HHS Public Access

Author manuscript

Clin Ther. Author manuscript; available in PMC 2017 April 01.

Published in final edited form as:

Clin Ther. 2016 April ; 38(4): 794–810. doi:10.1016/j.clinthera.2016.03.012.

Development of a Patient-Reported Outcome Instrument to Evaluate Symptoms of Advanced NSCLC: Qualitative Research and Content Validity of the Non-Small Cell Lung Cancer Symptom Assessment Questionnaire (NSCLC-SAQ)

Kelly P. McCarrier, PhD, MPH,
Health Research Associates, 6505 216thSt. SW, Suite 105, Mountlake Terrace, WA 98043, mccarrier@hrainc.net | Tel: 425-775-6565 Ext. 205 | Fax: 425-775-6734

Thomas M. Atkinson, Ph.D.,
Memorial Sloan Kettering Cancer Center, Dept. of Psychiatry and Behavioral Sciences, 641 Lexington Ave, 7th Floor, New York, NY 10022, atkinsot@mskcc.org

Kendra P.A. DeBusk, Ph.D.,
Senior Outcomes Research Scientist, Patient-Centered Outcomes Research, Genentech, 1 DNA Way, South San Francisco, CA 94080, debusk.kendra@gene.com

Astra M. Liepa, PharmD,
Principal Research Scientist, Global Patient Outcomes and Real World Evidence, Eli Lilly and Company, Indianapolis IN 46285 U.S.A., astra@lilly.com

Michael Scanlon, MA, and
Health Research Associates, 6505 216thSt. SW, Suite 105, Mountlake Terrace, WA 98043, scanlon@hrainc.net

Stephen Joel Coons, PhD on behalf of On behalf of the PRO Consortium's Non-Small Cell Lung Cancer Working Group
Critical Path Institute, Patient-Reported Outcome Consortium, 1730 E. River Rd., Tucson, AZ 85718, SJCoons@c-path.org

Correspondence to: Kelly P. McCarrier.

Conflict of Interest Disclosures

- Kendra P.A. DeBusk (Employee of Genentech, a member of the Roche group, and holder of Roche stock/stock options) and Astra M. Liepa (Employee and stockholder of Eli Lilly and Company) disclose their employment and other financial relationships with their respective firms.
- Kelly P. McCarrier, Thomas M. Atkinson, Michael Scanlon, and Stephen Joel Coons declare no conflicts of interest.

Author Contributions

All authors were involved in the conception and planning of the work that led to the manuscript, as well as interpretation of the data. KM, MS, and TA also led the collection and analyses of the study data. All authors participated in drafting or critical revisions of the manuscript and approval of the final submitted version.

Abstract

PURPOSE—To describe the process and results of the preliminary qualitative development of a new symptom-based PRO measure intended to assess treatment benefit in advanced non-small cell lung cancer (NSCLC) clinical trials.

METHODS—Individual qualitative interviews were conducted with adult NSCLC (Stage I–IV) patients in the US. Experienced interviewers conducted concept elicitation (CE) and cognitive interviews using semi-structured interview guides. The CE interview guide was used to elicit spontaneous reports of symptom experiences along with probing to further explore and confirm concepts. Interview transcripts were coded and analyzed by professional qualitative coders using Atlas.ti software, and were summarized by like-content using an iterative coding framework.

Data from the CE interviews were considered alongside existing literature and clinical expert opinion during an item-generation process, leading to development of a preliminary version of the NSCLC Symptom Assessment Questionnaire (NSCLC-SAQ). Three waves of cognitive interviews were conducted to evaluate concept relevance, item interpretability, and structure of the draft items to facilitate further instrument refinement.

FINDINGS—Fifty-one patients (mean age 64.9 [SD=11.2]; 51.0% female) participated in the CE interviews. A total of 1,897 expressions of NSCLC-related symptoms were identified and coded in interview transcripts, representing approximately 42 distinct symptom concepts. A 9-item initial draft instrument was developed for testing in three waves of cognitive interviews with additional NSCLC patients (n=20), during which both paper and electronic versions of the instrument were evaluated and refined. Participant responses and feedback during cognitive interviews led to the removal of 2 items and substantial modifications to others.

IMPLICATIONS—The NSCLC-SAQ is a 7-item PRO measure intended for use in advanced NSCLC clinical trials to support medical product labelling. The NSCLC-SAQ uses a 7-day recall period and verbal rating scales. It was developed in accordance with the FDA's PRO Guidance and scientific best practices, and the resulting qualitative interview data provide evidence of content validity. The NSCLC-SAQ has been prepared in both paper and electronic administration formats and a tablet computer-based version is currently undergoing quantitative testing to confirm its measurement properties and support FDA qualification.

Keywords

NSCLC; patient-reported outcome (PRO); content validity; qualitative research; scale development

1-Introduction/Background

Lung cancer is the leading cause of cancer-related mortality in the United States, with approximately 180,000 deaths expected to occur in 2015 [1]. Non-small cell lung cancer (NSCLC) is the most prevalent form of the disease and accounts for 85% of all lung cancers in the United States [2]. Early-stage NSCLC is often asymptomatic, or left undetected due to similar symptoms experienced by those with comorbid diseases (e.g., asthma, chronic obstructive pulmonary disease [COPD]) [3]. However, the degree of impairment that is experienced by patients with NSCLC is often impacted by the severity of their disease-

related symptoms. Therefore, accurate assessment and monitoring of these symptoms is an essential component when evaluating NSCLC treatment benefit in clinical studies [4].

Patient-reported outcomes (PROs), defined as the unfiltered subjective report of symptoms or health status by a patient, have been established as the "gold standard" for the capture of the patient symptom experience [5–7]. An increase in the assessment of PROs in clinical trials led the United States Food and Drug Administration (FDA) to release regulatory recommendations in its 2009 Guidance for Industry titled *Patient Reported Outcome Measures: Use in Medical Product Development to Support Labelling Claims* (hereinafter referred to as FDA PRO Guidance) [8]. The FDA PRO Guidance contains specific expectations for a given measures' psychometric properties, including conceptual framework, reliability, construct validity, and ability to detect clinically relevant score changes [8]. Most importantly, the FDA PRO Guidance recommends that content validity be established through the comprehensive qualitative elicitation of concepts from patients in the targeted disease population, as well as through cognitive interviewing to confirm respondent understanding of the PRO items assessing each measured concept. In addition, in 2014 the FDA released a *Guidance for Industry and FDA Staff on the Qualification Process for Drug Development Tools* (hereinafter referred to as FDA Qualification Guidance) [9]. Qualification, as defined by the FDA's Center for Drug Evaluation and Research (CDER), is a formal conclusion that the results obtained from the PRO instrument within a stated context of use can be relied upon to have a specific interpretation and application in drug development and regulatory review [9].

For NSCLC, a number of condition-specific PRO measures exist that capture disease-related symptoms, including the European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire – Lung Cancer (EORTC QLQ-LC13 [10]), Functional Assessment of Cancer Therapy – Lung (FACT-L [11]), Lung Cancer Symptom Scale (LCSS [12]), and M.D. Anderson Symptom Assessment Inventory – Lung Cancer (MDASI-LC [13]). Despite each of these measures being rigorously tested and widely-used, the development history, content, and comprehensiveness of these tools with respect to documenting symptom concepts that have been specifically elicited from first-hand accounts of the patients' experience with NSCLC may not necessarily satisfy the expectations of the FDA PRO Guidance. As such, the Critical Path Institute's (C-Path) PRO Consortium, with consultation from FDA advisors, identified the need for a well-defined and reliable PRO instrument to measure NSCLC symptoms and provide evidence for U.S. drug labeling.

