *"When I have those hard coughs like he's talking about I don't always bring up phlegm, but most of the time I'll bring up some. If you bring up that phlegm then it seems like it's all over…It seems like it's getting better." (306)* *"you have to work at it" (202)*

	• 1 participant mentioned difficulty bringing up sputum in the Phase I secondary analysis	*"Well, the coughing was a lot easier. I could bring up the phlegm a lot easier, but I just felt better. I could do things quicker than I could a few days before. I was feeling everyday there was an improvement, and my breathing, of course, was part of that whole thing." (EXACT FG #503)*
Chest Symptoms		
Item	**Supporting information**	**Relevant quotes**
Item 1: Did your chest feel congested today?	• 9 participants mentioned chest congestion in the Phase II FG's	*"A lot of congestion.* ***And a lot of congestion? And where do you feel that?*** *All in my chest." (402)* *"…when I wake up I've got to get up, that's all there is to it, and when I wake up I'm stopped up and I have to get the breathing machine.* ***What do you mean stopped up? Can you tell me more about that?*** *Congested, congested." (401)* *"Well, it's an every day thing. I get a certain amount of chest congestion every day…" (306)* ***Do you use any other words to describe when you're talking to your doctor?*** *"Congestion." (102) […]* *"Congestion is how you describe it." (101)*
	• 4 participants mentioned chest congestion in Phase I secondary analysis	*" I feel closed in and congested…all the time" (2:1 #101).* *"I'm congested most of the time" (CD #302)* *"Today is a bad day, I can't hardly breathe, I'm congested, I have chest pain and I'm full of stuff, because the weather got me read bad. I never feel good. I get up but I feel pretty good and then as I move around, get my oxygen and stuff on it helps me feel a little better. Tomorrow I probably be in the bed all day, you know." (CD #303)*
Item 5: Did you have chest discomfort today?	• 9 participants mentioned chest discomfort or heaviness in Phase II FGs	*"I have chest discomfort…* ***Do you tend to have it most days or just some days?*** *No, no just when I'm doing something or in the process of doing something that's really physically hard and then once I stop I just feel like whew, something's sitting on my chest." (303)* *"I try to live life good, as normal as you can, but there's a lot of discomfort in it.* ***…Can you tell me more about where the discomfort is?*** *"Well there's always tightness and all.* ***…tightness and discomfort in your chest?*** *Yeah, chest discomfort." (306)* *"But my lungs, if it gets bad enough where my chest starts hurting, that's it. Sometimes I'll be out there mowing and all of a sudden I just cut the mower off and fall on the ground." (101)* *"It feels like an elephant sitting on your chest" (104)*

	• 4 participants mentioned chest discomfort or heaviness in Phase I secondary analysis	*"What does it feel like?...It feels something like a little heaviness. Feels heavy. And it might even be sometimes you take a deep breath and you might even get a little pain with it and your throat gets dry." (CD #308)* *"At night and in the mornings, it feels like somebody standing on my chest when I get up" (CD #304).* *I will say yes because I can tell when I'm congested and I can't breathe, but the discomfort, it's just the pain in my chest from the coughing. Every time I cough my chest hurts so much." (CD #202)*
Item 6: Did your chest feel tight today?	• 13 participants mentioned chest tightness in Phase II FGs	*"Well, there's always tightness and all…Well, I do. I feel like somebody's squeezing your chest like this and it's about an every day thing all-in-all, maybe once a day or something like that." (306)* *"Oh, I've had chest tightness before, that usually means it's time for a rescue inhaler or something like that." (305)* *"Well, it tightens up right here like somebody-* ***Tightens up. You're pointing to your chest?*** *Like-yeah. Right here and I-like somebody just closed up the breathing pipe and my shoulders are hitting my ears." (203)* *"you feel like it's so tight that you think you're going to stop breathing" (403)*
	• 4 participants mentioned chest tightness in Phase I secondary analysis	*"You feel like you have phlegm and you have shortness of breath and you have to cough. You feel some sort of pressure or tightness. It doesn't feel right. It doesn't feel smooth, like you're breathing normally." (CD #201)* *"I can only explain it this way. You feel as though your air is up to here and you can't get any more than that, especially when your chest feels really tight and you have congestion, as well. You're in for some trouble then. You have to just sit back and let it run its course." (CD #302).*

6.5 Review of evidence of saturation

Qualitative studies were performed involving two sets of patients with the goal of understanding how patients with COPD describe exacerbations. For Phase I, saturation was defined as two focus group discussions and two 2:1 or 1:1 structured interview groups in which no new concepts were introduced by the participants, beyond those identified by previous participants, documented in the form of a saturation grid. Patients who had experienced a clinician-confirmed exacerbation during the previous 6 months participated in focus groups and patient interviews to generate the initial set of data. For Phase II, saturation was defined as two focus group discussions in which no new concepts were introduced by the participants. In Phase II, patients had not experienced an exacerbation in the previous 12 months

The Phase II saturation grid is below (Table 5). As evident here, saturation of the respiratory symptoms concept was reached after completion of the third focus group, after which no new respiratory symptoms emerged. In the saturation grid, when a symptom was indicated, the participant ID number was marked with an X for the particular symptom.

Table 5 E-RS Focus Group (Phase II) Saturation Grid

	FG 1	FG 2	FG 3	FG 4	101	102	103	104	105	201	202	203	204	205	206	207	301	302	303	304	305	306	401	402	403
Breathlessness																									
Breathless / short of breath	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X		X	
Short of breath during activities	X	X	X	X	X	X	X	X	X		X	X	X	X	X	X	X	X	X	X	X	X		X	
Cough and Sputum																									
Cough (frequency)	X	X	X	X					X	X	X		X		X	X		X		X	X	X	X	X	X
Sputum/mucus/ phlegm	X	X	X	X			X	X	X	X	X				X			X		X	X	X	X	X	X
Difficult bringing up sputum/mucus/ phlegm	X	X	X	X	X		X	X		X	X				X			X		X	X	X	X		X
Chest Symptoms																									
Chest congestion	X	X	X	X	X	X					X				X					X		X	X	X	X
Chest discomfort/ heavy	X	X	X		X	X		X			X				X	X			X	X		X			
Chest tightness	X	X	X	X	X	X		X		X	X				X	X		X	X		X	X	X		X
Other																									
Wheezing		X	X	X							X		X		X			X	X	X	X	X	X	X	X

6.6 Review of item-tracking matrix

An item tracking matrix is a record of the development of items used in an instrument that can be helpful to document the changes or deletions in items and the reasons for those changes.

An item-tracking matrix was not provided as no modifications were made to the items from the final parent instrument, EXACT, following the additional concept elicitation in stable patients.

6.7 Review of final item pool development

The eleven items focusing on respiratory symptoms which comprise the E-RS scoring algorithm were selected for inclusion based on the literature review, secondary qualitative analysis from EXACT development data, and analysis of qualitative data from focus groups with stable patients.

Different sources of information guided item pool development. These sources included expert input, input from the FDA, input from patients (Phase I and II of the qualitative research), and Rasch methodology used in the development of the EXACT-PRO.

The Item Summary Table from the EXACT review (Table 3) provides a description of the original items and modifications that were made, resulting in the final version of the EXACT and E-RS.[5]

Comments: The transcripts, as well as the Atlas.ti summaries, were evaluated as part of this review. The goal was to ensure that elements from the transcripts and summaries were incorporated into the assessment. Transcripts were only provided for Phase II.

In Phase II, the patients noted it was difficult to distinguish between chest and lungs for the symptom of chest tightness/hurting, some of the patients indicated that they only have a dry cough, which is not included in the coding dictionary.

The major complaint voiced was difficulty breathing in the Phase II transcript. Many times difficulty breathing was mentioned in the context of having difficulty with their sleep, indoor or outdoor activities.

Patients were not always able to clearly distinguish between the response options "severely" and "extremely". This was found in qualitative research done by instrument developers and summarized under section 6.10 of this review. To address this, "severely" and "extremely" are scored the same (both given a score of 3) for some items. In addition, "a little" and "some" are both assigned as score of 1. Shortness of breath while performing activities utilized 6 response options. In each of the cases where 6 response options appear, 2 or 3 of the more "severe responses" are collapsed for scoring purposes.

The response options for question #8 (Describe how breathless you were today) include both "breathless during light activity" and "breathless when washing and dressing". The cognitive interview transcripts from Phase I indicate some variability in the interpretation of "light activity" where some participants viewed light activity as washing and dressing whereas others had a different interpretation. For clarity, examples of light activity could be presented as part of that response option. The last response option is "present when resting", but the scoring is the same as "breathless when washing or dressing". In future studies, the submitter is encouraged to investigate the scoring for these response options.

One possible limitation with the instrument design is that it does not allow for skip patterns, which could result in potentially inconsistent responses. This lack of skip patterns has led to a few minor inconsistencies in responses to some of the questions. For example, there were inconsistencies relating to Question #7 (Were you breathless today?) in the validation study and in the three clinical trials that included the E-RS. For the subsequent questions that asked questions about shortness of breath while performing

5. http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM380961.pdf

different activities, an analysis of the data showed that there was a small percentage of participants in the studies that answered that they were slightly breathless during these activities even though their answer to Question #7 was "not at all". The analysis demonstrated that a consistently low percentage of "logically inconsistent profile" occurred and remained low over time (data not shown).

Since the percentage of logically inconsistent responses was low, it is unlikely to significantly affect the validity of the instrument. It may be prudent to confirm these findings in subsequent clinical trials to ensure that logical inconsistencies do not affect the validity and interpretation of the results from the instrument.

