Consequences of Screening in Lung Cancer: Development and Dimensionality of a Questionnaire

John Brodersen, MD, GP, PhD, Hanne Thorsen, MD, PhD, Svend Kreiner, PhD

Institute of Public Health, Department and Research Unit of General Practice, University of Copenhagen, Copenhagen, Denmark

ABSTRACT

Objective: The objective of this study was to extend the Consequences of Screening (COS) Questionnaire for use in a lung cancer screening by testing for comprehension, content coverage, dimensionality, and reliability.
Methods: In interviews, the suitability, content coverage, and relevance of the COS were tested on participants in a lung cancer screening program. The results were thematically analyzed to identify the key consequences of abnormal and false-positive screening results. Item Response Theory and Classical Test Theory were used to analyze data. Dimensionality, objectivity, and reliability were established by item analysis, examining the fit between item responses and Rasch models.
Results: Eight themes specifically relevant for participants in lung cancer screening results were identified: "self-blame," "focus on symptoms," "stigmatization," "introvert," "harm of smoking," "impulsivity," "empathy," and "regretful of still smoking." Altogether, 26 new items for

part I and 16 new items for part II were generated. These themes were confirmed to fit a partial-credit Rasch model measuring different constructs including several of the new items.
Conclusion: In conclusion, the reliability and the dimensionality of a condition-specific measure with high content validity for persons having abnormal or false-positive lung cancer screening results have been demonstrated. This new questionnaire called Consequences of Screening in Lung Cancer (COS-LC) covers in two parts the psychosocial experience in lung cancer screening. Part I: "anxiety," "behavior," "dejection," "sleep," "self-blame," "focus on airway symptoms," "stigmatization," "introvert," and "harm of smoking." Part II: "calm/relax," "social network," "existential values," "impulsivity," "empathy," and "regretful of still smoking."
Keywords: lung cancer, psychometrics, public health, questionnaire development, secondary prevention.

Introduction

Participation in cancer screening programs has an impact on people's lives regardless of the results of the screening. For example, women have reported negative emotional consequences as a result of receiving an invitation to breast cancer screening [1–4]. By definition, screening tests are imperfect: screening can be unresponsive to some cancers (false negatives or the so-called false assurance) and, conversely, not infrequently raises the suspicion of cancer among healthy persons (false positives or the so-called false alarms). The impact of screening on people's lives is more severe for those receiving abnormal (including false positive) screening results than for those receiving normal results [5–8].

Psychosocial consequences have been widely explored in many different aspects of breast cancer and cervical cancer screening [5], but little is known about the long-term psychosocial consequences of cancer screening [9]. In breast cancer screening, the results from different studies on the long-term psychosocial consequences of false-positive results are inconsistent [6,10]. This is most likely caused by the inadequacy of the instruments used to measure the psychosocial aspects [10,11].

In general, generic measures have been used in most surveys to investigate potentially harmful psychosocial side effects of cancer screening programs [5,6]. Nevertheless, these generic measures lack content validity in terms of cancer screening [10,11]. Furthermore, various studies have shown that the psychometric properties of generic instruments are inconsistent across different populations [12–14]. The use of generic measures becomes even more problematic if a group of persons that fare least well are also those who find the generic instruments of least relevance [15]. It has been shown that specific instruments are more sensitive than

generic instruments [16]. Therefore, when measuring psychosocial consequences of screening, questionnaires with high content validity and adequate psychometric properties are needed [17].

Nonrandomized trials on lung cancer screening using low-dose spiral computed tomography (CT) scanning as the screening test have given indications of substantial harm and no benefits. In a statistical prediction model, screening with spiral CT scans resulted in an overdiagnosis of lung cancer: for every patient with a diagnosis of lung cancer in the screening program, more than two persons were captured with inconsequential lung cancers [18]. Screening with spiral CT scans has also revealed a high number of lung lesions among healthy persons, resulting in many false-positive screening results [19–21]. None of these nonrandomized trials have established sufficient evidence for any benefits of screening for lung cancer with spiral CT scanning [22]. Therefore, five randomized control trials have been launched on lung cancer screening, using low-dose spiral CT scanning as the screening test [23–27].

Three studies have investigated potentially harmful psychosocial aspects of lung cancer screening. A survey from the Pittsburgh Lung Screening Study concludes that persons with abnormal screening results exhibit negative psychological effects as a result of being screened—increased anxiety, increased perceived risk of cancer, and fear of cancer after screening [28]. A study from the Belgian-Dutch randomized controlled lung cancer screening trial concludes that there are no negative Health Related Quality of Life (HRQoL) effects of participation in the screening program [29]. And finally, a study from the Danish Lung Cancer Screening Trial revealed that participants having a false-positive spiral CT scan result change their habits by smoking less, which was interpreted as an indication of the emotional impact of false-positive screening results [30]. A substantial limitation of the two first studies is the use of generic instruments to capture potentially adverse effects of lung cancer screening. In the study from Pittsburgh, the State-Trait Anxiety Inventory (STAI) was used to measure anxiety [31]. The STAI has

previously been shown to be less sensitive than a breast cancer screening specific instrument, the Psychological Consequences Questionnaire (PCQ) [32,33]. The STAI has also demonstrated inadequacy in relation to language and content validity in cancer screening [10], and insufficient psychometric properties [34]. The Belgian-Dutch study used the 12-item Short Form (SF-12), the EuroQol questionnaire (EQ-5D), the STAI-6, and the Impact of Event Scale, which are all generic health-related quality-of-life instruments or utility measures [29]. No test on content relevance and content coverage of these instruments was reported in the article, neither were the psychometric properties of the generic instruments [29]. Nevertheless, the studies from Pittsburgh [28] and Denmark [30] confirm lack of content validity for the generic questionnaires used in a context of lung cancer screening: items adapted from the PCQ about the perceived risk of cancer and fear of cancer, and items about smoking habits all capture significant differences between persons with normal and persons with abnormal screening results. Because generic measures do not adequately capture the psychosocial aspects of cancer screening, a valid comprehensive measure needs to be developed to capture the potential psychosocial harm of lung cancer screening.