To address this gap, the PRO Consortium established the NSCLC Working Group, with the objective of *qualifying* a PRO instrument to be used in assessing clinical benefit in advanced NSCLC clinical trials[14]. The NSCLC Working Group is comprised of pharmaceutical firm representatives and C-Path personnel. As part of a competitive bidding process, Health Research Associates (HRA) was awarded a contract to provide research services for the working group.

The development team for the NSCLC Symptom Assessment Questionnaire (NSCLC-SAQ) included members of the NSCLC Working Group and PRO measurement scientists from HRA, who employed rigorous methodological approaches similar to those used in the Major

Depressive Disorder Working Group's PRO instrument development efforts [15]. In addition, an advisory panel of clinical and measurement experts was engaged and the FDA convened a qualification review team (QRT) to review progress and provide input at each key PRO measure development milestone. The QRT is composed of representatives from FDA's Office of Hematology and Oncology Products, Clinical Outcome Assessment Staff, and Office of Biostatistics.

The purpose of this paper is to describe the initial steps in the development of the NSCLC-SAQ to assess disease-related symptom change in patients with NSCLC. The initial steps include: (1) the decision to develop a new PRO measure for NSCLC rather than select/modify an existing measure, (2) methodological steps/findings from concept elicitation (CE) interviews, including clinical input and item generation, (3) development of a preliminary version of the NSCLC-SAQ, and (4) findings from cognitive interviews and resulting modifications to the NSCLC-SAQ.

2-METHODS

2.1-Study Design and Development Steps

The major activities involved in the qualitative development of the initial draft of the NSCLC-SAQ and the evidence supporting its content validity are detailed below (see Figure 1). Preliminary quantitative evaluation of the NSCLC-SAQ's item performance is currently in progress and will be reported separately upon completion.

For the first developmental step, two distinct systematic literature reviews were conducted. The first examined available peer-reviewed qualitative research in NSCLC to identify symptom concepts and domains noted as relevant from the patient perspective. This literature review established a preliminary set of disease-defining NSCLC symptom concepts to be examined alongside concepts arising directly from NSCLC patients in qualitative concept elicitation efforts.

The second literature review examined evidence for previously-published NSCLC-targeted PRO instruments to assess their potential suitability for FDA qualification, and to identify potential content that could be considered during the construction of a new instrument if needed. The PRO instruments identified in the review varied in recall period, specific concepts assessed, questionnaire length, response scales, anchoring, and scoring algorithms. Given the limited information available in the published literature regarding the level and extent of direct patient involvement in item development, there was concern that the existing PRO measures were unlikely to meet the expectations described in the FDA PRO Guidance [8]. Therefore, documented qualitative work that is consistent with the FDA PRO Guidance, as well as the scientific best practices put forth by the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) PRO Good Research Practices Task Force [16, 17] was deemed necessary to determine if adequate evidence of content validity was available for any of the NSCLC legacy measures.

The literature and instrument reviews informed the development of the study protocol and interview guide for qualitative research, adhering to the methodological expectations set

forth in the FDA Guidance. Quorum Review IRB (Seattle, WA) and Memorial Sloan Kettering Cancer Center's institutional review board reviewed and approved the study protocol, recruitment forms and qualitative interview materials. All study participants provided written informed consent prior to participation in study activities. Individual qualitative CE interviews were conducted to identify NSCLC-specific symptom concepts relevant to patients diagnosed with the condition, documenting the terms and language used by patients to describe their symptoms, and exploring the severity, frequency, and duration with which the symptoms are experienced.

Following transcript coding and analysis of qualitative interview data, a determination was made by the working group that development of a novel PRO instrument for NSCLC was the best path forward. While several existing measures contained many of the core symptom concepts, no single instrument was deemed adequate for achieving the working group's goals; all would have required some degree of alteration. In addition, comments received from the QRT led the working group to believe that FDA's very specific expectations regarding supporting evidence of content validity could only be met by a carefully designed PRO instrument development process. Therefore, the qualitative data were utilized in an item generation process to develop the initial draft of the NSCLC-SAQ.

This initial draft of the NSCLC-SAQ was further tested and refined via three iterative waves of cognitive interviews to evaluate respondent understanding, acceptability of formatting, clarity of instruction text, appropriateness of recall period, and suitability of response options. During the second and third wave of cognitive interviews, an electronic data capture (ePRO) format of the draft NSCLC-SAQ was used alongside the paper version, and additional interview exercises explored the equivalence of the paper and ePRO versions. The working group's advisory panel and FDA's QRT were consulted during this process.

2.2-Study Participants

The qualitative study sought to enroll a diverse sample of subjects with NSCLC; including participants with early-stage disease (Stage I–II) as well as those with advanced-stage disease (III–IV), with representation of major histological subtypes of NSCLC (i.e., squamous cell carcinoma and adenocarcinoma). Although the objective was to develop an instrument primarily for use in advanced disease, all stages were included so as to document a continuum of symptoms. Eligibility criteria for participants in CE and cognitive interviews were identical, and reflected common entry criteria for clinical trials testing treatments for NSCLC. Specifically, the study included female and male adult (≥18 years) participants with Eastern Cooperative Oncology Group (ECOG) performance status of 0–2, representing the common eligibility criteria for clinical trials of advanced NSCLC intended for regulatory approval. Study participants diagnosed with Stage I or II cancer were required to be naïve to treatment for NSCLC; and those diagnosed with Stage III or IV cancer were required to be either treatment-naïve or to have recovered from any prior treatment-related toxicity/adverse events to Common Terminology Criteria for Adverse Events (CTCAE) v4.03 grade 1 (mild) or better [18].

All participants were required to be capable of reading, writing, and speaking English at a level allowing them to provide written informed consent and actively contribute in an

interview. To assure that the concepts elicited were related to NSCLC, individuals with current or past history of a personality disorder, bipolar disorder, schizophrenia or other psychotic disorder, obsessive compulsive disorder, post-traumatic stress disorder, or mental retardation were excluded. Also excluded were persons with a recent history (within 12 months) of significant alcohol or drug abuse; concurrent (or within the previous 30 days) participation in an investigational drug, device, or biologics product study; or other significant medical comorbidities with the potential to confound patients' description of their NSCLC symptoms (COPD was allowed, as detailed below).

Oncology clinic sites in six US states (Alabama, Idaho, Illinois, Montana, New York, and North Dakota) recruited subjects between May and December 2013. Recruitment quotas were employed to ensure representation within NSCLC stage and ECOG performance status groups, as well as to ensure appropriate representation of patients with clinically-important comorbidities, such as COPD. Each clinic site aimed to enroll patients with diverse NSCLC treatment histories and a broad range of demographic characteristics (sex, age, ethnicity, race, educational attainment, marital status, and employment status).

2.3-Concept Elicitation (CE) Interviews

The study protocol and qualitative interview guide were based on findings from the literature reviews, a hypothesized conceptual framework, and input from the expert advisory panel (see *Acknowledgments*). CE interviews were conducted by trained qualitative research staff and took place in private rooms within each clinic. In total, five different experienced qualitative interviewers conducted the CE interviews, and three of those conducted the cognitive interviews.