6.8 Description of instrument scoring

The E-RS is a subset of the respiratory symptoms from the 14-item EXACT-PRO daily diary. The developer of the instrument indicates that the qualitative study is supportive of the existing items in the EXACT which comprise the E-RS. The raw scoring is the same between the EXACT and E-RS for the symptoms. The final scoring for E-RS will be obtained by taking the straight sum of the 11 E-RS symptom raw scores. The total score ranges from 0 to 40 with a higher score indicating more severe respiratory symptoms. Each domain score is calculated in a similar manner to the total score using the question level raw score. The respective questions and score ranges for the domains are as follows:

- Breathlessness domain: questions 7, 8, 9, 10, 11; (0 to 17)
- Cough and Sputum domain: questions 2, 3, 4; (0 to 11)
- Chest domain: questions 1, 5, 6; (0 to 12)

Assignment of item-level raw scores can be found in Appendix A.

In order to confirm that the IRT (item response theory) analysis used to develop the EXACT was appropriate for the E-RS and could inform E-RS scoring, the submitter conducted a confirmatory factor analysis. This analysis was conducted using the data from the validation (quantitative) study with the Stable patients using data from Day 1 of the study. The data regarding score distribution and item to item correlation from the validation (quantitative) study is shown in Tables 6 and 7.

The submitter indicated that additional data was required in order to determine how the E-RS will respond over time when the scores will be used as secondary endpoints in clinical trials.

Table 6 Item Analysis: E-RS Item Descriptive Statistics – Stable Group Day 1

Item[a]	N	Mean (SD)	Range	Median	Mode	Floor (%)	Ceiling (%)	% Missing	Item-total correlation
RS-Chest									
Item - 1 Chest feel congested	188	1.0 (0.87)	0-3	1	1	63 (33.5%)	10 (5.3%)	0.0%	0.65
Item - 5 Have chest discomfort	188	0.7 (0.82)	0-4	1	0	86 (45.7%)	2 (1.1%)	0.0%	0.60
Item - 6 Chest feel tight	188	0.8 (0.81)	0-4	1	0	81 (43.1%)	1 (0.5%)	0.0%	0.63
RS-Cough & Sputum									
Item - 2 How often cough	188	1.7 (1.06)	0-4	2	2	30 (16.0%)	6 (3.2%)	0.0%	0.63
Item - 3 How much mucus when cough	188	0.8 (0.62)	0-2	1	1	62 (33.0%)	0 (0.0%)	0.0%	0.53
Item - 4 Difficult bring up mucus	188	1.1 (1.14)	0-4	1	0	72 (38.3%)	7 (3.7%)	0.0%	0.60
RS-Breathlessness									
Item - 7 Breathless today	188	1.4 (0.90)	0-4	1	1	30 (16.0%)	3 (1.6%)	0.0%	0.70
Item - 8 Describe how breathless	188	1.5 (0.92)	0-3	2	2	31 (16.5%)	25 (13.3%)	0.0%	0.55
Item - 9 Short of breath - personal care	188	1.0 (0.91)	0-4	1	1	68 (36.2%)	2 (1.1%)	0.0%	0.63
Item - 10 Short of breath - indoor	188	1.0 (0.89)	0-3	1	1	61 (32.4%)	10 (5.3%)	0.0%	0.60
Item - 11 Short of breath - outdoor	188	1.1 (0.97)	0-3	1	0	65 (34.6%)	18 (9.6%)	0.0%	0.56

[a] RS-domain score ranges are as follows with the higher values indicating greater severity of respiratory symptoms: RS-Breathlessness scores range from 0 to 17; RS-Cough & Sputum scores range from 0 to 11; and RS-Chest scores range from 0 to 12.

Table 7 Item Analysis: E-RS Item Descriptive Statistics – Stable Group Day 1

Item	RS-Chest			RS-Cough & Sputum			RS-Breathlessness				
	Item 1	Item 5	Item 6	Item 2	Item 3	Item 4	Item 7	Item 8	Item 9	Item 10	Item 11
RS-Chest											
Item - 1 Chest feel congested	1.00	--	--	--	--	--	--	--	--	--	--
Item - 5 Have chest discomfort	0.63***	1.00	--	--	--	--	--	--	--	--	--
Item - 6 Chest feel tight	0.52***	0.61***	1.00	--	--	--	--	--	--	--	--
RS-Cough & Sputum											
Item - 2 How often cough	0.60***	0.45***	0.49***	1.00	--	--	--	--	--	--	--
Item - 3 How much mucus when cough	0.46***	0.36***	0.35***	0.66***	1.00	--	--	--	--	--	--
Item - 4 Difficult bring up mucus	0.55***	0.45***	0.51***	0.56***	0.49***	1.00	--	--	--	--	--
RS-Breathlessness											
Item - 7 Breathless today	0.43***	0.40***	0.43***	0.38***	0.33***	0.42***	1.00	--	--	--	--
Item - 8 Describe how breathless	0.26**	0.23*	0.32***	0.32***	0.27**	0.32***	0.64***	1.00	--	--	--
Item - 9 Short of breath - personal care	0.37***	0.33***	0.37***	0.35***	0.36***	0.33***	0.60***	0.54***	1.00	--	--
Item - 10 Short of breath – indoor	0.37***	0.35***	0.34***	0.33***	0.22*	0.28***	0.61***	0.47***	0.65***	1.00	--
Item - 11 Short of breath – outdoor	0.31***	0.35***	0.41***	0.30***	0.16*	0.28***	0.56***	0.45***	0.50***	0.55***	1.00

[1] Pearson product moment correlation
*P<0.05; **P<0.001; ***P<0.0001

Comments: The developers conducted a confirmatory factor analysis in the Stable Group using data from Day 1, which did not confirm the uni-dimensionality of these 11-items. This was the developer's rationale for why no Rasch analysis was performed with the E-RS as was conducted with the EXACT, and the reasoning for the difference in scoring approaches between E-RS and EXACT.

We note that in a stable COPD population (Table 6), there were some floor effects in some of the symptom items. In particular, items about chest discomfort and chest tightness resulted in a score of 0 for 46% and 43% respectively. There were no major ceiling effects, so the tool may be better at detecting worsening than improvement, at least in a mildly

symptomatic patient population. It may be useful to enrich study populations with patients who are moderately symptomatic at baseline if the goal is to detect improvement with treatment.

The data in Table 6 seem to support the Phase II qualitative research in terms of the respiratory symptoms assessed. In the majority of the cases, the highest score of 4 was not attained and when it was, a very limited number of patients recorded that level of severity. On the whole, for the item-to-item correlation in Table 7, there was good correlation among the items in the respective domains.

6.9 Description of the recall period

E-RS evaluations of respiratory symptoms are based on a daily recall period. The E-RS is a subset of items from the EXACT daily diary to be completed each evening before bedtime and each item comprising the instrument references "today."

The need for a daily diary to assess respiratory symptoms of COPD was supported by the qualitative focus groups in stable patients with COPD where patients described variability, both within-day and day-to-day, in their respiratory symptoms based on factors such as activity level, stress levels and weather conditions.

Comment: It may be useful to clarify in training materials to describe what the timeframe "today" is intended to include (e.g., referring to the time from waking up that day to the evening when the diary is completed, or if it includes the overnight period from the last time the diary was completed).

6.10 Description of evidence of patient understanding

As discussed above, cognitive debriefing occurred only during the EXACT development and no cognitive interviews were conducted in a separate sample of stable, non-exacerbating patients. Therefore, the developer assumed that the exacerbating patients would be able to accurately recall their stable state condition and that patient understanding of the questions would not vary by their disease severity state.

Per the EXACT dossier (qualitative report), a total of 35 patients (23 from the initial cognitive and structured interviews with exacerbation ≤ 6 months + 3 from initial cognitive interviews only with exacerbation ≤ 6 months+ 9 from the revised draft EXACT item pool) participated the cognitive debriefings on the EXACT.

Timeframe

Participants cognitively debriefed on the draft EXACT versions dated June 5, 2006 and August 22, 2006 reported that the questionnaire was to be completed before bed every night, and that they were to reflect on the day's experiences. Patients also said they would have no trouble

recalling the events of the day during an exacerbation. To keep patients grounded in a single day, the word "today" was added to each item in the draft EXACT item pool dated August 22, 2006.

Instructions
Patients reported that the instrument was to be completed each evening before going to bed. Participants stated that they understood how to complete the items—i.e., by reflecting on their exacerbation experience that day and marking the response option that most closely matched that experience.

Item Stems:
Item stems were revised or modified based on input from patients and from instrument development, clinical, and translation experts. The majority of revisions occurred after each round of cognitive debriefings.

Response options:
In the first round of cognitive debriefing with the draft EXACT item pool dated June 5, 2006, response options for several items were tested by giving participants a set of index cards, each with a different response option.

Participants were asked to place each response option on the index cards along a 0–100 scale, with higher numbers indicating greater severity.

According to the qualitative study report, more than half of the participants who ordered the intensity response options (shown to participants with Item 1) confirmed that the order was appropriate. Some patients switched "severely" and "extremely."

In keeping with the intention that the EXACT be administered via electronic diary, the draft EXACT item pool was entered into a personal electronic device system (screenshots dated December 11, 2006) for patient evaluation. Directions for the draft EXACT item pool completed on the PDA instructed patients to tap on the boxes to record a response. There were nine patients that participated in the user acceptability evaluation of the PDA. Patients did not express any difficulty using the electronic implementation of the instrument.