A condition-specific questionnaire, Consequence of Screening in Breast Cancer (COS-BC), which measures the psychosocial consequences of abnormal and false-positive screening mammography, has been developed on the basis of a Danish version of the PCQ [35]. In a qualitative study using focus-group interviews with women who had previously had a false-positive screening mammography, the PCQ was considerably changed and extended, with new items forming the COS-BC [35]. The COS-BC was subsequently statistically validated in two quantitative studies using the Item Response Theory and the Classical Test Theory [7,8]. The COS-BC consists of two parts: part I is especially relevant for the critical period from abnormal screening until final diagnosis, and part II is only relevant after the person has been declared free from suspicion of cancer.

The Danish Lung Cancer Screening Trial was launched in October 2005 and has been designed to investigate potential benefits and harms of the screening program, including psychosocial consequences [26].

It has previously been shown that many Danes perceive all cancers as one and the same disease with one and the same bad prognosis [36]. Therefore, it could be anticipated that the core items of the COS-BC (i.e., the COS-BC without the breast-specific items hereafter referred to as COS) would also be relevant for participants in lung cancer screening. Therefore, the purpose of the present study was threefold:

1. to examine the content relevance and content coverage of the core items of the COS in a setting of lung cancer screening;
2. if lack of content coverage of the COS was revealed, to generate themes and new items especially relevant for participants in lung cancer screening and to test the items for suitability;
3. if new items were generated, to test the extended version of the COS for dimensionality using the Item Response Theory Rasch model.

Materials and Methods

Data Collection: Content Relevance and Content Coverage of the COS for Application in Lung Cancer Screening

Interviewees were recruited in the prevalence round of the Danish Lung Cancer Screening Trial [26] to test the relevance and content coverage of the COS in a setting of abnormal and false-positive lung cancer screening.

Participants who, during the first 6 months of the prevalence round, had had an abnormal screening indicating a noncalcified nodule between 5 mm and 15 mm, and who were recalled for another scan after 3 months were invited to participate in five group interviews. The first two authors of this article participated as moderators in the group interviews.

The group interviews lasted approximately 2 hours and consisted of two parts. The first part was an open-ended discussion on the psychosocial consequences of abnormal and false-positive lung cancer screening results. The conceptualization of "psychosocial consequences" was based on the biopsychosocial model in which people are not regarded as passive [37]. They are considered as able to both interact with, and influence, the environment [38,39].

In the second part of the first group interview, the participants were asked to complete the COS and to comment on the relevance of the items. Part I of the COS encompasses three single items and four domains (including 23 items), in total 26 items with each four response categories: "not at all," "a bit," "quite a bit," and "a lot." Part II encompasses four domains, including 13 items with each five response categories: "much less," "less," "the same as before," "more," and "much more" [35]. The COS items are ordered thematically in Table 1. The participants of the first group interview only completed part I of the COS because, at the time of the interview, they did not know that their abnormal screening results were false positive. In the four subsequent group interviews, the participants completed versions of both parts of the questionnaire that included new items generated after the preceding group interview. In the group interviews, cognitive interviewing was also carried out item by item and covered understandability, content relevance, and content coverage [40,41].

The response options were also reviewed for relevance and ease of completion. In part I, the response options are ordered as categorical variables on a continuum going from "Not at all" to "A lot" (Fig. 1).

The response options in part II are also categorical variables on a continuum, but as can be seen in Figure 2, they are not ordered on one continuum but two. In six group interviews with women who had false-positive results from screening mammography, it was revealed that the women's experiences in the critical period from abnormal screening mammography until final false-positive diagnosis were entirely different from their experiences after the final diagnosis. In addition, it was argued that some issues could only be raised after the women had been declared "free from" suspicion of cancer. The participants classified the issues exclusively relevant to the period after the final diagnosis as long-term psychosocial consequences of false-positive screening mammography [35]. The women also claimed that the consequences of the final diagnosis were not all positive. Compared with the time before participation in screening, the women argued that after the final diagnosis, there were negative as well as positive consequences [35].

The COS-BC part II was developed so that each item included response options indicating "no change" as an anchor relative to two other options of changes in opposing directions. It is well known that people's values and perceptions of life can change as a result of trauma and existential crisis. These changes can be interpreted by the individual as positive, negative, or a combination of both. This has been seen in a qualitative study of cancer patients' reactions one or more years after diagnosis [42]. Another study showed that women with a diagnosis of breast cancer experienced both positive and negative consequences of

Table 1 Content of the COS

Themes or single items	The items of the COS. The number indicates the order of appearance in the questionnaire
Part I	
Anxiety	2. Worried about my future
	3. Scared
	12. Upset
	13. Restless
	14. Nervous
	23. Terrified
Behavioral	4. Irritable
	5. Quieter than normal
	8. Hard to concentrate
	10. Change in appetite
	17. Withdrawn into myself
	20. Difficulty dealing with work or other commitments
	22. Difficulty doing things around the house
Sense of dejection	1. Worried
	9. Time passed slowly
	11. Sad
	15. Uneasy
	18. Unable to cope
	19. Depressed
Sleep	6. Slept badly
	16. Taken long time to fall asleep
	21. Woken up far too early in the morning
	24. Awake most of the night
Single items	7. Busy to take mind off things
	51. Less interest in sex
	52. Sick leave
Part II	
Lung cancer	3. Anxiety about lung cancer
	13. Not lung cancer
Relaxed/calm	4. Relaxed
	9. Calm
Social relations	6. Family
	7. Friends
	8. Other people
Existential values	1. Broader aspects of life
	2. Enjoyment of life
	5. Thought about future
	10. Well-being
	11. Awareness of life
	12. Value life

COS, Consequence of Screening.

their disease [43]. Their positive reactions could be seen as a kind of backward rationality, irrationality, or ambivalence. These well-known phenomena are, for example, described in the psychological theory of cognitive dissonance.