Utilizing the semi-structured CE interview guide, interviewers employed open-ended questions and day-reconstruction exercises to elicit spontaneous accounts of NSCLC-related symptom concepts. Following these open-ended inquiries, targeted probes were used to assess concepts not spontaneously reported by interview participants. Interview probe wording was based on concepts identified during the systematic review of NSCLC literature. Interview subjects were asked to rate the severity and level of bother or difficulty associated with each symptom they reported.

2.4-Analysis of Qualitative Data

Audio recordings of interviews were transcribed and independently reviewed by experienced qualitative coders to identify patient-expressed concepts. Employing an iterative coding framework, code assignment was conducted using ATLAS.ti™ software [19] to assist coders in tagging concepts, and facilitated the grouping of concepts with other codes of similar content to identify predominant patient expressions.

Consistency in the assignment of concept codes was evaluated through analyses of inter-rater agreement. A random selection of ten percent of transcripts were independently coded by two members of the coding team and compared to assess variation in code assignment. Consistency of coding was characterized by 1) agreement in concept identification, and 2) agreement in code assignment for each identified concept.

To assess concept saturation, transcripts were ordered chronologically and split into six groups for sequential examination. Codes reported for each subsequent group of transcripts were compared to codes applied in the prior group to identify the point at which no novel coded symptom concepts (thus no new concept-level information) were observed. The data collection and analysis techniques used were based on current best practice recommendations for establishing evidence of content validity for PRO instruments intended for use in the evaluation of medical products. [16, 20].

2.5-Determination of Measurement Strategy and Process for Item Generation

The core development team and the expert panel reviewed the CE interview results and determined whether to pursue qualification of an existing measure or develop a new instrument. CE interview findings were reviewed alongside concepts identified during the review of published literature and existing PRO measures to inform selection of NSCLC symptom concepts for inclusion in PRO measurement. This review of data considered the overall intent for the final PRO instrument to accurately assess treatment-related changes in clinically-meaningful and patient-relevant symptoms of NSCLC, with sufficient evidence to enable PRO instrument qualification for use in supporting label claims for drug products in the U.S. The targeted concepts were then cross-referenced against the content coverage of existing legacy PRO measures in the NSCLC therapeutic area to determine if a previously-developed instrument could meet the working group's PRO measurement needs.

Patient language in the CE interview transcripts was used to construct the wording of draft questionnaire items for each targeted concept. During this process of concept selection and drafting of item text, the development team considered the appropriateness of each potential item based on key criteria, including: 1) relevance to patients with NSCLC, as determined by the frequency with which the item was expressed in the CE interviews, particularly when mentioned spontaneously; 2) patient ratings of importance or bother and/or other sources of qualitative patient-based evidence indicating relevance; 3) the item assesses a single, rather than multidimensional, symptom; 4) the item is written using words and phrases commonly expressed and understood by patients with NSCLC, as informed by the transcripts from CE interviews; 5) the core development team and expert panel agree the item will likely be sensitive to symptomatic change occurring from treatment for NSCLC; 6) the item is unlikely to be vulnerable to floor or ceiling effects among patients with NSCLC; 7) the item is likely to have conceptual or semantic equivalence in other languages; and 8) the recall period used by the item is appropriate given the anticipated rate of symptom change experienced by patients. In subsequent steps, the targeted concepts and preliminary item text was further refined to address synonymous/duplicative concepts. The formatted initial version of the NSCLC-SAQ was prepared for evaluation via cognitive interviews and the translatability assessment process.

2.6-Cognitive Interviews and Instrument Refinement

Cognitive interviews were conducted to evaluate concept relevance and comprehensiveness; as well as the understandability of the wording and structure of the draft item stems, response options, and instructions to support subsequent refinement of the instrument. Participants were recruited from two of the participating clinical sites for cognitive

interviews. Each face-to-face cognitive interview lasted 60 to 90 minutes. Interviews began with participants self-administering the NSCLC-SAQ, after which they were asked a series of interview questions crafted to gain insight into the cognitive process undertaken by respondents with each questionnaire item.

An interview guide was used to standardize the semi-structured cognitive interviews, and utilized a think-aloud process to evaluate the items of the draft instrument. Interview questions assessed comprehension and relevance of the individual items; fit of the response scales; appropriateness of the recall period and item wording; and helped to identify any lack of clarity in instructions, item terminology, or sentence structure. To evaluate the comprehensiveness of the included concepts, interview items also asked whether participants experienced any other NSCLC symptoms that they felt were missing from the draft instrument. For some symptom concepts, participants were presented alternate item stems to consider, which used modified references to the recall period or different phrasing of the symptom concept.

After the first wave of cognitive interviews, instrument modifications, translatability assessment (TA), and input from the advisory panel, a revised draft NSCLC-SAQ was programmed for self-completion on an ePRO tablet device. Exercises were added to the cognitive interview guide to evaluate the conceptual equivalence of the two administration formats (ePRO and paper) during the remaining cognitive interviews. These exercises focused on a comparison of the “think aloud” narratives provided by participants for items administered in each format, and through direct probing to identify differences in the thought process, understanding of the concept, or response selection between the two formats.

Transcripts were prepared from the cognitive interview audio recordings, and were used to construct summary tables of participant quotes employed in evaluating the NSCLC-SAQ. Three iterative waves were conducted, with 4 to 10 interviews completed in each. After each subsequent wave, the core development team reviewed the interview findings and refined the NSCLC-SAQ. Parallel with the overall cognitive interview process, experienced PRO linguistics consultants executed a TA in five languages (Chinese, Hindi, Japanese, Russian, and Spanish) to evaluate the potential for difficulty in maintaining conceptual equivalence when translating each item. These five languages were selected as both representatives of key language families [21] and likely languages of need for the clinical trial programs of the NSCLC Working Group member firms. Findings from the TA facilitated revisions to items ahead of the completion of the cognitive interview phase.

3-RESULTS

3.1-Concept Elicitation Findings

3.1.1-Demographic and Clinical Characteristics—CE interviews were conducted with 51 participants. The average age of the participants was 64.9 (range 46–86), and 51.0% were female (Table 1). Three-quarters (74.5%) of CE interview participants were White, 15.7% were Black/African American, and 9.8% reported being of Hispanic/Latino ethnicity.

At the time of their interview, the majority of participants (51.0%) had Stage IV NSCLC; and 35.3% had a diagnosis of comorbid COPD.

3.1.2-Content Analysis Results—Analysis of the interview transcripts resulted in 1,897 coded symptom expressions, grouped into 43 different concepts based on content and similarity of patient expression, within five hypothesized symptom sub-domains. Inter-rater agreement on five dual-coded transcripts showed 95.9 to 98.9% agreement between the two coders. Given the 5,837 individual codes that were assigned across the 51 transcripts, these results can be interpreted as a high level of agreement between the coders.

Concept saturation was achieved after the 27th of the 51 coded transcripts (i.e., no novel concepts were observed after the third of six transcript groups). In the first group of nine transcripts, 40 (93%) of the coded concepts arose. Two additional concepts arose in the second transcript group, and the last coded concept appeared in the third transcript group. No new information was provided by the three remaining transcript groups, suggesting that additional concepts are unlikely to emerge from additional CE interviews and that the 51 interviews were adequate to achieve comprehensiveness of concepts in this target population.