6.11 Review of respondent burden

Patients with COPD who are experiencing acute exacerbation often have difficulty with the basic activities of daily living. According to the EXACT PRO dossier, the instrument developers considered twice daily administration, however, to reduce respondent burden, it was decided to administer the EXACT once a day, in the evening prior to bedtime. Since the E-RS is to always be administered as part of the 14-item EXACT, the schedule of administration is identical to EXACT.

The instrument developers also sought to minimize the length of the questionnaire from 25 questions with the EXACT, while preserving the tool's reliability and validity. The final tool for

E-RS contains a total of 14 items that patients complete daily and of these 11 items are included in the E-RS total score.

Comments: We agree with the approach to balance the development of a comprehensive instrument with the need to limit patient burden. We agree that daily (rather than twice daily) and the inclusion of the selected subset items from the larger item pool is appropriate.

7 Assessments of Other Measurement Properties

7.1 Validation and clinical trial study design, patient disposition, compliance and patient characteristics

The E-RS was psychometrically evaluated following analysis plans submitted to the Agency describing the analyses to be conducted in support of reliability, construct validity and ability to detect change.

Initial psychometric assessment of the E-RS was based on secondary analyses from a two-group observational study that had been previously conducted for the validation of the EXACT that included a total of 410 COPD patients (222 acute patients and 188 stable patients). The pool of 25 items used to develop the electronic EXACT daily diary was the instrument used in the study.

In addition, the E-RS (final version) was used in subsequent clinical trials contributing further evidence to the available psychometric data.

The Stable Group (N=188) from the initial observational study was used in the secondary analysis for the E-RS. Key eligibility criteria are as follows:

3.2.1 Key Inclusion Criteria for Stable Group

1. History of one or more acute exacerbations within the past 24 months:
    a. Exacerbation defined as a sustained worsening of the patient's condition, from the stable state and beyond normal day-to-day variations, that is acute in onset and necessitates a change in regular medication in a patient with underlying COPD;
    b. Exacerbation may be associated with a telephone call or an unexpected clinic, ER, or hospital visit.
    c. Exacerbation history was clinician-determined.
2. Current diagnosis of COPD and/or chronic bronchitis:
    a. COPD defined as by the GOLD Initiative during stable state (at least 60 days before or 60 days after acute exacerbation event):
        i. GOLD-0 At Risk - characterized by chronic cough and sputum production. Lung function, as measured by spirometry, is still normal.
        ii. GOLD-1 indicates –mild COPD. This stage is characterized by mild airflow limitation and usually, but not always, chronic cough and sputum production. $FEV_1/FVC < 70\%$; $FEV_1 \geq 80\%$ predicted.
        iii. GOLD-2 indicates —moderate COPD. This stage is characterized by worsening airflow and usually the progression of symptoms, with shortness of breath typically developing on exertion. $FEV_1/FVC < 70\%$; $50\% \leq FEV_1 < 80\%$ predicted.
        iv. GOLD-3 indicates —severe COPD. This stage is characterized by further worsening of airflow limitation, increased shortness of breath, and repeated exacerbations. $FEV_1/FVC < 70\%$; $30\% \leq FEV_1 < 50\%$ predicted.

v. GOLD-4 indicates "very severe COPD." This stage is characterized by severe airflow limitation or the presence of respiratory failure or clinical signs of right heart failure. $FEV_1/FVC < 70\%$; $FEV_1 < 30\%$ predicted or $FEV_1 < 50\%$ predicted plus chronic respiratory failure**;**

3. > 40 years of age;
4. Smoking for at least 10 pack/years;
5. Willing and able to provide written informed consent;
6. Has a telephone land line; and
7. Able to speak and read English.

3.2.2 Key Exclusion Criteria for Stable Group

1. Participant experienced an exacerbation within 60 days prior to enrollment;
2. Concurrent diagnosis of asthma with no obstructive disease (post bronchodilator >80%; $FEV_1/FVC \geq 70\%$), and no chronic bronchitis;
3. Concurrent diagnosis of clinically relevant bronchiectasis;
4. Concurrent medical or psychiatric condition that, in the investigator‘s opinion, may affect participation in the study; or
5. Visual or cognitive impairment that would interfere with completing questionnaires.

The final version of the E-RS was used for three phase 2 clinical trials. These phase 2 trials had varied study designs. One trial, Mpex (NCT00739648), was a 6-month trial conducted in the US testing MP-376 (Levofloxacin) Inhalation Solution administered for 5 days every 28 days to reduce exacerbations in high risk COPD patients, with exacerbation rate over the study period serving as the primary efficacy endpoint.

In addition, there were two phase 2 trials sponsored by AstraZeneca (AZ) to evaluate AZD9668 (a neutrophil elastase inhibitor) for 12 weeks with an additional two weeks of follow-up. The AZ trials used a medication to treat patients who had experienced an exacerbation in the previous 1-12 months requiring treatment with a corticosteroid, but receiving different maintenance treatments. These are subsequently referred to as AZ 12 (NCT00949975) and AZ 20 (NCT01023516).

This review focuses on the psychometric analyses (i.e., assessment of reliability, construct validity, and ability to detect change) obtained from the two phase 2 AZ trials, because these trials included a patient population more representative of those included in clinical trials enrolling patients with stable COPD.

The specifics of the two AZ trials are below:

AZ 12 study:

- 12 week parallel-group multinational trial (NCT00949975)
- Dose ranging study with placebo as a comparator, twice daily dosing

- Patients on a maintenance treatment of tiotropium
- Primary efficacy endpoint was pre-bronchodilator FEV1
- Key inclusion criteria:
    - Age 40-80 years inclusive, either gender
    - Post-bronchodilator FEV1% predicted of 40-80%
    - 1 or more clinical visit or hospitalization required for exacerbation in the previous 12 months
    - Total COPD symptom score of ≥2 per day for at least a week (7 days) in the prior two weeks before the randomization/enrollment visit.
    - Morning recordings of daily FEV1 for a minimum of 10 days in the prior two weeks before the randomization/enrollment visit.

AZ 20 study:

- 12 week parallel-group multinational trial (NCT01023516)
- Testing 1 dose against placebo twice daily
- Patients on a maintenance treatment of budesonide/formoterol
- Primary efficacy endpoint was pre-bronchodilator FEV1
- Key inclusion criteria:
    - Age 40-80 years inclusive, either gender
    - Post-bronchodilator FEV1% predicted of 40-80%
    - 1 or more clinical visit or hospitalization required for exacerbation in the previous 12 months
    - Total COPD symptom score of ≥2 per day for at least a week (7 days) in the prior two weeks before the randomization/enrollment visit.
    - Morning recordings of daily FEV1 for a minimum of 10 days in the prior two weeks before the randomization/enrollment visit.
    - Received inhaled corticosteroid (ICS) as monotherapy or in combination with any long acting bronchodilator in the prior 3 months

Descriptive Statistics:

Compliance:

Compliance was defined as the total number of actual completed diaries divided by the number of diary entries expected.

Overall, compliance with E-RS completion was good (i.e., >90% across both studies in non-exacerbating patients). As expected, compliance dropped in those patients who were hospitalized, which was a small proportion of the total patient population (data not shown).

Demographics:

The patient characteristics including demographics, clinical characteristics as patient reported are shown below in Table 8.

Table 8 AZ12 and AZ 20: Demographic and Clinical Characteristics

	AZ 12 (N=749)	**AZ 20 (N=597)**
Age		
Mean (range)	62 (40-80)	62 (41-80)
Gender (n, %male)	572 (76%)	443 (74%)
Race		
White	536 (72%)	592 (99%)
Asian	212 (28%)	
Other	1 (0.1%)	5 (0.8%)
Years since COPD diagnosis		
Mean (range)	9 (1-51)	7 (1-41)
Spirometry		
FEV1 % Predicted		
Mean (range)	59 (23-107)	54 (23-106)
GOLD stage		
1=mild	23 (3%)	22 (4%)
2=moderate	503 (69%)	300 (52%)
3=severe	197 (27%)	240 (42%)
4=very severe	5 (0.7%)	16 (3%)
Smoking status		
Yes	340 (45%)	264 (44%)
No	409 (55%)	333 (56%)
Number of acute exacerbation in past 12 months		
Mean (range)	1.4 (1-14)	1.2 (1-4)
Time since most recent exacerbation (days)		
Mean (range)	164 (28-376)	156 (28-366)

Comments: The demographic and clinical characteristics are consistent between AZ 12 and AZ 20, and appear typical of clinical trial populations in this context. In both studies, the mean age was 64 years and both included a higher proportion of men (74-76%). AZ 20 included 99% White patients, while AZ 12 included more racial diversity. Mean FEV1% predicted was 54-59% across the two studies. AZ 20 included a higher proportion of GOLD Stage 3-4 (severe and very severe) compared to AZ 12. While these samples are adequate and support qualification for exploratory use, additional psychometric testing

performed in more racially diverse samples would be helpful to further evaluate these findings.

7.2 E-RS Performance in Clinical Trials

Both E-RS total and domain scores were presented in the dossier with the total score being primary and domain scores functioning in a supportive manner.

The E-RS was completed as part of the EXACT instrument and self-administered daily in the evening. The duration of data collection varied from study to study. For AZ 12, E-RS data was collected from the screening visit (14-17 days prior to the randomization visit) up to the last treatment visit at 12 weeks. In AZ 20, E-RS data was collected from the screening visit (3 weeks before the randomization visit) up to the follow-up visit at 14 weeks.

The theoretical range for the E-RS total score is 0-40 and for the breathlessness, cough & sputum and chest symptoms are 0-17, 0-11 and 0-12 respectively. Higher values for either the total or domain score indicate greater severity of either total respiratory symptoms or for the particular domain being assessed.

Tables 9 and 10 provide results of the E-RS total score and domain scores from the AZ studies.