Antonovsky also reported similar complex reactions in interviews with former Jewish victims of the Second World War who had spent time in concentration camps. Antonovsky described a concept of a "sense of coherence" as a salutogenic model of the relationship between health and disease [44]. Surprisingly, Antonovsky found that not all the victims had a low sense of coherence as a consequence of their traumas. Some victims had a high sense of coherence despite their stay in concentration camps [44,45].

Antonovsky did not use the positive reactions or greater sense of coherence as justification for sending people to concentration camps. In comparison, Gram et al. reported an overall positive impact on life as a long-term consequence of false-positive screening mammography. Nevertheless, the authors emphasized that it would be unreasonable to put this impact on the positive side of the balance sheet of breast cancer screening, "since first the fear, then the relief, are induced by the same screening" [46]. Therefore, part II of the COS requires a particular item-scoring system. It is important not to score the changes so that they neutralize each other: therefore, an assessment of a mean score of the dimensions does not express the actual distribution of changes. The Rasch model presupposes that changes occur in one direction only. It follows that any change from "The same as before" is to be regarded as a long-term psychosocial consequence of a false-positive screening result. Therefore, the responses to part II should be recoded "laterally reversed": a response to "Much less" or "Much more" becomes a response to one variable of "much less/more change," a response to "Less" or "More" becomes a response to one variable of "less/more change," and finally, a response to "The same as before" becomes a response to a variable of "no change."

The test version of the questionnaire including the new items was subsequently field-tested [47,48]. The field tests were carried out as individual interviews with five of the participants from the group interviews. The interviewees were asked if the test version of the questionnaire was easy to complete and if the layout was easy to comprehend.

The group interviews were audio recorded, and after each interview, the recording was independently audited by the first two authors, who conducted thematic analyses to determine the key consequences of abnormal and false-positive lung cancer screening results. In the subsequent group interviews, the identified themes were discussed in detail. Furthermore, the participants' verbatim comments were used to define a construct, specifying a range of intensity from, for example, "little" to "severe" "self-blame." The items belonging to the construct were compared in pairs to avoid redundancy by ensuring that they qualitatively did not have the same intensity. Finally, the results of the analyses were compared and, if they did not correspond,

Figure 1 The layout of the items in part I of Consequence of Screening plus the new items in the new themes in part I.

	Not at all	A bit	Quite a bit	A lot
I have been worried.	☐	☐	☐	☐

Figure 2 The layout of the items in part II of Consequence of Screening plus the new items in the new themes: "impulsivity" and "empathy."

After the examinations……:

☐	☐	☐	☐	☐
Much less	Less	**The same as before**	More	Much more
Less……..			**More…….**	

the relevant sequences from the audio-recording were re-audited and discussed until consensus was reached.

Data Collection for Statistical Psychometric Analysis

Data were collected from the years 2006 and 2007 from the first, second, and third screening incidence rounds in the Danish Lung Cancer Screening Trial [26].

Part I

Persons with abnormal, false-positive, and true-positive screening results. All those with an abnormal screening result were eligible. These persons were asked to complete part I of the questionnaire 1 week after receiving the abnormal screening result. The same persons, later diagnosed either with a false-positive screening result or with lung cancer (true-positive result), were all asked to complete part I 1 month after the final diagnosis.

Persons with normal screening results. For each person with an abnormal screening result who was asked to complete the questionnaire, another two persons with normal screening results were asked to complete part I of the questionnaire 1 week after receiving their screening results. The two persons with normal screening results were selected based on the following criteria:

1. Persons receiving a normal screening result within 1 week of the person receiving the abnormal screening result.
2. Persons having the same sex as the person with the abnormal screening result.
3. Persons of the same age (+/−3 years) as the person with the abnormal screening result.

Persons with normal screening results were also asked to complete part I of the questionnaire 1 month after receiving their screening results.

Part II

Persons with false-positive and true-positive screening results. These persons were asked to complete part II of the questionnaire at 1, 6, and 18 months after the final diagnosis.

Persons with normal screening results. These participants were recruited, as described above, and were asked to complete part II of the questionnaire at 1, 6, and 18 months after receiving their screening results.

All persons were sent the questionnaire in the post and were asked to complete and return it in an enclosed stamped addressed envelope. A reminder was posted to those who had not returned the questionnaire within 2 weeks.

Statistical Analyses on Dimensionality

Item responses were analyzed by the conditional distribution of items given total person scores to avoid assumptions on the distribution of the latent trait being measured. The pair-wise estimation procedure implemented in the software program RUMM2020 (Version 4.0 for Windows, RUMM Laboratory Pty Ltd., Perth, WA, Australia) was used to estimate the item parameters [49,50]. The analysis of the fit of item responses to the partial-credit Rasch model was based on analyses of residuals comparing observed to expected item responses, both for separate individuals and for different score groups. The overall fit of the model was assessed by the Wright–Panchapakesan χ^2 statistic summarizing standardized residuals over score groups and items [51]. Item fit statistics summarizing standardized residuals in different score groups were used to identify items that did not fit the partial-credit Rasch model.

An analysis on differential item functioning (DIF) related to time of assessment was conducted to determine if the items functioned consistently over time and if there was any local response dependence between the points of assessment at 1 week and 1, 6, and 18 months after diagnosis/screening. If no DIF was identified in relation to the time of assessment, then it was found valid to pool the data from the different administrations for the analyses of dimensionality and DIF in relation to other covariates [52,53]. The DIF relative to person covariates was checked by analyses of variance examining the degree to which individual residuals for specific items depended on the covariates [54]. The absence of an evidence of interaction between the covariates and the estimated trait parameters was taken as an evidence of DIF being uniform [55]. If one or more items in a dimension possessed uniform DIF, then the items were split stepwise, beginning with the item possessing the greatest magnitude of uniform DIF [56]. Items possessing nonuniform DIF were deleted from the dimension.