3.1.3-Selection of Concepts and Generation of Items for the NSCLC-SAQ—Key findings from the literature reviews and the CE interviews, along with input from the advisory panel, were reviewed to select the symptom concepts to be considered for inclusion in the NSCLC-SAQ. To identify the most strongly-supported symptom concepts from the qualitative interviews, the following factors were considered: the overall number of coded expressions of a given symptom concept within the interview transcripts, the number of participants expressing each symptom concept, the percent of participants offering each concept spontaneously (rather than as a result of specific probing by the interviewer), the severity ratings for each symptom concept, and the bothersome ratings for each symptom concept. Table 2 presents this concept-level information used by the development team to help guide the concept selection process.

This process resulted in the selection of nine symptom concepts from the initial set of 40. Table 3 presents the nine selected symptom concepts along with the key findings from the qualitative interview data and examples of interview participants' quotes.

After reaching consensus on the list of selected concepts, item wording was developed for each concept based on the interview participants' quotes in order to form an initial draft of the NSCLC-SAQ. At this stage, two parallel items were constructed for each symptom concept; one using a five-point verbal rating scale (VRS), and another using an 11-point (0 to 10) numeric rating scale (NRS).

The working group chose a 7-day recall period for the NSCLC-SAQ. This decision was based on participant responses during the CE interviews supporting a one-week period, the recall period employed by existing NSCLC-focused symptom measures, advice from the advisory panel and the QRT, and a desire to avoid the additional burden on respondents associated with a daily symptom diary. A 7-day recall period has been repeatedly shown to

be equivalent to 24-hour/daily reporting of PROs across multiple disease types and settings [22–24].

The draft items were reviewed by the working group and advisory panel members. Recommended changes were proposed and adjudicated, and a preliminary version of the NSCLC-SAQ was constructed for the first wave of cognitive interviews.

3.2-Evaluation and Refinement of the Preliminary NSCLC-SAQ

3.2.1-Preliminary (Wave 1) Cognitive Interviews—The initial set of cognitive interviews included 4 participants. Participant characteristics are described in Table 1. Participant responses during the first wave of interviews expressed at least some difficulty with the NRS version of 8 of the 9 items. Based on these findings, the working group decided to proceed with the VRS for subsequent cognitive interview waves.

Participant responses led to several other minor refinements to the VRS version of the instrument in order to increase the clarity of the items as follows: 1) both the stem and the response options of the first three items (general pain, chest pain, and cough) were revised to clarify the focus on the intensity/magnitude of the symptom being assessed (as opposed to the frequency), 2) the reference to the recall period “last 7 days” was included at the end of each item stem, in order to maintain consistency, 3) the cough item was moved into the first position in sequence so the NSCLC-SAQ began with a chest symptom item rather than a general pain item, and 4) the general pain item was replaced by an item assessing pain “in areas other than your chest” in order to serve as a mutually-exclusive complement to the item assessing pain in the chest.

The revised version of the NSCLC-SAQ that utilized the VRS and contained the updated wording was presented to the expert advisory panel for review and discussion with the development team. During these discussions, the development team confirmed the use of the VRS format and considered the deletion of the hemoptysis item. Discussions from the expert panel focused on two primary reasons for suggesting deletion of this item. First, among the first four patient interviews, only one subject recognized having ever coughed up blood, which confirmed the clinical experts’ opinion that the symptom is experienced infrequently by few patients and is unlikely to be sensitive to treatment effects. Therefore, it was decided by the working group to remove this item from the instrument prior to the Wave 2 interviews and TA.

3.2.2-Wave 2 and 3 Cognitive Interviews and ePRO Evaluation—During the second wave, the 8-item NSCLC-SAQ was evaluated with 10 participants (Table 1). Participant responses during the Wave 2 interviews supported the overall relevance of the included concepts, provided evidence of conceptual equivalence between the paper and ePRO formats of the instrument, and facilitated refinement of the wording for several items. Specifically, the development team made the following key changes based on interview findings and input from the QRT: 1) the three severity/intensity-focused items (cough, chest pain, non-chest pain) were reworded to assess the peak (“worst”) intensity of the symptom, 2) the two dyspnea-focused items were combined to result in a single item that assesses the frequency of feeling “short of breath during usual activities,” 3) the appetite-focused item

was reworded from assessing “good appetite” to “poor appetite” to allow the response options to remain directionally consistent with the other items in the instrument.

The revised, 7-item NSCLC-SAQ was further evaluated through 6 additional cognitive interviews (Wave 3). Subjects in Wave 3 confirmed the relevance of items, expressed no difficulty with comprehension of the items or response options, and noted no noteworthy differences in meaning or response between the paper and ePRO format of the instrument.

Therefore, the finalized preliminary PRO instrument, the NSCLC-SAQ, is a 7-item instrument that measures each concept using a 5-point VRS. The instrument specifies a 7-day retrospective recall period for each of the items. Three of the items focus on the peak intensity of symptoms with a rating scale from “no [concept] at all” to “very severe.” The remaining 4 items focus on the frequency or the amount of time a symptom was experienced and employ a rating scale from “never” to “always.” Examples of each of these two item types are presented in Figure 2. This preliminary NSCLC-SAQ is currently undergoing additional testing through a quantitative pilot study designed with input from the QRT to provide data to support individual item analysis and the initial assessment of measurement properties.

4-DISCUSSION

In the FDA PRO Guidance, content validity is defined as the extent to which a measure appropriately and comprehensively captures all aspects of the concept to be measured relative to the intended context of use [8]. The FDA considers direct, unfiltered input from the targeted patient population as an essential component of establishing the content validity of a PRO measure to be fit for purpose for medical product labeling claims [8]. The present study describes our rigorous efforts to qualitatively establish the content validity of a PRO instrument that satisfies FDA recommendations to be qualified for use as a primary endpoint measure to assess treatment benefit in advanced NSCLC clinical trials.

Based on the findings from the literature review, qualitative evidence collected during the CE interviews, and input from clinical experts, concept saturation was sufficiently achieved. Subsequent cognitive interviews and feedback from the QRT led to refinements in item content and instructions to ensure that the NSCLC-SAQ assesses symptoms that are important to patients with advanced NSCLC and has response options that fit the way patients think about the severity of those symptoms. Participants of the cognitive interviewing phase considered the 7-day recall period to be appropriate. However, a daily symptom diary may be useful in some contexts, particularly where a treatment benefit claim may be linked to time (e.g., time to symptom relief).

This study carries with it a number of limitations. While only six subjects with Stage I NSCLC were recruited for the CE phase, and none with Stage II NSCLC, the disease-related concepts identified by subjects with advanced-stage disease (i.e., Stage III–IV) encompass those that were elicited from subjects with Stage I NSCLC [25]. It is therefore unlikely that additional concepts would have been captured through additional interviews with subjects with early-stage disease (i.e., Stage I–II). Additionally, during Waves 2 and 3 of cognitive

interviewing, the ePRO version of the NSCLC-SAQ was only tested using a handheld tablet device. While other widely used ePRO devices such as laptops and mobile phones were not tested, the single-item-per-screen formatting of the tablet device was a suitable approximation of a device-neutral platform for screen-based electronic administration of the measure.

The preliminary version of the NSCLC-SAQ contains 7-items that address the clinically relevant core symptoms of NSCLC that patients deem important and for which effective relief would be meaningful. Efforts are underway to complete the next steps in FDA's Clinical Outcome Assessment (COA) qualification process: collection of quantitative evidence to refine and confirm item content, exploration of response scale distribution anomalies and potential subscale structure, as well as establishment of key psychometric properties, including internal consistency, test-retest reliability, construct validity, and clinical responsiveness. Additionally, guidelines for interpreting and defining clinically meaningful NSCLC-SAQ score changes will be established.