Table 9. AZ 12: E-RS Scores at Baseline and Final Week of Study (N=749)

E-RS	N	Mean	SD	SEM	Min-Max
Baseline Daily Scores[2]					
RS-Total Score	746	15.8	6.6	0.2	1-37
RS-Breathlessness Score	749	7.8	3.5	0.1	0-17
RS-Cough & Sputum Score	749	4.1	1.8	0.1	0-9
RS-Chest Symptoms Score	749	3.8	2.2	0.1	0-12
Baseline Mean Weekly Scores[3]					
RS-Total Score	749	15.9	6.0	0.2	2-35
RS-Total Score Intra-individual variability	749	2.5	1.6	0.1	0-10
RS-Breathlessness Score	749	7.9	3.2	0.1	0-17
RS-Cough & Sputum Score	749	4.2	1.6	0.1	0-9
RS-Chest Symptoms Score	749	3.8	1.9	0.1	0-11
Final Week[4]					
RS-Total Score	732	14.7	6.7	0.2	1-38
RS-Total Score Intra-individual variability	732	1.8	1.5	0.1	0-9
RS-Breathlessness Score	732	7.5	3.6	0.1	0-17
RS-Cough & Sputum Score	732	3.7	1.7	0.1	0-10
RS-Chest Symptoms Score	732	3.5	2.2	0.1	0-11

[1]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 0-40; RS-Breathlessness scores range from 0-17; RS-Cough & Sputum scores range from 0 to 11; and RS-Chest scores range from 0 to 12
[2]Baseline for daily scores is defined as (Visit 2, Day -1), representing the participant's stable or usual state.
[3] Baseline for mean weekly scores is defined as the mean of scores for the seven days preceding Visit 2(Day -7 to Day -1). Data for ≥4 days is required to calculate baseline using E-RS guidelines.
[4] Final weeks is defined as the 7 days prior to Visit 6 or last 7 days prior to final Visit/Early Termination. Data for ≥4 days is required to calculate Final week.

Table 10. AZ20: E-RS Scores at Baseline and Final Week of Study (N=597)

E-RS	N	Mean	SD	SEM	Min-Max
Baseline Daily Scores[2]					
RS-Total Score	597	17.7	6.3	0.3	1-40
RS-Breathlessness Score	597	8.7	3.4	0.1	0-17
RS-Cough & Sputum Score	597	4.5	1.8	0.1	0-11
RS-Chest Symptoms Score	597	4.5	2.1	0.1	0-12
Baseline Mean Weekly Scores[3]					
RS-Total Score	597	18.2	6.0	0.2	3-40
RS-Total Score Intra-individual variability	597	2.5	1.6	0.1	0-11
RS-Breathlessness Score	597	8.9	3.1	0.1	0-17
RS-Cough & Sputum Score	597	4.7	1.6	0.1	0-11
RS-Chest Symptoms Score	597	4.6	1.9	0.1	0-12
Final Week[4]					
RS-Total Score	586	16.1	6.8	0.3	1-39
RS-Total Score Intra-individual variability	586	1.9	1.5	0.1	0-9
RS-Breathlessness Score	586	8.1	3.6	0.2	0-17
RS-Cough & Sputum Score	586	3.9	1.8	0.1	0-11
RS-Chest Symptoms Score	586	4.1	2.2	0.1	0-12

[1]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 0-40; RS-Breathlessness scores range from 0-17; RS-Cough & Sputum scores range from 0 to 11; and RS-Chest scores range from 0 to 12
[2]Baseline for daily scores is defined as (Visit 2, Day -1), representing the participant's stable or usual state.
[3] Baseline for mean weekly scores is defined as the mean of scores for the seven days preceding Visit 2(Day -7 to Day -1). Data for ≥4 days is required to calculate baseline using E-RS guidelines.
[4] Final week is defined as the 7 days prior to Visit 6 or last 7 days prior to final Visit/Early Termination. Data for ≥4 days is required to calculate Final week.

Comments: The baseline mean total score for E-RS across the two studies ranged from 15.8 to 17.7; the slightly higher mean total score in the AZ 20 study compared with AZ 12 study is consistent with a higher proportion of patients with more severe GOLD scores. Mean E-RS total scores at the end of the each study were only slightly improved (1.1-1.6 points improved) compared with baseline scores.

7.3 Description of Evidence of Reliability in Clinical Studies

Internal Consistency Reliability:

Cronbach's alpha was used to describe internal consistency reliability with a target value of greater than 0.70. The results for AZ 12 are presented in the following table (Table 11).

AZ12

Table 11. AZ 12: Internal Consistency Reliability of E-RS Daily Scores (Visit 2, Day -1)

E-RS Scores	N	Cronbach's Alpha
RS-Total Score	749	0.92
RS-Breathlessness	749	0.90
RS-Cough & Sputum	749	0.68
RS-Chest Symptoms	749	0.89

[1] Or last day of data before Visit 2(Day-7 to Day -1).

Comments: E-RS total score demonstrated high internal consistency (Cronbach's alpha >90).

AZ 20 showed similar internal consistency values as AZ 12 (data not shown). The cough and sputum domain score is consistently lower than any of the other domain scores for both studies, though not substantially lower than 0.70. As stated earlier, the primary focus of the psychometric property review is on the E-RS total score which demonstrated high internal consistency and is the score recommended for qualification.

Test-retest Reliability:

Only those patients whose COPD status was classified as unchanged were part of the analysis for reproducibility. Day -7 to Day-1 prior to treatment initiation was used to define the stable state. For the AZ studies, the effect size (ES) was computed using the following equation (Day -7 E-RS score – Day -1 E-RS score)/SD of Day -1 E-RS score. Both the intraclass correlation coefficient (ICC) and ES were calculated for the AZ studies using a random effect model. The results can be found in the tables below (Tables 12 and 13).

Table 12. AZ 12: Seven-day Reproducibility of E-RS Scores during Study Run-in Period

E-RS Scores	N	Day -7 Mean(SD)	Day -1 Mean(SD)	Mean Difference[2] (SD)	Effect Size[3]	ICC[4]
RS-Total Score	715	16.0(6.42)	15.9(6.55)	0.2(4.66)	0.03	0.74
RS-Breathlessness	715	8.0(3.39)	7.9(3.50)	0.1(2.64)	0.03	0.71
RS-Cough & Sputum	715	4.2(1.76)	4.2(1.77)	0.0(1.35)	0.02	0.71
RS-Chest Symptoms	715	3.8(2.12)	3.8(2.14)	0.0(1.64)	0.00	0.71

[1] Pre-treatment, Visit 2, Day -7 to Day -1.
[2] Mean difference = average of (Day -7 E-RS score - Day -1 E-RS score).
[3] Effect size = Day -7 E-RS score - Day -1 E-RS score/SD of Day -1 E-RS score.
[4] Intraclass correlation coefficient (ICC) based on random-effect model.

Table 13. AZ 20: Seven-day Reproducibility of the E-RS scores during the Study Run-in Period (N=597)

E-RS Scores	N	Day -7 Mean(SD)	Day -1 Mean(SD)	Mean Difference[2] (SD)	Effect Size[3]	ICC[4]
RS-Total Score	584	18.3 (6.85)	17.6 (6.32)	0.6 (4.70)	0.10	0.74
RS-Breathlessness	584	9.0 (3.55)	8.7 (3.38)	0.3 (2.47)	0.10	0.74
RS-Cough & Sputum	584	4.7 (1.84)	4.5 (1.77)	0.2 (1.41)	0.10	0.69
RS-Chest Symptoms	584	4.6 (2.26)	4.5 (2.05)	0.1 (1.70)	0.07	0.69

[1] Pre-treatment, Visit 2, Day -7 to Day -1.
[2] Mean difference = average of (Day -7 E-RS score - Day -1 E-RS score).
[3] Effect size = Day -7 E-RS score - Day -1 E-RS score/SD of Day -1 E-RS score.
[4] Intraclass correlation coefficient (ICC) based on random-effect model.

Comments: The E-RS demonstrated acceptable reliability for the overall score across the two AZ clinical trials.

7.4 Description of Factor Analysis

Confirmatory factor analysis (CFA) was conducted to determine if the 11 respiratory symptom candidate items of the E-RS comprised a single underlying factor in patients with stable COPD. CFA was conducted as a secondary analysis using the data from the "Stable Group" from the prospective observational study conducted for the development and validation of the EXACT. Patients in the Stable Group were required to have one or more acute exacerbations within the past 24 months, but patients were excluded if they had experienced COPD exacerbation 60 days prior to enrollment.

The item threshold for potential deletion was a standardized coefficient < 0.30. The model fit was assessed using the comparative fit index (CFI), root mean square error approximation (RMSEA), and average weighted correlation residuals (SRMR) with the cut-offs for consideration of good fit as CFI>0.95, RMSEA<0.5, and SRMR<0.8.

In order to evaluate any potential underlying factors if the CFA demonstrated that the factor structure was not unidimensional, an exploratory factor analysis was planned with the 11 candidate respiratory items from Day 1. The assessment was conducted by evaluating the Scree plot (Figure 2) with their respective Eigen values to determine the number of factors in E-RS with model fit examined through the root mean square residual (RMSR) and RMSEA with an a priori threshold of less than 0.5, as seen in Table 14.