Reliability was assessed by Cronbach's alpha defining the lower bound for the test–retest correlation of the raw scores [57,58] and by the Person Separation Index (PSI) calculating the lower bound for the test–retest correlation of the estimated values of the latent trait being measured [59].

When one Rasch analysis is conducted on two or more items, information about the potential psychometric properties of the scale and the items is revealed. In one Rasch analysis, it is possible to generate information about the scale's overall fit to the model expressed as a *P*-value, the total-item chi-square, and the degrees of freedom. Two estimates about the scale's reliability are computed: the Cronbach's alpha and the PSI. The information generated about the items includes the item's fit to the Rasch model, the individual item's chi-square, degrees of freedom, the item's ability to discriminate (including under-discrimination and over-discrimination), the functionality of the item's response categories (order of thresholds of the response categories), the item's location relative to the other items, and finally, DIF (uniform or non-uniform). Furthermore, information about how the items and the responders are targeted in relation to each other is revealed, including how much "information" the items are capturing from the responders. In addition, the "item information" captured by one item can be compared with the "information" that the remaining items are capturing separately.

Items that are "problematic" in relation to the revealed information described above, e.g., items that present a misfit to the partial-credit Rasch model (defined as $P < 0.05$ after a correction of the Benjamini–Hochberg procedure [60]), items possessing DIF, disordered thresholds, etc., are categorized as "poor" items. If a scale encompasses one or more "poor" items, then these items will affect the evidence concerning the other items on the scale. For example, if a "poor" item has a strong DIF in one direction, then evidence will suggest that the DIF will operate in the opposite direction for other items on the scale: the DIF will level out for the remaining items on the scale. Therefore, an item possessing real DIF can affect other items on the scale to show DIF—a DIF that is artificial. Should an item possessing real DIF be split, the fit to the Rasch model should increase—and vice versa if an item with artificial DIF was to be split.

The items included in each theme in the COS and the items included in each new lung cancer screening-specific theme were analyzed individually to test whether the items in a theme fitted the partial-credit Rasch model. Except for items possessing

Table 2 Participants in the group interviews

Sex	Women	Women	Men	Men	Women
Number	5	4	4	3	4
Mean age and range	57.2 (53–63)	58.4 (55–63)	62.5 (59–66)	62.5 (62–64)	59.2 (58–63)
Stage in relation to the screening result	Critical period	3 to 4 months after final result	4 to 5 months after final result	2 months after final result	2 months after final result

uniform DIF, "poor" items possessing the greatest magnitude of psychometric "problems" were deleted from the theme stepwise, and a Rasch analysis was conducted with the remaining items composing the theme.

The item on sick leave (no. 52, Table 1) and the two other single items (nos. 7 and 51, Table 1) were not included in the Rasch analyses because these items did not belong to any of the other dimensions.

The Cronbach's alpha was calculated using the software SPSS for Windows version 16.0 (SPSS, Chicago, IL). For Rasch analyses, the software RUMM2020 was used [50]. The study was approved by the Danish Data Protection Agency.

Results

Results from the Interviews

Altogether, nine women and seven men participated in five group interviews. Five women were interviewed before they knew that their abnormal screening results were false positive. Four of these women participated again in the last group interview, which was conducted 2 months after they had been told that their screening results were false positive (Table 2).

Three women and two men participated in the field test. During the field test, only minor linguistic changes were made in part I, e.g., "more than usual" was added to the item "I have been aware of my weight," a phrase that was already included in several other items. Another example was that the word "other(s)" had to be highlighted in items 32, 34, and 46 (see Table 3). No items were changed in part II.

The items in the COS were all found relevant by the participants. Five themes especially relevant for the critical period were extracted from the interviews: "self-blame," "focus on symptoms," "stigmatization," "introvert," and "harm of smoking." All five themes were generated in the first group interview. Two themes especially relevant for the period after known diagnosis were identified: "impulsivity" and "empathy." In addition, "regretful of still smoking" occupied the minds of those participants who were currently smokers. These three themes were generated in the second group interview. The themes and the generated items matching these themes are listed in Table 3.

Results of the Data Collection for the Statistical Psychometric Analysis

There was a sample of 195 persons eligible to complete part I of the questionnaire. Of those asked to participate, 183 (93.8%) agreed to complete the questionnaire 1 week after screening, and 173 (88.7%) agreed to complete the questionnaire 1 month after screening.

There was a sample of 195 persons eligible to complete part II of the questionnaire at 1 and 6 months after screening. Of these, 173 (88.7%) agreed to complete the questionnaire at 1 month after screening, and 164 (84.1%) agreed to complete the questionnaire at 6 months after screening. At the time of their inclusion in the present study (March 25, 2009), 111 of the 195 persons had known their screening result for 18 months. Of these, 91 (82.0%) completed part II of the questionnaire.

Table 3 The new themes and the new items generated in the group interviews

Themes	The content of the new items. The number indicates the order of appearance in the questionnaire
Part I	
Self-blame	35. Blamed oneself*
	36. Felt guilty*
	38. Disappointed in oneself*
	48. Angry with oneself*
	50. Been annoyed with oneself*
Focus on (airway) symptoms	26. Aware of one's weight
	27. Aware of being short of breath
	37. Considered going to the doctor
	40. More colds than usual
	42. More tired than usual
	45. Been aware of one's coughing
	47. Felt ill
	49. Been aware of coughing up mucous
Stigmatization	32. Been told off by other people*
	34. A finger wagging from others*
	43. Felt stigmatized*
	46. Blamed by other people*
Introvert	29. Fear of lung cancer at the back of one's mind
	31. Insecure
	33. Sorry for oneself
	39. Thought one's situation hopeless
	41. Mood swings
	44. Keeping things from those who are close to you
Harm of smoking	25. Thought of smoking as harmful
	30. Sorry*
Anxiety†	28. Shocked
Part II	
Impulsivity	14. energy
	16. Lived life to the full
	19. Being impulsive
	21. Desire to venture into something new
	22. Desire to venture into something risky
	23. Done some things that overstepped one's bounds
Empathy	15. Responsibility for one's family
	18. Understand other people's problems
	20. Ability to listen to other people's problems
Relaxed/calm‡	17. Relieved
24. Are you currently smoking? If yes, please complete the questions below:	
Regretful still smoking	25. Thought about quitting smoking
	26. Feel guilty for smoking
	27. Annoyed with oneself for smoking
	28. Disappointed in oneself for smoking
	29. Change in one's attitude toward smoking
	30. Having second thoughts about one's smoking

*For having smoked for many years.
†New item possibly belonging to the "old" anxiety scale in COS.
‡New item possibly belonging to the "old" relaxed/calm scale in COS.
COS, Consequence of Screening.