The patient-centered approach to establishing the content validity of the NSCLC-SAQ ensures the instrument has the potential to accurately capture the patient-reporting of treatment benefits in NSCLC clinical trials. Upon completion of quantitative testing, the final version of the NSCLC-SAQ will be submitted to the FDA for the purposes of review for qualification. Once qualified, the NSCLC-SAQ will be publicly available to capture patient-reported NSCLC-related symptoms via electronic data entry platforms. Although not encouraged for use in assessing a clinical trial endpoint, a paper version of the instrument will also be available

Acknowledgments

This research was funded by the following members of the PRO Consortium: Abb Vie, Boehringer Ingelheim, Bristol-Meyers Squibb, Eli Lilly and Company, Genentech, Merck Sharp & Dohme Corp., and Novartis Pharmaceuticals. Critical Path Institute's PRO Consortium is supported by grant No. U01FD003865 from the United States Food and Drug Administration. This project was also supported by a National Institutes of Health Support Grant (P30 CA08748-49), which provides partial support for the MSK Behavioral Research Methods Core Facility used in conducting this investigation. A portion of these results were presented at the 2015 Conference of the International Society for Quality of Life Research in Vancouver, BC.

The authors would like to thank David Cella, PhD; Richard Gralla, MD; Donald L. Patrick, PhD; and Suresh Ramilingham, MD for providing their clinical experience and insight to the development process as members of the expert advisory panel. We would also like to thank Mona Martin, Carla Ascoytia, and Lisa Peterson for their valuable contributions to study design, data collection, and analysis efforts.

REFERENCES

1. ACS. Cancer Facts and Figures. Atlanta: American Cancer Society; Available from: http://www.cancer.org/acs/groups/content/@editorial/documents/document/acspc-044552.pdf [Accessed December 9, 2015]
2. Molina JR, Yang P, Cassivi SD, Schild SE, Adjei AA. Non-small cell lung cancer: epidemiology, risk factors, treatment, and survivorship. Mayo Clin Proc. 2008 May.83:584–594. [PubMed: 18452692]
3. Ironmonger L, Ohuma E, Ormiston-Smith N, et al. An evaluation of the impact of large-scale interventions to raise public awareness of a lung cancer symptom. Br J Cancer. 2015 Jan 6.112:207–216. [PubMed: 25461805]

4. Masters GA, Temin S, Azzoli CG, et al. Systemic therapy for Stage IV non-small-cell lung cancer: American Society of Clinical Oncology Clinical Practice Guideline Update. J Clin Oncol. 2015 Oct 20.33:3488–3515. [PubMed: 26324367]
5. Basch E. The missing voice of patients in drug-safety reporting. N Engl J Med. 2010 Mar 11.362:865–869. [PubMed: 20220181]
6. Basch E. Beyond the FDA PRO Guidance: Steps toward integrating meaningful patient-reported outcomes into regulatory trials and US drug labels. Value Health. 2012 May.15:401–403. [PubMed: 22583448]
7. Trotti A, Colevas AD, Setser A, Basch E. Patient-reported outcomes and the evolution of adverse event reporting in oncology. J Clin Oncol. 2007; 25:5121–5127. [PubMed: 17991931]
8. US Department of Health and Human Services. [Accessed January 5, 2016] Guidance for Industry. Patient-reported outcome measures: Use in medical product development to support labeling claims. 2009 Dec. Available from: http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM193282.pdf
9. US Department of Health and Human Services. [Accessed December 9, 2015] Guidance for Industry and FDA Staff. Qualification Process for Drug Development Tools. 2014 Jan. Available from: http://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm230597.pdf
10. Bergman B, Aaronson NK, Ahmedzai S, Kaasa S, Sullivan M. The EORTC QLQ-LC13: a modular supplement to the EORTC core quality of life questionnaire (QLQ-C30) for use in lung cancer clinical trials. Eur J Cancer. 1994; 30:635–642. [PubMed: 8080679]
11. Cella DF, Bonomi AE, Lloyd SR, et al. Reliability and validity of the Functional Assessment of Cancer Therapy—Lung (FACT-L) quality of life instrument. Lung Cancer. 1995; 12:199–220. [PubMed: 7655830]
12. Hollen PJ, Gralla RJ, Kris MG, Potanovich LM. Quality of life assessment in individuals with lung cancer: testing the Lung Cancer Symptom Scale (LCSS). Eur J Cancer. 1993; 29A(Suppl 1):S51–S58. [PubMed: 8381294]
13. Mendoza TR, Wang XS, Lu C, et al. Measuring the symptom burden of lung cancer: the validity utility of the lung cancer module of the M. D. Anderson Symptom Inventory. Oncologist. 2011; 16:217–227. [PubMed: 21285393]
14. Coons SJ, Kothari S, Monz BU, Burke LB. The Patient-reported Outcome (PRO) Consortium: filling measurement gaps for PRO end points to support labeling claims. Clin Pharmacol Ther. 2011 Nov.90:743–748. [PubMed: 21993428]
15. McCarrier KP, Deal LS, Abraham L, et al. Patient-centered research to support the development of the symptoms of major depressive disorder scale (SMDDS): Initial qualitative research. The Patient. 2015 Jun 26.
16. Patrick DL, Burke LB, Gwaltney CJ, et al. Content validity-Establishing and reporting the evidence in newly developed patient-reported outcomes (PRO) instruments for medical product evaluation: ISPOR PRO Good Research Practices Task Force Report: Part 1-Eliciting concepts for a new PRO instrument. Value Health. 2011 Dec.14:967–977. [PubMed: 22152165]
17. Patrick DL, Burke LB, Gwaltney CJ, et al. Content validity-Establishing and reporting the evidence in newly developed patient-reported outcomes (PRO) instruments for medical product evaluation: ISPOR PRO Good Research Practices Task Force Report: Part 2-Assessing respondent understanding. Value Health. 2011 Dec.14:978–988. [PubMed: 22152166]
18. National Cancer Institute, National Institutes of Health, U.S. Department of Health and Human Services. [Accessed January 28, 2016] Common Terminology Criteria for Adverse Events (CTCAE) Version 4.0. Published. 2009 May 28. Revised Version 4.03 June 14, 2010. Available from: http://evs.nci.nih.gov/ftp1/CTCAE/CTCAE_4.03_2010-06-14_QuickReference_8.5x11.pdf
19. Muir, T. User's Manual for ATLAS.ti 5.0. Berlin: ATLAS.ti Scientific Software Development GmbH; 2004. Available from: http://www.qsrinternational.com/products_nvivo.aspx
20. Brod M, Tesler LE, Christensen TL. Qualitative research and content validity: developing best practices based on science and experience. Qual Life Res. 2009 Nov.18:1263–1278. [PubMed: 19784865]

21. Basse SJ, Martin ML, McCarrier KP. The notion of representative languages in the context of translatability assessment. Value Health. 2013; 16:A2.

22. Bennett AV, Patrick DL, Lymp JF, Edwards TC, Goss CH. Comparison of 7-day and repeated 24-hour recall of symptoms of cystic fibrosis. J Cysc Fibr. 2010 Dec.9:419–424.