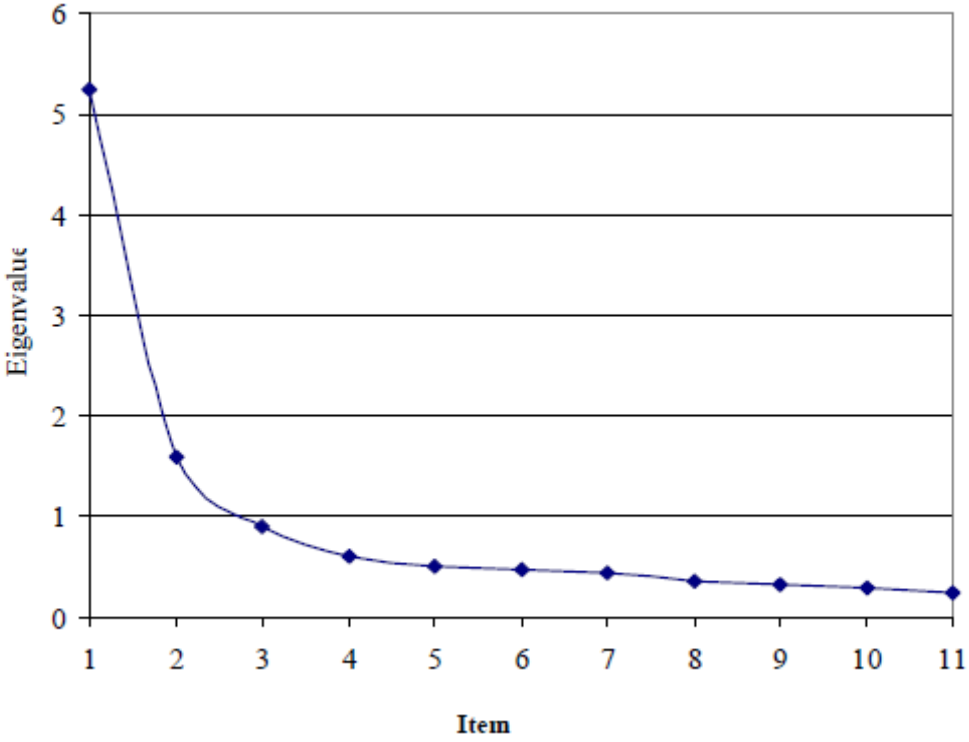


Figure 2. Psychometric study: Scree Plot of Post-hoc EFA of the 11 E-RS Items

Comments: The submitter's choice of three factors for the instrument seems reasonable based on the Scree plot in Figure 2.

Table 14 Psychometric study CFA: Standardized Coefficient for E-RS Items (Stable Group)

Item Number	Factor 1
Item 1 - Chest feel congested	0.673
Item 5 - Have chest discomfort	0.561
Item 6 - Chest feel tight	0.646
Item 2 - How often cough	0.696
Item 3 - How much mucus when cough	0.646
Item 4 - Difficult bring up mucus	0.670
Item 7 - Breathless today	0.740
Item 8 - Describe how breathless	0.594
Item 9 - Short of breath – personal care	0.675
Item 10 - Short of breath – indoor	0.647
Item 11 - Short of breath – outdoor	0.603

Model Fit Statistics*	
CFI	0.747
RMSEA	0.172
SRMR	0.104

*CFI greater than 0.95 will be considered a good fit, as well as RMSEA < 0.05, and SRMR < 0.08.

Comments: The correlations as seen in Table 14 demonstrate that there is moderate correlation if only 1 factor is used.

The results from the confirmatory factor analysis demonstrated that assumption of uni-dimensionality was not met, but the submitter indicated that the difference between E-RS and EXACT was expected due to the different time course of COPD (stable vs. exacerbated).

Post-hoc exploratory analysis with the entire sample from the Stable Group on Day 1 without pre-specifying the number of factors resulted in a four factor solution that had best model goodness-of-fit ($\chi 2 = 22.37$ (DF = 17), P = 0.171, RMSEA = 0.041, and RMSR = 0.018). Due to the item content and group, the submitter decided that a three factor solution would be a better model ($\chi 2=44.24$ (DF = 25), P =0.010, RMSEA = 0.064, and RMSR = 0.027). The three factors found were very similar to EXACT, RS-Chest Symptoms (Factor 1), RS-Cough & Sputum (Factor 2) and RS-Breathlessness (Factor 3) with factor loadings for all items > 0.30, as seen in Table 15.

The developer notes, while the failure to show uni-dimensionality was different than the Rasch analysis of the EXACT, this difference was expected due to the different patient populations (stable vs. exacerbating patients) and smaller item pool.

Table 15 Post-Hoc Exploratory Factor Analysis: Promax Factor Loading for E-RS Items (Stable Group)

Item Number	Factor 3	Factor 2	Factor 1
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
Q2 – How often cough	0.017	**0.717**	0.156
Q3 – How much mucus when cough	0.004	**0.850**	-0.082
Q4 - Difficult bring up mucus	0.071	**0.447**	0.284
Q1 - Chest feel congested	0.006	0.290	**0.575**
Q5 - Have chest discomfort	-0.029	-0.043	**0.876**
Q6 - Chest feel tight	0.113	0.069	**0.622**

* $\chi 2$=44.24 (DF = 25), P =0.010;RMSEA=0.064, RMSR=0.027.

RS-Breathlessness

RS-Cough & Sputum

RS-Chest

Comments: The three factors identified appear reasonable based on the items loading onto each factor as well as from a content perspective.

7.5 Construct Validity

Construct validity is evidence that relationships among items, domains, and concepts conform to *a priori* hypotheses concerning logical relationships that should exist with other measures or characteristics of patients and patient groups.

For convergent validity, the measure should be highly correlated with a measure of the same or a similar concept. Evidence of discriminant validity requires that a measure does not correlate too highly with measures from which it is supposed to differ.

For known groups (discriminant) validity, we review evidence that the instrument can differentiate (differs as predicted) between clinically distinct groups.

Construct validity was demonstrated through the relationship between E-RS scores and established measures to assess health related quality of life (SGRQ-C), underlying airway obstruction (FEV1%), breathlessness severity with activity using the breathlessness, cough and sputum using the Breathlessness Cough and Sputum Scale (BCSS), and rescue medication use. The analysis was conducted in only the stable patients with both the overall E-RS score as well as the domain scores using Spearman's rank-order correlation.

<u>Evidence of Convergent Validity</u>

SGRQ-C

For both AZ studies, data from the E-RS daily and weekly mean total and domain score prior to treatment administration was being compared to the patient's baseline report using the St. George's Respiratory Questionnaire (SGRQ) (Cycle 1, Day 1) total and domain scores. Both AZ12 and AZ20 studies used the 1-year recall version of the SGRQ. The relationships were assessed via the Spearman's rank-order correlation coefficient. The submitter hypothesized that the correlations should be greater than 0.50 for both the E-RS total and domain scores compared with the SGRQ total score. It is also expected that in general, the respective domain scores with the SGRQ scores will be smaller than the correlations found with the E-RS total score. For ease of presentation, only the data for the E-RS daily score is presented below in Tables 16-17. There are similar scores between the daily and the weekly average method of score calculation.

Table 16. AZ 12: Correlations[1] between E-RS Daily and SGRQ-C Scores (N=749)

E-RS Scores[2]	SGRQ-C Domains[3]			
	Total (N=745)	Symptom (N=745)	Activity (N=745)	Impact (N=745)
RS-Total Score	0.47, p <.0001	0.39, p <.0001	0.42, p <.0001	0.42, p <.0001
RS-Breathlessness	0.49, p <.0001	0.37, p <.0001	0.47, p <.0001	0.42, p <.0001
RS-Cough & Sputum	0.37, p <.0001	0.38, p <.0001	0.28, p <.0001	0.34, p <.0001
RS-Chest Symptoms	0.35, p <.0001	0.31, p <.0001	0.31, p <.0001	0.32, p <.0001

[1] Spearman's rank-order correlation and p value.
[2] E-RS scores for Visit 2, Day -1 or last day of data before Visit 2 (Day -7 to Day -1).
[3] SGRQ-C scores for Visit 2, Day 1(baseline) collected prior to treatment.

Table 17. AZ 12: Correlations between E-RS Daily and SGRQ-C Item Scores (N=749)

E-RS Scores[2]	SGRQ-C Scores[3]		
	Item 1-Cough (N=745)	Item 2-Sputum (N=745)	Item 3-SOB (N=745)
RS-Total Score	0.22, p <.0001	0.15, p <.0001	0.26, p <.0001
RS-Breathlessness	0.17, p <.0001	0.09, p=0.01	0.30, p <.0001
RS-Cough & Sputum	0.28, p <.0001	0.25, p <.0001	0.15, p <.0001
RS-Chest Symptoms	0.18, p <.0001	0.13, p=.0005	0.21, p <.0001

[1] Spearman's rank-order correlation and p value.
[2] E-RS scores for Visit 2, Day -1 or last day of data before Visit 2 (Day -7 to Day -1).
[3] SGRQ-C scores for Visit 2, Day 1 (baseline) collected prior to treatment.

AZ20

Table 18. AZ20: Correlations Between E-RS Daily and SGRQ-C Scores (N=597)

E-RS Scores[2]	SGRQ-C Domains[3]			
	Total (N=586)	Symptom (N=596)	Activity (N=588)	Impact (N=593)
RS-Total Score	0.45, p <.0001	0.42, p <.0001	0.36, p <.0001	0.43, p <.0001
RS-Breathlessness	0.46, p <.0001	0.38, p <.0001	0.43, p <.0001	0.42, p <.0001
RS-Cough & Sputum	0.33, p <.0001	0.39, p <.0001	0.20, p <.0001	0.32, p <.0001
RS-Chest Symptoms	0.38, p <.0001	0.35, p <.0001	0.27, p <.0001	0.36, p <.0001

[1] Spearman's rank-order correlation and p value.
[2] E-RS scores for Visit 2, Day -1 or last day of data before Visit 2 (Day -7 to Day -1).
[3] SGRQ-C scores for Visit 2, Day 1(baseline) collected prior to treatment.