Altogether, 26 new items for part I were generated. The subject matter for these items was generated in the first group interview and was accepted in the following group interviews. In all, 16 new items for part II were generated, and the subject matter for these items was generated in the second group interview and was accepted in the three subsequent group interviews.

Table 4 Wright–Panchapakesan (WP) fit statistics, Person Separation Index, and the Cronbach's alpha of nine dimensions in part I of the Consequence of Screening in Lung Cancer and the "still regretful smoking" dimension in part II

Dimensions (Number of items)	WP chi-square	Degrees of freedom	P	Person Separation Index	Cronbach's alpha (time 1/2/3)
Anxiety (7)—before splitting items	57.87	34	0.007*	0.89	0.90 (0.87/92)
Anxiety (7)—split item 3 for sex and items 2, 14, and 28 for diagnosis†	66.21	55	0.143	0.89	
Behavioral (7)	47.82	35	0.073	0.87	0.88 (0.79/0.92)
Sense of dejection (6)—before splitting item	33.91	30	0.284	0.91	0.88 (0.91/0.86)
Sense of dejection (6)—split item 1 for diagnosis (abnormal) and item 15 for sex†	43.31	40	0.332	0.91	
Sleep (4)	47.47	16	0.000058*	0.85	0.86 (0.88/0.83)
Sleep (3)—minus item 24 "awake most of the night"	22.15	12	0.035832‡	0.80	0.75 (0.64/0.85)
Self-blame (5)	31.63	20	0.047372‡	0.97	0.91 (0.94/0.89)
Focus on airway symptoms (2)	10.08	7	0.184	0.71	0.67 (0.048/0.82)
Stigmatization (4)	30.65	20	0.060	0.89	0.78 (0.74/0.81)
Introvert (4)	25.72	16	0.058	0.82	0.86 (0.80/0.91)
Harm of smoking (2)	6.94	7	0.435	0.81	0.84 (0.78/0.92)
Still regretful smoking (6)	72.55	12	<0.0000005*	0.84	0.86 (0.86/0.85/0.86)
Still regretful smoking (5)—minus item 29 "change in one's attitude toward smoking"	26.06	10	0.004*	0.85	0.87 (0.87/0.87/0.86)
Still regretful smoking (4)—minus items 29 and 5 "feel guilty for smoking"	9.40	4	0.052	0.84	0.86 (0.86/0.87/0.83)

*Misfit after a correction of Benjamini–Hochberg procedure [60].

†The patients with lung cancer were excluded from these analyses.

‡Adjusting the *P*-values in the table to control the false discovery rate and to avoid spurious significant results due to multiple testing suggested that this result should be regarded as insignificant [60].

Results from the Rasch Analyses

No DIF was identified in relation to the time of assessment. Therefore, the data from the different administrations were pooled for the Rasch analyses on dimensionality and DIF in relation to covariates other than time of assessment.

Part I

Dimensionality of part I of the core-questionnaire COS. In part I, only the sleep dimension from the core-questionnaire COS did not include all original items. Item no. 24, "awake most of the night," (Table 1) did not fit the Rasch model, whereas the remaining three sleep items fitted the partial-credit Rasch model (Table 4). All the other items from the COS fitted their respective three dimensions, forming subscales of anxiety, sense of dejection, and negative impact on behavior (Table 4). One of the new items, "shocked," which was generated in the interviews in the present study, belonged to the anxiety scale.

Some items in the anxiety and dejection dimensions possessed DIF in relation to sex and diagnosis. Splitting those items increased the fit to the Rasch model (Table 4).

Dimensionality of part I of the lung cancer screening-specific items. All five items covering the theme "self-blame" fitted the partial-credit Rasch model (Table 4). Only two of the eight items, covering the theme "focus on symptoms," fitted the Rasch model. Because these two items, "aware of being short of breath" and "been aware of one's coughing," both described symptoms from the airway, the subscale was renamed to "focus on airway symptoms" (Table 4). The four "stigmatization" items, four out of the six "introvert" items, and the two "harm of smoking" items fitted the Rasch model (Table 4).

Of the 26 new items in part I, 25 items covered the five themes: "self-blame," "focus on symptoms," "stigmatization," "introvert," and "harm of smoking." Eight of these 25 items were identified to be "poor" or single items. The summary results of the item fit statistics, their location, and order in the dimensions of part I are listed in Table 5.

Part II

Dimensionality of part II of the core-questionnaire COS plus the new "impulsivity" and "empathy" themes. The response pattern in the items of part II of the COS and the items covering the two new themes "impulsivity" and "empathy" was skewed so that 0% to 12.1% responded to the two response categories "Much less" and "Less" (Fig. 2). As described in the method section, Rasch analyses were performed after transforming the data by recoding the item responses as "laterally reversed." However, because the data were skewed, it was also decided to investigate the fit of responses to Rasch models in two additional ways (Table 6).