23. Bennett AV, Amtmann D, Diehr P, Patrick DL. Comparison of 7-day recall and daily diary reports of COPD symptoms and impacts. Value Health. 2012 May.15:466–474. [PubMed: 22583457]

24. Wood WA, Deal AM, Bennett AV, et al. Comparison of seven-day and repeated 24-hour recall of symptoms in the first 100 days after hematopoietic cell transplantation. J Pain Symptom Manage. 2015 Mar.49:513–520. [PubMed: 25128608]

25. Atkinson TM, Martin ML, McCarrier KP, et al. Use of concept elicitation interviews to determine potential differences in disease-related symptom concepts between early- versus advanced-stage non-small cell lung cancer (NSCLC) patients. Qual Life Res. 2015 Oct.24(Suppl 1):51.

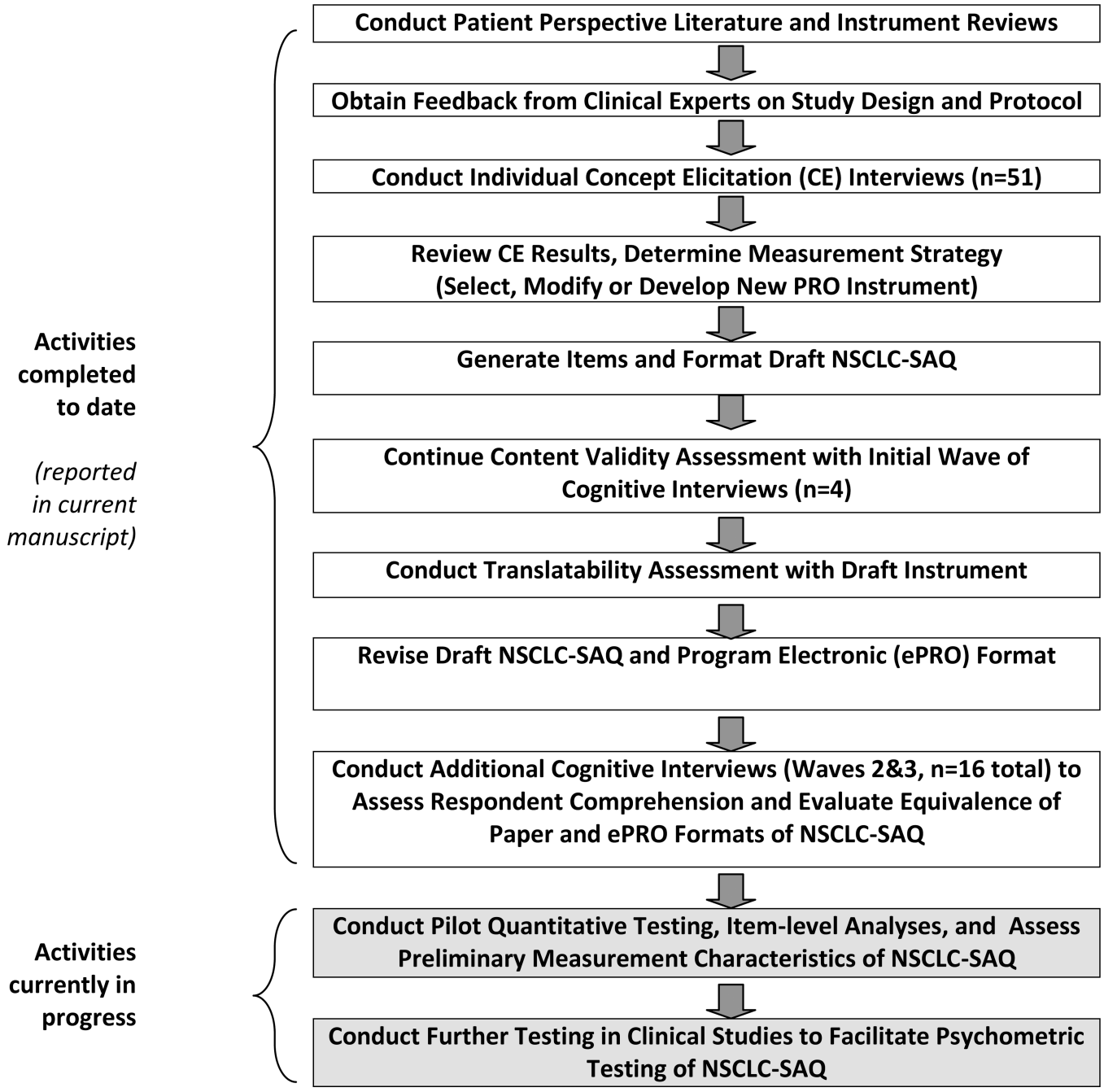

Figure 1.
Chronology of NSCLC-SAQ Development Activities

Symptom Intensity Item	Symptom Frequency Item
2. How would you rate the worst pain in your chest over the last 7 days? ☐ No Pain at All ☐ Mild Pain ☐ Moderate Pain ☐ Severe Pain ☐ Very Severe Pain	4. How often did you feel short of breath during usual activities over the last 7 days? ☐ Never ☐ Rarely ☐ Sometimes ☐ Often ☐ Always

Figure 2.

Example Items from Developmental Version of the NSCLC-SAQ

Source: Example items are from the Non-Small Cell Lung Cancer Symptom Assessment Questionnaire (NSCLC-SAQ©) and are used with permission of the Critical Path Institute.

Table 1

Characteristics of Study Participants

		Concept Elicitation Interviews N=51	Wave 1: Cognitive Interviews N=4	Waves 2 and 3: Cognitive Interviews + ePRO Assessment N=16
Age in Years:	- Mean (standard deviation)	64.9 (11.2)	68.0 (13.3)	64.8 (10.8)
	- Median	66	68	64.5
	- Range	46–86	56–80	44–83
Gender:	- Male	25 (49.0%)	1 (25.0%)	11 (68.8%)
	- Female	26 (51.0%)	3 (75.0%)	5 (31.3%)
Marital status:	- Married or Living with Partner	34 (66.7%)	1 (25.0%)	11 (68.8%)
	- Widowed	5 (9.8%)	2 (50.0%)	2 (12.5%)
	- Separated	1 (2.0%)	---	1 (6.3%)
	- Divorced	7 (13.7%)	1 (25.0%)	1 (6.3%)
	- Never Married	4 (7.8%)	---	1 (6.3%)
Highest Level of Education Completed:	- Less than High School	3 (5.9%)	1 (25.0%)	---
	- High School	25 (49.0%)	3 (75.0%)	10 (62.5%)
	- Some College	13 (25.5%)	---	6 (37.5%)
	- Bachelor's Degree	3 (5.9%)	---	---
	- Graduate or Professional School	7 (13.7%)	---	---
Employment outside home:	- Not Employed Outside Home	4 (7.8%)	---	2 (12.5%)
	- Employed for Wages (Full or Part time)	9 (17.6%)	1 (25.0%)	3 (18.8%)
	- Self-Employed	5 (9.8%)	---	---
	- Retired	21 (41.2%)	2 (50.0%)	8 (50.0%)
	- Unable to work	12 (23.5%)	1 (25.0%)	2 (12.5%)
	- Missing / Declined to Answer	---	---	1 (6.3%)
Ethnicity:	- Hispanic, Latino, or Spanish Origin	5 (9.8%)	---	1 (6.3%)
	- Not Hispanic or Latino	46 (90.2%)	4 (100%)	15 (93.8%)
Race:	- Asian	2 (3.9%)	---	---
	- Black/African American	8 (15.7%)	2 (50.0%)	3 (18.8%)
	- White	38 (74.5%)	2 (50.0%)	12 (75.0%)
	- Other	3 (5.9%)		1 (6.3%)
Household income:	- Under $4,999	2 (3.9%)	1 (25.0%)	---
	- $5,000 – $9,999	2 (3.9%)	1 (25.0%)	1 (6.3%)
	- $10,000 – $14,999	3 (5.9%)	1 (25.0%)	1 (6.3%)
	- $15,000 – $24,999	10 (19.6%)	---	1 (6.3%)
	- $25,000 – $34,999	8 (15.7%)	---	3 (18.8%)
	- $35,000 – $49,999	8 (15.7%)	1 (25.0%)	4 (25.0%)
	- $50,000 – $74,999	6 (11.8%)	---	4 (25.0%)