Table 19. Correlations[1] between E-RS Mean Weekly and SGRQ-C Scores (N=597)

E-RS Scores[2]	SGRQ-C Domains[3]			
	Total (N=586)	Symptom (N=596)	Activity (N=588)	Impact (N=593)
RS-Total Score	0.51, p <.0001	0.45, p <.0001	0.42, p <.0001	0.48, p <.0001
RS-Breathlessness	0.52, p <.0001	0.40, p <.0001	0.48, p <.0001	0.47, p <.0001
RS-Cough & Sputum	0.41, p <.0001	0.44, p <.0001	0.27, p <.0001	0.40, p <.0001
RS-Chest Symptoms	0.44, p <.0001	0.39, p <.0001	0.33, p <.0001	0.41, p <.0001

[1] Spearman's rank-order correlation and p value.
[2] Mean of scores for the seven days preceding Visit 2(Day -7 to Day -1). Data for ≥4 days is required to calculate baseline.
[3] SGRQ-C scores for Visit 2, Day 1(baseline) collected prior to treatment.

Comments: The two AZ clinical trials demonstrated borderline construct validity using SGRQ as the comparison instrument. The correlations between the E-RS daily total and domain scores with SGRQ scores were lower than expected and did not meet the pre-specified hypothesized value of 0.5 as indicated in the SAP (statistical analysis plan). There were slightly stronger correlations seen with the mean weekly average E-RS total and breathlessness domain scores, compared to the SGRQ Total Score. The correlations between mean weekly E-RS Total and RS-Breathlessness and SGRQ-C total scores slightly exceeded the expected value of 0.50.

The correlations were lower than expected even when evaluating the three specific questions in the SGRQ that corresponded to the particular domains of E-RS [similar between the AZ 12 and AZ 20 study (data not shown)]. It is likely the very long recall period (1-year) employed in the SGRQ impacted validity and reliability of the SGRQ. Thus, the lower than expected correlations could be in part explained by the differences in instrumentation between the E-RS, developed as a daily symptom diary, and the SGRQ, which was developed for a different purpose.

FEV1% predicted

Since FEV1% predicted is a measure of pulmonary function, the correlation of this measure with E-RS was calculated for descriptive purposes only. The submitter indicated that they did not expect that there would be a correlation since the E-RS is a symptoms measure and does not assess airflow obstruction. Their data confirmed their expectations that these two measures would be essentially uncorrelated (AZ12: -0.17 for the E-RS breathlessness domain and -0.14 for E-RS total score; AZ20: -0.16 for the E-RS breathlessness domain of E-RS and -0.11 for the E-RS total score) (data not shown).

Comments: The lack of correlation between FEV1% and E-RS scores is consistent with the submitter's *a priori* hypothesis.

BCSS score

Theoretically, those patients with higher E-RS domain and total scores would be expected to have a higher BCSS scores due to the similarity of the concepts being assessed by each instrument. For the AZ studies, Spearman-rank order correlations were used to evaluate the relationship between daily (Visit 2, Day -1) between and mean weekly (Visit 2, Day -7 to -1) RS-Total and RS-domain scores with the Breathlessness, Cough, and Sputum Scale (BCSS) total and domain scores at Visit 2, Day -1, as seen in Table 19. The submitter expected that the RS-Total and domain scores would have large correlations with the BCSS total and domain scores respectively.

Table 19 AZ 12: Correlations[1] between E-RS Daily Scores and Patient-measured eFEV1 2, ePEF2, Breathlessness, Cough and Sputum Scale (BCSS) Scores, and Rescue Medication Utilization3 at Day -1 4 (N=749)

E-RS Scores[5]	FEV_1 (N=735)	PEF (N=735)	BCSS Score Total (N=735)	Breathlessness (N=735)	Cough (N=735)	Sputum (N=735)	Rescue Medication Utilization (N=735)
RS-Total Score	-0.10, p=0.005	-0.11, p=0.003	0.84, p <.0001	0.75, p <.0001	0.67, p <.0001	0.69, p <.0001	0.39, p <.0001
RS-Breathlessness	-0.15, p <.0001	-0.15, p <.0001	0.73, p <.0001	0.74, p <.0001	0.55, p <.0001	0.55, p <.0001	0.38, p <.0001
RS-Cough & Sputum	-0.05, p=0.14	-0.06, p=0.11	0.84, p <.0001	0.58, p <.0001	0.77, p <.0001	0.76, p <.0001	0.35, p <.0001
RS-Chest Symptoms	-0.06, p=0.09	-0.06, p=0.09	0.71, p <.0001	0.63, p <.0001	0.55, p <.0001	0.62, p <.0001	0.32, p <.0001

[1] Spearman's rank-order correlation and p value.
[2] Mean of morning and evening FEV_1 and PEF measurements.
[3] Mean number of rescue medication inhalations for the day (#Morning puff+#PM puff)/2.
[4] All measurements for Visit 2, Day -1.

Table 20 **Correlations[1] between E-RS Mean Weekly Scores and Patient-measured eFEV$_{1,2}$, ePEF$_2$, Breathlessness, Cough, and Sputum Scale (BCSS) Scores$_3$, and Rescue Medication Utilization[4] (N=749)**

E-RS Scores[5]	FEV$_1$	PEF	BCSS Score				Rescue Medication
	(N=749)	(N=749)	Total (N=749)	Breathlessness (N=749)	Cough (N=749)	Sputum (N=749)	Utilization (N=749)
RS-Total Score	-0.10, p=0.01	-0.11, p=0.002	0.89, p <.0001	0.84, p <.0001	0.73, p <.0001	0.79, p <.0001	0.43, p <.0001
RS-Breathlessness	-0.14, p=0.000	-0.16, p <.0001	0.80, p <.0001	0.85, p <.0001	0.62, p <.0001	0.66, p <.0001	0.42, p <.0001
RS-Cough & Sputum	-0.06, p=0.09	-0.07, p=0.07	0.89, p <.0001	0.65, p <.0001	0.83, p <.0001	0.86, p <.0001	0.38, p <.0001
RS-Chest Symptoms	-0.04, p=0.26	-0.04, p=0.22	0.75, p <.0001	0.71, p <.0001	0.61, p <.0001	0.70, p <.0001	0.33, p <.0001

[1] Spearman's rank-order correlation and p value.
[2] Mean of morning and evening FEV$_1$ and PEF measurements, Day -7 to Day -1
[3] All measurements for Visit 2, Day -7 to Day -1.
[4] Mean number of rescue medication inhalations for the day (#Morning puff+#PM puff)/2, Day -7 to Day -1
[5] Mean of scores for the seven days preceding Visit 2(Day -7 to Day -1). Data for ≥4 days is required to calculate baseline.

Comments: The correlations among daily E-RS scores and BCSS scores are as expected. The daily E-RS Total Score was correlated with all three domains of the BCSS (r ranged from 0.67 to 0.84). The corresponding domain scores of the E-RS were also highly correlated between the E-RS and the BCSS. Correlations with the BCSS using the mean weekly scores were all higher than those observed using the daily E-RS Scores. AZ 12 showed consistent findings with AZ 20 (data not shown).

Overall Comments: Adequate construct validity was demonstrated and consistent with pre-specified hypotheses concerning relationships that should exist among measures.

7.6 Ability to Detect Change

Review an instrument's ability to detect change using data that compare change in PRO scores to change in other similar measures that indicate that the patient's state has changed with respect to the concept of interest. A review of the ability to detect change includes evidence that the instrument is equally sensitive to gains and losses in the measurement concept and to change at all points within the entire range expected for the clinical trial population.

The ability of an instrument to detect change influences the sample size for evaluating the effectiveness of treatment.

The following tables (Tables 22 and 23) shows the change in E-RS weekly total scores and domain scores according by patient subgroups defined by the following categories on their SGRQ scores.

Table 21 AZ 12: Change from Visit 2 (Day -7 to Day-1) to Visit 6: Statistical Indicators of Magnitude of Change for Mean Weekly E-RS Total and Subscale Scores by Patient-Change in SGRQ-C Total Score (N=749)

SGRQ-C Change	E-RS Scores[1]			
	RS-Total	RS-Breathlessness	RS-Cough & Sputum	RS-Chest
Improved(change ≤-4 points)				
N	331	331	331	331
Mean Change(SD) p value[2]	-2.7 (5.4) <.0001	-1.1 (2.8) <.0001	-0.8 (1.4) <.0001	-0.7 (1.7) <.0001
Mean % Change	-13.4	-9.62	-17.1	-7.94
Effect Size	-0.44	-0.35	-0.53	-0.39
No Change				
N	195	195	195	195
Mean Change(SD) p value[2]	-0.8 (4.6) 0.0128	-0.2 (2.4) 0.2132	-0.4 (1.2) <.0001	-0.2 (1.6) 0.0621
Mean % Change	0.91	4.93	-5.70	15.49
Effect Size	-0.14	-0.07	-0.27	-0.12
Worsened(change >4 points)				
N	185	185	185	185
Mean Change(SD) p value[2]	0.9 (5.9) 0.0312	0.5 (2.9) 0.0209	0.0 (1.7) 0.8301	0.4 (1.9) 0.0050
Mean % Change	10.16	11.61	8.00	29.33
Effect Size	0.16	0.16	0.02	0.22
p value from ANCOVA[3]	<.0001	<.0001	<.0001	<.0001

[1] 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 0-40; RS-Breathlessness scores range from 0-17; RS-Cough & Sputum scores range from 0 to 11; and RS-Chest scores range from 0 to 12
[2] From t-test.
[3] For comparison of differences in E-RS Scores from baseline to Visit 6 among the three groups adjusting for baseline scores.