In part II, the two items making up the "lung cancer" dimension did not fit a partial-credit Rasch model in two of the three data sets (Table 7). Despite the fit in the data set recoded "laterally reversed," the reliability was low: PSI = 0.38 and Cronbach's alpha = 0.47. Low values of the PSI and Cronbach's alpha could be caused by bad targeting between items and persons, however, this was not the case for the "lung cancer" dimension in any of the three recoded data sets. Therefore, these two items have to be regarded as single items. The new item "relieved" did not fit the original "relax/calm" dimension in any of the three data sets and, therefore, has to be regarded as a single or "poor" item (Table 7). All six dimensions fitted the partial-credit Rasch model in all three data sets except for the "existential" dimension in the data set recoded "the same as before" (Table 7).

The item "broader aspects of life" did not fit the "existential values" dimension. There was a marginal misfit for these three other items and their respective dimensions: "thought about future" and "awareness of life" both belong to the "existential values" dimension, and the item "energy" belong to the "impulsivity" dimension. To avoid spurious significant results because of multiple testing, the Benjamini–Hochberg procedure [60] was used to adjust the *P*-values to control the false discovery rate. After adjusting the *P*-values for the three above-mentioned items, these results should be regarded as insignificant (Table 8).

All the items' thresholds were in order in all the Rasch analyses on the three recoded data sets.

Table 5 Summary of result of item fit statistics and ordering in the dimensions of part I of the COS-LC

Dimension	The items of part I of the COS-LC in location order in the dimensions. The number indicates the order of appearance in the questionnaire	Chi-square	Degrees of freedom	Probability of fit to the Rasch model	Item location
Anxiety	2. Worried about my future (split for diagnosis: "abnormal" and "false positive")	1.266	5	0.938	−2.064
	14. Nervous (split for diagnosis: "abnormal")	6.007	5	0.306	−1.130
	3. Scared (split for sex: "female")	7.728	5	0.172	−0.345
	13. Restless	8.223	5	0.144	−0.195
	28. Shocked (split for diagnosis: "abnormal" and "false positive")	8.366	5	0.137	−0.030
	3. Scared (split for sex: "male")	8.873	5	0.114	0.166
	14. Nervous (split for diagnosis: "normal" & "false-positive")	14.447	5	0.013008*	0.189
	28. Shocked (split for diagnoses: "normal")	3.002	5	0.700	0.348
	12. Upset	2.008	5	0.848	0.615
	23. Terrified	3.268	5	0.659	0.691
	2. Worried about my future (split for diagnosis: "normal")	3.025	5	0.696	1.754
Behavioral	22. Difficulty doing things around the house	11.960	5	0.035344*	−0.239
	20. Difficulty dealing with work or other commitments	5.684	5	0.338	−0.146
	5. Quieter than normal	9.082	5	0.106	−0.145
	8. Hard to concentrate	4.189	5	0.523	−0.038
	17. Withdrawn into myself	3.595	5	0.609	0.132
	10. Change in appetite	5.836	5	0.322	0.184
	4. Irritable	7.475	5	0.188	0.252
Sense of dejection	1. Worried (split for diagnosis: "abnormal")	5.175	5	0.395	−1.79
	15. Uneasy (split for sex: "female")	7.482	5	0.187	−0.621
	1. Worried (split for diagnosis: "normal" and "false positive")	7.776	5	0.169	−0.583
	11. Sad	7.100	5	0.213	−0.21
	19. Depressed	2.841	5	0.725	−0.151
	9. Time passed slowly	5.430	5	0.366	0.386
	18. Unable to cope	3.888	5	0.566	0.886
	15. Uneasy (split for sex: "male")	3.619	5	0.605	2.083
Sleep	21. Woken up far too early in the morning	4.426	3	0.219	−0.191
	6. Slept badly	14.863	4	0.004996*	−0.046
	16. Taken long time to fall asleep	2.865	5	0.721	0.237
Self-blame	36. Felt guilty[†]	6.305	4	0.178	−0.557
	35. Blamed oneself[†]	11.104	4	0.025424*	−0.43
	50. Been annoyed with oneself[†]	3.525	4	0.474	0.169
	38. Disappointed in oneself[†]	2.846	4	0.584	0.228
	48. Angry with oneself[†]	7.853	4	0.097	0.589
Focus on airway symptoms	27. Aware of being short of breath	4.706	4	0.319	−0.073
	45. Been aware of one's coughing	5.371	3	0.147	0.073
Stigmatization	43. Felt stigmatized[†]	10.226	5	0.069	−0.195
	32. Been told off by other people[†]	3.272	5	0.658	−0.117
	34. A finger wagging from others[†]	10.328	5	0.066	0.105
	46. Blamed by other people[†]	6.819	5	0.234	0.207
Introvert	31. Insecure	7.509	4	0.111	−0.48
	41. Mood swings	4.678	4	0.322	0.008
	39. Thought one's situation hopeless	5.607	4	0.230	0.113
	33. Sorry for oneself	7.928	4	0.094	0.36
Harm of smoking	25. Thought of smoking as harmful	2.648	4	0.618	−0.261
	30. Sorry[†]	4.292	3	0.232	0.261

*Adjusting the *P*-values in the table to control the false discovery rate and to avoid spurious significant results due to multiple testing suggested that this result should be regarded as insignificant [60].
[†]For having smoked for many years.
COS-LC, Consequence of Screening in Lung Cancer.

Dimensionality of the Six New Items in Part II Covering the Theme "Regretful of Still Smoking"

Four of the six new items covering "regretful of still smoking" for those participants who were currently smokers fitted a partial-credit Rasch model (Table 4). The two items, "feel guilty for smoking" and "change in one's attitude toward smoking," must be regarded as "poor" or single items. The remaining four items in the dimension are listed in order of location in Table 9.

None of the items in part II possessed DIF in relation to sex, age, and diagnosis.