		Concept Elicitation Interviews N=51	Wave 1: Cognitive Interviews N=4	Waves 2 and 3: Cognitive Interviews + ePRO Assessment N=16
	- $75,000 and Over	11 (21.6%)	---	2 (12.5%)
	- Missing / Declined to Answer	1 (2.0%)	---	---
Stage at initial NSCLC diagnosis	- Stage I	6 (11.8%)	---	---
	- Stage II	1 (2.0%)	---	---
	- Stage III	25 (49.0%)	2 (50.0%)	8 (50.0%)
	- Stage IV	19 (37.3%)	2 (50.0%)	8 (50.0%)
Histological Subtype of NSCLC	- No histological evidence in record	1 (2.0%)	---	1 (6.3%)
	- Adenocarcinoma	36 (70.6%)	1 (25.0%)	5 (31.3%)
	- Squamous cell carcinoma	13 (25.5%)	3 (75.0%)	9 (56.3%)
	- Adenocarcinoma & Squamous cell carcinoma	1 (2.0%)	---	1 (6.3%)
Current stage of NSCLC	- Stage I	6 (11.8%)	---	1 (6.3%)
	- Stage II	---	---	---
	- Stage III	19 (37.3%)	2 (50.0%)	7 (43.8%)
	- Stage IV	26 (51.0%)	2 (50.0%)	8 (50.0%)
	- Early-stage (treatment naïve)	19 (37.3%)	1 (25.0%)	3 (18.8%)
Current line of NSCLC treatment	- 1st-line advanced/metastatic	18 (35.3%)	2 (50.0%)	5 (31.3%)
	- 2nd- line advanced/metastatic	9 (17.6%)	1 (25.0%)	3 (18.8%)
	- 3rd-line advanced/metastatic	3 (5.9%)	---	3 (18.8%)
	- Other (e.g., observation, subsequent)	2 (3.9%)	---	2 (12.5%)
Current ECOG performance status	- ECOG=0	17 (33.3%)	2 (50.0%)	5 (31.3%)
	- ECOG=1	24 (47.1%)	---	11 (68.8%)
	- ECOG=2	10 (19.6%)	2 (50.0%)	---
	- ECOG=3	---	---	---
	- ECOG=4	---	---	---
Patient smoking history	- Never a regular smoker	8 (15.7%)	---	3 (18.8%)
	- Current smoker	7 (13.7%)	1 (25.0%)	2 (12.5%)
	- Former smoker	36 (70.6%)	3 (75.0%)	10 (62.5%)
	- Missing/Unknown	---	---	1 (6.3%)
Number of Pack Years Smoked? (Current / Former Smokers only)	- Mean (standard deviation)	32.5 (22.0)	130.0 (156.8)	35.5 (13.1)
	- Median	35.0	65.0	30.0
	- Range	0–90	30–360	20–56
Indicator of Comorbid COPD in Medical Record	- Yes	18 (35.3%)	3 (75.0%)	2 (12.5%)
	- No	38 (64.7%)	1 (25.0%)	14 (87.5%)

Clin Ther. Author manuscript; available in PMC 2017 April 01.

Table 2

Summary of Symptom Concepts Identified in Concept Elicitation Interviews

NSCLC SYMPTOM DOMAIN[1]	SYMPTOM CONCEPT[2]	Predominance in coding		Patient assessment during interview		
		NUMBER OF CODED EXPRESSIONS OF CONCEPT[3] (Starred [*] if above mean number [44] of expressions per concept)	NUMBER OF PATIENTS EXPRESSING CONCEPT[4] (Starred [*] ≥15 [approx. 30% of sample])	PERCENT OF PATIENTS EXPRESSING CONCEPT SPONTANEOUSLY[5] (with probed percent if larger; starred [*] if above mean [≥12%] spontaneous)	MEAN SEVERITY RATING[6] (Starred [*] if above mean [>6.6])	MEAN BOTHERSOME RATING[7] (Starred [*] if above mean [>6.1])
Fatigue-Related Concepts	Exhaustion	16	7	4%	6.3	8.0*
	Fatigue	56*	20*	14%*	6.9*	6.1
	Low / Lack of Energy	41	15*	8%	7.5*	7.8*
	Low Stamina	10	7	2%	8.0*	8.0*
	Tire Easily / Tiredness	172*	39*	24%*	6.9*	6.8*
	Weakness	22	11	4%	4.0	5.3
Pain and Discomfort	***Back Pain***	47*	12	10%	7.7*	6.9*
	Bone Pain	9	6	2%	7.5*	6.0
	Pain in Chest	67*	17*	27%*	7.1*	6.8*
	General Pain	59*	18*	16%*	8.0*	6.6*
	Muscle Pain	39	13	10%	5.4	4.5
Respiratory Symptoms	Bronchitis	18	9	25%*	8.0*	3.0
	Coughing Up Blood	56*	12	12%*	4.5	5.7
	Cough	206*	41*	57%*	6.5	5.6
	Difficulty Breathing	91*	21*	18%*	6.1	6.0
	Emphysema	9	7	0%	NR	NR

NSCLC SYMPTOM DOMAIN[1]	SYMPTOM CONCEPT[2]	Predominance in coding		Patient assessment during interview		
		NUMBER OF CODED EXPRESSIONS OF CONCEPT[3] (Starred [*] if above mean number [44] of expressions per concept)	NUMBER OF PATIENTS EXPRESSING CONCEPT[4] (Starred [*] ≥15 [approx. 30% of sample])	PERCENT OF PATIENTS EXPRESSING CONCEPT SPONTANEOUSLY[5] (with probed percent if larger; starred [*] if above mean [≥12%] spontaneous)	MEAN SEVERITY RATING[6] (Starred [*] if above mean [>6.6])	MEAN BOTHERSOME RATING[7] (Starred [*] if above mean [>6.1])
	Phlegm	46*	16*	10%	5.1	4.5
	Pneumonia	62*	16*	16%*	8.2*	4.6
	Shortness of Breath	149*	35*	43%*	6.7*	6.9*
	Wheezing	35	12	10%	6.2	6.4*
Digestive Symptoms	***Poor Appetite***	92*	26*	22%	7.5*	6.9*
	Diarrhea	12	5	0%	NR	10.0*
	Difficulty Swallowing	16	7	2%	6.0	4.0
	Nausea	51*	15*	6%	5.3	7.2*
	Vomiting	19	6	4%	7.5*	6.0
Other Symptoms	Cognition	41	12	2%	8.0*	5.0
	Dizziness and Fainting	35	11	8%	8.8*	8.4*
	Headache	17	6	10%	6.0	6.0
	Heart Problems	12	3	2%	5.0	10.0*
	Heat Sensitivity	33	8	6%	7.7*	6.0
	Hoarseness	49*	20*	4% (20%)	4.9	5.5
	Immunity Lessened	37	12	4%	6.0	0.5
	Lump	19	5	6%	4.7	4.7
	Numbness	25	8	6%	7.3	8.7*
	Restlessnes s	5	2	0%	NR	NR