Comments: The E-RS total score and the breathless domain appear to demonstrate greater ability to detect change (improvement) than the cough and sputum as well as the chest symptom domains, which is consistent with the other measures of validity and reliability. As mentioned earlier, the primary analysis of ability to detect change focuses on E-RS total score as this is the score that is being recommended in this qualification.

Table 22 AZ 20: Change from Visit 2 (Day -7 to Day-1) to Visit 6: Magnitude of Change for Mean Weekly E-RS Total and Subscale Scores by Patient-Change in SGRQ-C Total Score (N=597)

SGRQ-C Change	E-RS Scores[1]			
	RS-Total	RS-Breathlessness	RS-Cough & Sputum	RS-Chest
Improved(change ≤-4 points)				
N	264	264	264	264
Mean Change(SD) p value[2]	-3.4 (6.1) <.0001	-1.5 (3.1) <.0001	-1.1 (1.7) <.0001	-0.9 (1.8) <.0001
Mean % Change	-16.7	-15.0	-20.0	-14.6
Effect Size	-0.61	-0.50	-0.72	-0.51
No Change				
N	166	166	166	166
Mean Change(SD) p value[2]	-1.7 (4.4) <.0001	-0.7 (2.3) 0.0002	-0.6 (1.4) <.0001	-0.4 (1.6) 0.0008
Mean % Change	-6.96	-3.48	-10.5	-4.03
Effect Size	-0.26	-0.20	-0.37	-0.21
Worsened(change >4 points)				
N	124	124	124	124
Mean Change(SD) p value[2]	-0.4 (5.3) 0.3670	-0.0 (2.8) 0.8861	-0.3 (1.5) 0.0090	-0.1 (1.7) 0.7193
Mean % Change	1.60	5.75	-5.37	3.08
Effect Size	-0.07	-0.01	-0.21	-0.03
p value from ANCOVA[3]	<.0001	<.0001	0.0002	0.0002

[1] 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 0-40; RS-Breathlessness scores range from 0-17;RS-Cough & Sputum scores range from 0 to 11; and RS-Chest scores range from 0 to 12
[2] From t-test.
[3] For comparison of differences in E-RS Scores from baseline to Visit 6 among the three groups adjusting for baseline scores.

Comments: Patients who improved by 4 points on the SGRQ on average improved by 3.4 points on the E-RS total score. On the other hand, patients who worsened by 4 points on the SGRQ on average worsened by only 0.4 points on the E-RS total score. The SGRQ total score measures a broader concept than just respiratory symptoms. Therefore, it is unclear whether this reflects a lack of sensitivity to detect worsening on the part of the E-RS, or perhaps an SGRQ worsening of 4 points may not reflect the same degree of symptomatic change.

E-RS total score improvement was 2.7-3.4 across the AZ 12 and AZ 20 among patients whose SGRQ improved by 4 points providing preliminary evidence of an ability to detect change in the direction of improvement only.

Both AZ 20 and AZ 12 failed to meet their primary endpoints and the investigational drug failed to demonstrate efficacy. For this reason, it is not surprising that the symptom severity scores were only slightly improved. Additional analysis is needed in future clinical trials of effective products to demonstrate that the E-RS is sensitive to change.

8 Interpretation of Scores

Planning for interpretation of clinical trial results includes developing a responder definition for the context of use, i.e., the individual patient score change over a predetermined time period that should be interpreted as a treatment benefit.

A responder definition for the targeted patient population was not included in the submission; this should be evaluated in longitudinal studies and pre-specified before use of the E-RS as a primary or secondary endpoint in confirmatory clinical trials. Guidelines for interpretation of meaningfulness of changes in score at the individual patient level, both in the direction of improvement as well as in the direction of worsening, are needed.

In version 3 of the user manual, there is limited guidance to the user to how to handle missing data. In the analyses of the two AZ clinical trials, a minimum of 4 days of data are required to compute a mean weekly score. Daily E-RS Total scores of 0 are set to missing. The submitter states that this scoring rule is based on their previous validation work demonstrating that moderate to severe COPD patients will experience symptom(s) each day, and a score of zero on all 14 EXACT items is likely to represent a situation where in order to complete the diary quickly, the respondent did not accurately report their daily symptom(s).

The extent to which this assumption is true is unknown. This could be problematic for patients who have a true score of zero. However, it is anticipated that very few patients will have a true score of zero, therefore this concern is not viewed as a critical flaw.

If no diary entry exists for a given day, the E-RS Total score is set to missing. Information was not provided on the extent to which missing days of data would impact scores when they are aggregated as an average over multiple days and used as an endpoint. Therefore, additional information on how to establish endpoints in the setting of missing data is needed.

9 Language Translation and Cultural Adaptation

The E-RS was developed in the United States. The E-RS is not to be completed in isolation, but rather is self-administered by study participants as part of the EXACT daily diary, which is completed each evening just prior to bedtime.

A complete list of available translations as well as information on methodology of translation should be available from the submitter and described in the instrument user manual. FDA encourages careful adherence to good practices for translation and cultural adaptation as described in an ISPOR Task Force Report (Wild et al, 2005), including item definition, dual forward translation; reconciliation; dual back translation; back translation review; harmonization; in-person cognitive testing with COPD patients in each target country using a standardized interview script; analysis of cognitive testing results; clinician review as-needed to verify terminology; finalization; and dual proofreading. This methodology is to ensure that the translated versions of a PRO instrument are both conceptually equivalent to the source version (in this case English) and easily understood by the target population.

10 Data Collection Method

The E-RS is designed for self-administration as an electronic daily diary to be completed each evening just prior to bedtime. The E-RS is comprised of the first 11-items of the 14-item EXACT. A user manual has been developed for E-RS.

Comments: Since the majority of the testing has been conducted using an electronic platform, this qualification recommendation is similarly limited to the electronically administered version.

12 References

Global Initiative for Chronic Obstructive Lung Disease (GOLD). *Global Strategy for the Diagnosis, Management, and Prevention of COPD*. Dec 2009. Available at: http://www.goldcopd.org.

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. *Am Rev Respir Dis.* Jun 1992;145(6):1321-1327.

Jones PW, Quirk FH, Baveystock CM. The St. George‘s Respiratory Questionnaire. *Resp Med.* 1991;85(suppl):25-31.

Leidy NK, Schmier JK, Jones MK, Lloyd J, Rocchiccioli K. Evaluating symptoms in chronic obstructive pulmonary disease: validation of the Breathlessness, Cough and Sputum Scale. *Respir Med*. 2003;97 Suppl A:S59-S70 (2003a)

Leidy NK, Rennard SI, Schmier J, Jones MK, Goldman M. The breathlessness, cough, and sputum scale: the development of empirically based guidelines for interpretation. *Chest*. 2003;124(6):2182-91 (2003b)

Appendix A. Annotated E-RS for Raw Score Assignment

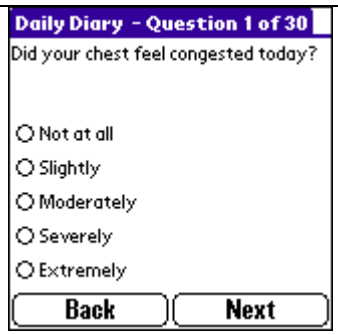
The following annotates the raw score values associated with each response category for the E-RS items. Please take note of items with collapsed response scale scoring, highlighted in **bold.** The E-RS is a subset of the 14 item EXACT.

1. Did your chest feel congested today?	0. Not at all
	1. Slightly
	2. Moderately
	3. Severely
	4. Extremely
2. How often did you cough today?	0. Not at all
	1. Rarely
	2. Occasionally
	3. Frequently
	4. Almost constantly
3. How much mucus (phlegm) did you bring up when coughing today?	0. None at all
	1. A little
	1. Some
	2. A great deal
	3. A very great deal
	NOTE: Score "a little" and "some" the same.
4. How difficult was it to bring up mucus (phlegm) today?	0. Not at all
	1. Slightly
	2. Moderately
	3. Quite a bit
	4. Extremely
5. Did you have chest discomfort today?	0. Not at all
	1. Slight
	2. Moderate
	3. Severe
	4. Extreme
6. Did your chest feel tight today?	0. Not at all
	1. Slightly
	2. Moderately
	3. Severely
	4. Extremely

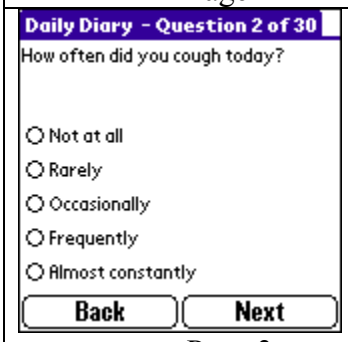
7. Were you breathless today?	0. Not at all
	1. Slightly
	2. Moderately
	3. Severely
	4. Extremely
8. Describe how breathless you were today:	0. Unaware of breathlessness
	1. Breathless during strenuous activity
	2. Breathless during light activity
	3. Breathless when washing or dressing
	3. Present when resting
	NOTE: Score "Breathless when washing or dressing" and "Present when resting" the same.
9. Were you short of breath today when performing your usual personal care activities like washing or dressing?	0. Not at all
	1. Slightly
	2. Moderately
	3. Severely
	3. Extremely
	4. Too breathless to do these
	NOTE: Score "severely" and "extremely" the same.
10. Were you short of breath today when performing your usual indoor activities like cleaning or household work?	0. Not at all
	1. Slightly
	2. Moderately
	3. Severely
	3. Extremely
	3. Too breathless to do these
	NOTE: Score "severely", "extremely", and "Too breathless to do these" the same.
11. Were you short of breath today when performing your usual activities outside the home such as yard work or errands?	0. Not at all
	1. Slightly
	2. Moderately
	3. Severely
	3. Extremely
	3. Too breathless to do these
	NOTE: Score "severely", "extremely", and "Too breathless to do these" the same.