Discussion

All four dimensions from part I of the core-questionnaire COS—anxiety, sense of dejection, negative impact on behavior, and sleep—were found qualitatively relevant and psychometri-

Table 6 Three different ways to recode the item responses in part II to able to conduct Rasch analysis

	Much less	Less	The same as before	More	Much more
Data set recoded as "missing"	Missing	Missing	0	1	2
Data set recoded as "the same as before"	0	0	0	1	2
Data set recoded as "laterally reversed"	2	1	0	1	2

Table 7 The Wright–Panchapakesan (WP) chi-square fit statistics, the Person Separation Index, and the Cronbach's alpha of the six dimensions in Consequence of Screening in Lung Cancer part II

Dimensions (no. of items)	WP chi-square	Degrees of freedom	P	Person Separation Index	Cronbach's alpha (time 1/2/3)
Recoding—Data set recoded as "missing"					
Lung cancer (2)	18.30	5	0.002589*	0.57	0.56 (0.53/0.54/0.63)
Relaxed/calm (3)	33.35	12	0.000854*	0.69	0. 75 (0.70/0.78/0.79)
Relaxed/calm (2)	6.71	5	0.243	0.68	0.75 (0.71/0.82/0.67)
Social relations (3)	7.63	10	0.665	0.86	0. 85 (0.81/0.90/0.87)
Existential values (6)	44.94	30	0.039087†	0.81	0.84 (0.82/0.82/0.89)
Impulsivity (6)	27.64	24	0.276	0.84	0.83 (0.85/0.77/0.87)
Empathy (3)	14.13	10	0.167	0.66	0.69 (0.71/0.56/0.82)
Recoding—Data set recoded as "the same as before"					
Lung cancer (2)	19.73	5	0.001409*	0.60	0.56 (0.53/0.54/0.61)
Relaxed/calm (3)	37.35	12	0.000196*	0.70	0.73 (0.68/0.76/0.79)
Relaxed/calm (2)	7.19	4	0.126	0.69	0.75 (0.73/0.80/0.68)
Social relations (3)	7.89	9	0.546	0.87	0.84 (0.79/0.89/0.87)
Existential values (6)	77.34	30	0.000005*	0.80	0.82 (0.80/0.80/0.87)
Impulsivity (6)	36.00	24	0.055	0.82	0.85 (0.87/0.82/0.86)
Empathy (3)	10.61	6	0.101	0.66	0.68 (0.71/0.55/0.81)
Recoding—Data set recoded as "laterally reversed"					
Lung cancer (2)	9.46	5	0.092	0.38	0.47 (0.43/0.45/0.56)
Relaxed/calm (3)	25.14	12	0.014184*	0.68	0.73 (0.66/0.76/0.81)
Relaxed/calm (2)	4.72	4	0.317	0.61	0.70 (0.66/0.75/0.70)
Social relations (3)	12.30	9	0.197	0.86	0.84 (0.81/0.86/0.87)
Existential values (6)	38.22	30	0.144	0.83	0.85 (0.84/0.83/0.90)
Impulsivity (6)	23.83	24	0.471	0.85	0.88 (0.88/0.88/0.90)
Empathy (3)	9.89	6	0.129	0.66	0.69 (0.71/0.56/0.84)

*Misfit after a correction of Benjamini–Hochberg procedure [60].
†Adjusting the P-values in the table to control the false discovery rate and to avoid spurious significant results due to multiple testing suggested that this result should be regarded as insignificant [60].

cally valid for persons having abnormal and false-positive findings in screening for lung cancer. This was also valid for three of the four dimensions from part II—calm/relax, social network, and existential values. One new item was generated for the anxiety scale and, together with the other six anxiety items, fitted a Rasch model. The two items involving belief and anxiety about lung cancer did not measure the same construct and must be regarded as single items.

Five new dimensions for part I of the questionnaire specifically relevant for participants in lung cancer screening were developed and confirmed in a partial-credit Rasch model to measure different constructs—"self-blame," "focus on airway symptoms," "stigmatization," "introvert," and "harm of smoking." Three new dimensions for part II were also developed and fitted a Rasch model—"impulsivity," "empathy," and "regretful of still smoking."

There is a trade-off between developing a single scale that includes many items and developing several scales that have a smaller number of items. In the present study, the overall aim was to develop a psychosocial measure for those having abnormal and false-positive lung cancer screening results. The conceptualization of "psychosocial consequences" was based on the biopsychosocial model, whose name in itself indicates something multidimensional. Another aim of the study was to develop a measure with high content validity, including all relevant themes for participants with abnormal and false-positive lung cancer screening results. Linguistically, it is difficult to pose many different questions for all themes; we have, for example, identified four different qualities within the theme "sleeping problems." Therefore, developing a multidimensional measure in a health-care area in which members of the target population are regarded as the experts of the content of the measure may result in scales with few items. Scales with few items tend to be less reliable than scales with many items. Nevertheless, if many of the scales with few items measure the same general construct operating in the same direction, then the con-

struct validity might be higher for a multidimensional measure than for a unidimensional measure. When developing and validating the COS-BC, confirmatory factor analyses (CFA) revealed that the scales included in part I and the scales included in part II of the COS measured the same overall construct, respectively [7,8]. In the future, such CFA could also be conducted with data collected using the COS in Lung Cancer (COS-LC).

The field testing was carried out as face-to-face interviews with five participants from the group interviews. These interviewees were selected because they had been "good informants" in the group interviews. A limitation of the field test could be a degree of dependence in the interviewees' responses because they had completed parts of the COS-LC, or a version close to the final version of the COS-LC previously. Nevertheless, the purpose of the field test was only to test the "user-friendliness" and not the relevance and irrelevance of the items. Therefore, we do not think that this hypothetical response dependence has caused any problems.

All the new themes and the subject matter of all the new items were generated during the first two group interviews. The new themes and the new items were accepted in the subsequent group interviews. These resulted in robust evidence of high data saturation and, thus, evidence of high content validity of the COS-LC.