NSCLC SYMPTOM DOMAIN[1]	SYMPTOM CONCEPT[2]	Predominance in coding		Patient assessment during interview		
		NUMBER OF CODED EXPRESSIONS OF CONCEPT[3] (Starred [*] if above mean number [44] of expressions per concept)	NUMBER OF PATIENTS EXPRESSING CONCEPT[4] (Starred [*] ≥15 [approx. 30% of sample])	PERCENT OF PATIENTS EXPRESSING CONCEPT SPONTANEOUSLY[5] (with probed percent if larger; starred [*] if above mean [≥12%] spontaneous)	MEAN SEVERITY RATING[6] (Starred [*] if above mean [>6.6])	MEAN BOTHERSOME RATING[7] (Starred [*] if above mean [>6.1])
	Feeling Sick	6	7	4%	NR	NR
	Skin Change	14	3	0% (4%)	5.0	6.5*
	Sore Throat	7	3	2%	2.0	1.5
	Swelling	31	12	4%	8.0*	5.0
	Taste Change	15	5	2%	NR	8.5*
	Twitching	14	3	0%	NR	NR
	Voice Change	32	8	2% (14%)	5.9	5.2
	Weight Loss	53*	22*	18%*	6.6	6.2*

[1] NSCLC = Non-Small Cell Lung Cancer

[2] ***Bolded/Italicized*** entries in this column represent the NSCLC symptom concepts most highly supported by patient interview data (based on receipt of starred "high" designations in ≥3 of the following 5 categories: number of patients expressing, number of coded expressions, percent of spontaneous mentions, mean severity rating, and mean bothersome rating).

[3] In total, 1,897 expressions were coded for the 43 symptom concepts. Each concept had between 2 and 206 coded expressions, with a mean of 44 mentions per concept.

[4] Among n=51 total concept elicitation interview participants.

[5] Expressed symptom concepts were identified as "spontaneous" if described by the patient during the open-ended portion of the interview discussion, as opposed to those concepts arising in the discussion after specific symptom probing by the interviewer. Concepts in this column are starred if expressed spontaneously by >20% of interview participants.

[6] For each NSCLC symptom they experienced, interview participants rated the severity of the symptom "at its worst" on a 0 to 10 scale anchored with "none" and "extremely severe." Across all concepts, ratings ranged from 0 to 10, with a mean of 6.6.

[7] For each symptom they experienced, participants rated "how much that particular symptom bothers you" on a 0 to 10 scale anchored with "not bothersome at all" and "extremely bothersome." Across all concepts, the bothersome ratings given ranged from 0 to 10, with a mean of 6.1.

NRNo Rating: No patients provided rating data during interview.

Table 3

NSCLC Symptom Concepts Included in the NSCLC-SAQ (Through the Cognitive Interview Phase[1])

NSCLC SYMPTOM DOMAIN[1]	SYMPTOM CONCEPT	EXAMPLE PATIENT LANGUAGE SUPPORTING CONCEPT EXPRESSED DURING CONCEPT ELICITATION INTERVIEWS
Pain and Discomfort	Pain in Chest	- I could feel pain in my chest - had a lot of chest pain - it felt like someone was stabbing a knife into your chest - pain in my chest, my right lung, its real sensitive and sore around that area - I have pain in the right side of my chest radiating up to my right shoulder - pain in my chest
	General Pain (non-chest)	- I would say just general body pain - I hurt all over, it was so bad - it was just pain, just pain, engulfing my whole body - pain from the top of my head to the bottom of my feet - feeling like the time I had bad pain, sometimes I cry - get little short pains on numerous places in my body - I just cry because the pain is hard
Cough	Cough	- It's a constant cough there all the time - its better but I still have the cough - little hacking kind of cough like you have to clear your throat - only the coughing - still had that constant cough - if I start coughing and carrying on that's what it's tough - if I'm coughing I gotta wait, clear that - it was a low level cough, but constant - it's a dry air cough, a deeper cough - persistent back of the throat type of a dry cough - the cough would be so bad - the really bad cough
Hemoptysis	Coughing Up Blood[2]	- coughing and blood - the coughing and the blood - I'd cough up a little bit of blood you know - I'm still spitting up blood - spit up blood streaked sputum - when I coughed up that blood, that got my attention - I was coughing, I was coughing up blood
Dyspnea	Shortness of Breath	- I had a little shortness of breath - I can't hardly do anything without getting short winded - I feel a little short of breath

NSCLC SYMPTOM DOMAIN[1]	SYMPTOM CONCEPT	EXAMPLE PATIENT LANGUAGE SUPPORTING CONCEPT EXPRESSED DURING CONCEPT ELICITATION INTERVIEWS
		- shortness of breath, definite shortness of breath - the shortness of breath - the shortness of breath hasn't gone away - had some shortness of breath with walking - I got the shortness of breath - my breath got kinda short
	Difficulty Breathing[2]	- difficulty to breathe, very tiring - I noticed I have difficulty of breathing when I've walked a ways and I'm tired - *could go through a whole flea market breathing and walk, I can't do that* (*now*) - it's difficult for me to breathe - once they see me not breathing well, whatever we're (his friends and he) doing slows down basically - walk with my daughter, I can't keep up, I've got to stop and catch breath
Fatigue	Low/Lack of Energy	- lack of energy - I don't have the energy, that's a minus - just do not have the energy to do your daily tasks - just not having energy to want to do anything - I don't have the energy - I just have no energy - *my energy level, it's not 100%* - my energy was so low
	Tire Easily/Low Stamina	- I might get tired just coming in from the car - I noticed that I was tired - I would be so tired - I'm tired, very tired after I walk a ways, I'm tired - certain things I start doing and feel like I get a little tired - feel tired when I was talking on the phone - get tireder a little bit sooner that I used getting in the way of enjoying activities because you are so tired, you don't want to do anything - I get so tired, I have to just go sit down - I get tired a little quicker - I get tired real quick - just get tired real easy
Digestive Symptoms	Poor Appetite	- can look at food on the plate, it doesn't interest you at all - don't have an appetite - get up in the morning, you don't want anything to eat - I didn't really want to anything to eat - I had lost my appetite - it's hard to get an appetite

NSCLC SYMPTOM DOMAIN[1]	SYMPTOM CONCEPT	EXAMPLE PATIENT LANGUAGE SUPPORTING CONCEPT EXPRESSED DURING CONCEPT ELICITATION INTERVIEWS
		- my appetite was just gone - sometimes me appetite isn't very good - there was no appetite - weak appetite - lack of appetite - don't have hunger for anything, don't have any urge to eat - food repulsed me, I knew I had to eat, but I just couldn't - I have way less appetite - I just have no appetite, I have no interest in any food - I really have no appetite, have to make myself eat if I'm going to

[1] NSCLC = Non-Small Cell Lung Cancer

[2] Although included in the initial draft of the NSCLQ-SAQ, items for "coughing up blood" and "difficulty breathing" were removed based on patient responses during the cognitive interviews, and do not appear in the current 7-item version of the NSCLC-SAQ.