Appendix B. EXACT (including E-RS) PDA Screenshots

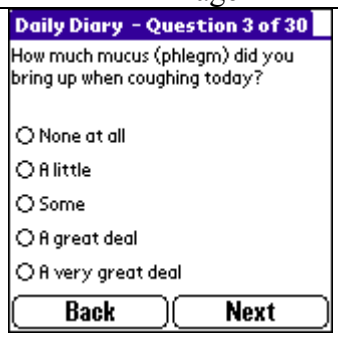
Page 1



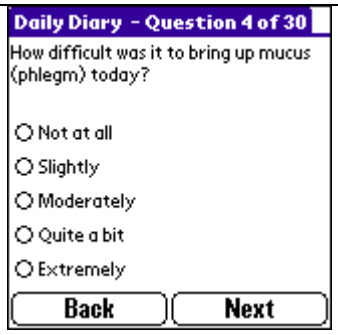
Page 2



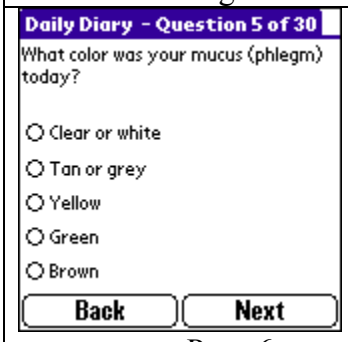
Page 3



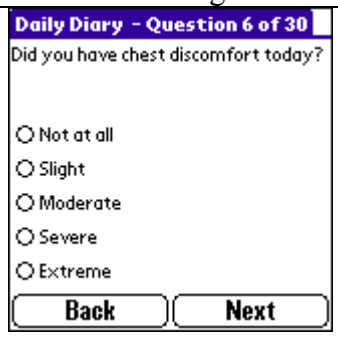
Page 4



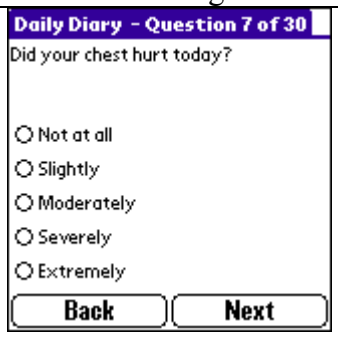
Page 5



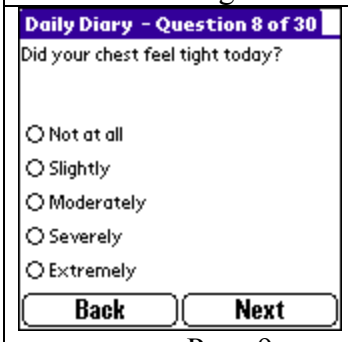
Page 6



Page 7

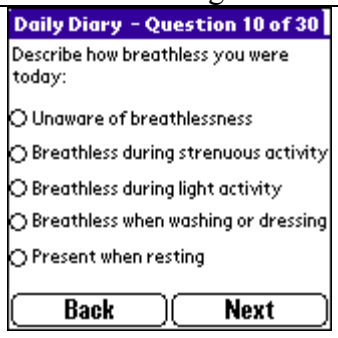


Page 8

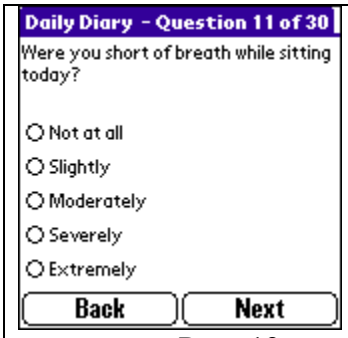


Page 9

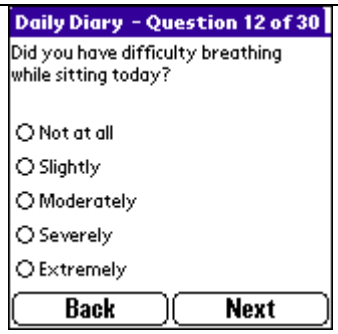
Page 10



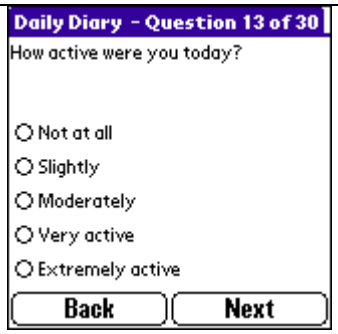
Page 11



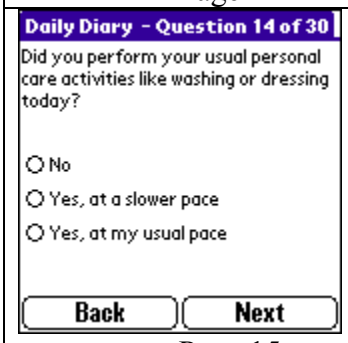
Page 12



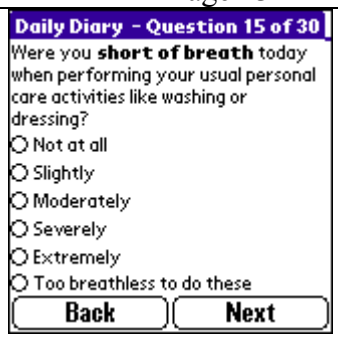
Page 13



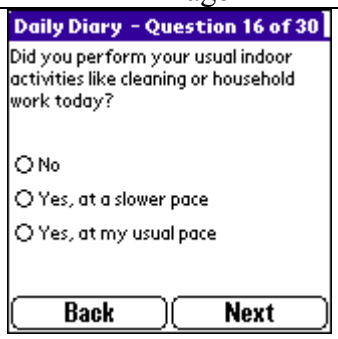
Page 14



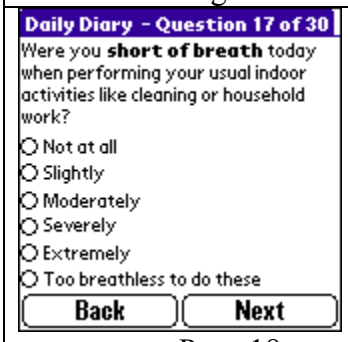
Page 15



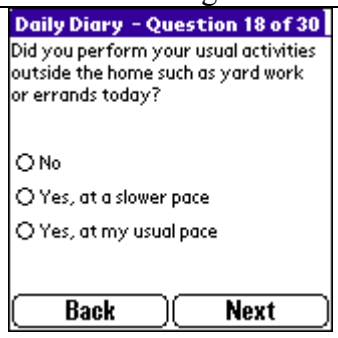
Page 16



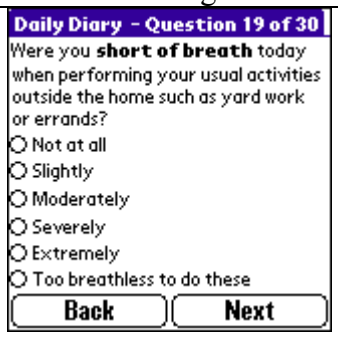
Page 17



Page 18

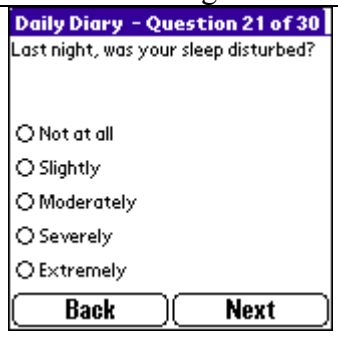


Page 19

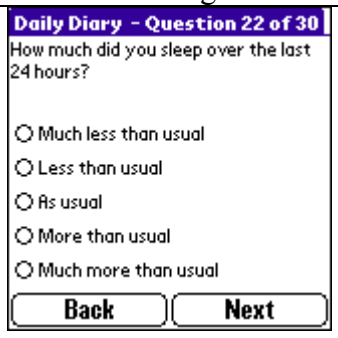


Page 20

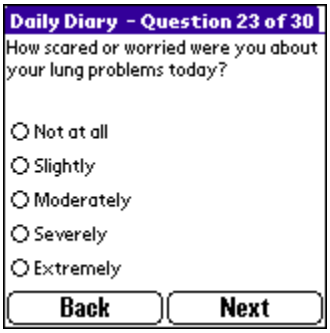
Page 21



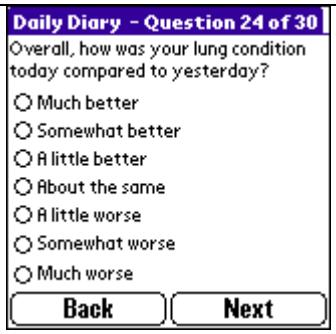
Page 22



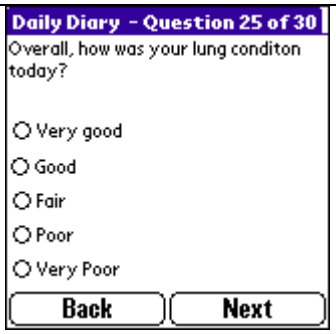
Page 23



Page 24



Page 25



Page 26

Page 27