Eight new items in part I and one new item in part II did not fit the dimensions for which they were developed. These items could either be regarded as "poor" or single items. This depends on the items' responsiveness because they were all found relevant in the group interviews. When the psychosocial impact of lung cancer screening is to be estimated from the longitudinal data collected from all four incidence screening rounds in the Danish Lung Cancer Screening Trial [26], the new items' responsiveness could be compared with the responsiveness of the Rasch fitting dimensions. If any statistically significant differences are found between persons having normal and abnormal screening results

Table 8 Individual item fit statistics for items numbered 1 to 23 of part II of the COS-LC

The items of part II of the COS-LC in location order as in the data set recoded as "laterally reversed." The number indicates the order of appearance in the questionnaire.

Dimensions		WP chi-square	Degrees of freedom	P	Item location	WP chi-square	Degrees of freedom	P	Item location	WP chi-square	Degrees of freedom	P	Item location
		Data set recoded as missing				Data set recoded as "the same as before"				Data set recoded as laterally reversed			
Relaxed/calm	4. Relaxed	2.16	3	0.539	−0.666	2.26	2	0.323	−0.670	2.30	2	0.317	−0.325
	9. Calm	4.55	2	0.103	0.666	4.93	2	0.085	0.670	2.42	2	0.298	0.325
Social relations	6. Family	1.93	4	0.748	−2.367	1.93	3	0.586	−2.422	2.43	3	0.488	−2.198
	7. Friends	2.66	3	0.448	−0.665	3.27	3	0.352	−0.633	6.21	3	0.102	−0.645
	8. Other people	3.04	3	0.386	3.032	2.69	3	0.443	3.056	3.67	3	0.300	2.843
Existential values	1. Broader aspects of life	10.98	5	0.052	−0.351	23.64	5	0.000256*	−0.480	8.06	5	0.153	−0.340
	12. Value life	6.91	5	0.228	−0.162	11.96	5	0.035†	−0.247	6.43	5	0.267	−0.074
	2. Enjoyment of life	6.31	5	0.277	−0.230	10.73	5	0.057	−0.168	5.10	5	0.404	−0.061
	11. Awareness of life	10.59	5	0.060	−0.094	14.85	5	0.011026†	−0.173	7.21	5	0.205	0.000
	5. Thought about future	6.67	5	0.246	0.109	11.87	5	0.036552†	0.232	7.61	5	0.179	0.061
	10. Well-being	3.49	5	0.625	0.728	4.29	5	0.508	0.836	3.81	5	0.576	0.414
Impulsivity	14. Energy	11.88	4	0.018246†	−0.379	15.68	4	0.003486†	−0.287	3.71	4	0.446	−0.557
	16. Lived life to the full	2.48	4	0.649	−0.063	4.95	4	0.293	−0.143	4.44	4	0.350	0.023
	19. Being impulsive	2.13	4	0.712	−0.382	3.04	4	0.551	−0.317	3.09	4	0.543	0.078
	22. Desire to venture into something risky	4.31	4	0.366	0.925	3.34	4	0.502	0.835	6.18	4	0.186	0.09
	21. Desire to venture into something new	4.33	4	0.363	−0.049	6.23	4	0.183	−0.022	4.32	4	0.364	0.122
	23. Done some things that overstepped one's bounds	2.51	4	0.643	−0.051	2.75	4	0.600	−0.065	2.09	4	0.720	0.243
Empathy	18. Understand other people's problems	4.80	4	0.309	−0.426	2.34	2	0.313	−0.449	2.24	2	0.327	−0.442
	15. Responsibility for one's family	6.08	4	0.193	−0.093	5.07	2	0.079	−0.091	4.37	2	0.113	−0.095
	20. Ability to listen to other people's problems	3.26	2	0.196	0.52	3.21	2	0.201	0.54	3.29	2	0.193	0.537

*Misfit after a correction of Benjamini–Hochberg procedure [60].
†Adjusting the P-values in the table to control the false discovery rate and to avoid spurious significant results due to multiple testing suggested that this result should be regarded as insignificant [60].
COS-LC, Consequence of Screening in Lung Cancer; WP, Wright–Panchapakesan.

Table 9 Summary of result of item fit statistics and ordering in the "regretful of still smoking" dimension in part II of the COS-LC

The items composing the "regretful of still smoking" dimension in part II of the COS-LC. The items are listed in location order, and the number indicates the order of appearance in the questionnaire	Chi-square	Degrees of freedom	Probability of fit to the Rasch model	Item location
24. Are you currently smoking? If yes, please complete the questions below:	Single item	Single item	Single item	Single item
25. Thought about quitting smoking	2.51	1	0.113	−1.662
28. Disappointed in oneself for smoking	2.22	1	0.136	−0.423
27. Annoyed with oneself for smoking	1.38	1	0.240	0.846
30. Having second thoughts about one's smoking	3.28	1	0.070	1.239

COS-LC, Consequence of Screening in Lung Cancer.

in the dimensional sum scores, then the same or greater impact also needs to be identified in the new items if these are not to be regarded as "poor." Therefore, decisions about which new items to delete from the questionnaire and which items to keep as single items must be made when the results of the psychosocial consequences of lung cancer screening have been assessed.

Several items in part I of the questionnaire were identified as possessing DIF. After splitting those items in relation to the covariables possessing DIF, the fit to the Rasch model increased in all cases. These results confirmed that the identified DIF was real and was not caused by some other bias.

When developing and validating part I of the COS-BC, there were still some questions to be answered [8]. Some of these questions concerned the core-questionnaire COS. All the items in the COS were found relevant for participants in lung cancer screening. Therefore, the unanswered questions may not only be relevant in a population of women participating in breast cancer screening, but also for participants in lung cancer screening. The unanswered questions were the following:

1. Do the four items covering sleep form one dimension?
   Answer: No, one of the items, "awake most of the night," did not fit a Rasch model and must be regarded as a single item.
2. Will the item "felt sad" still have uniform DIF?
   Answer: This item did not possess DIF among participants in lung cancer screening.
3. Will some items still have thresholds that are not in order?
   Answer: All items in part I had their thresholds in order.

As expected, the themes "risk of cancer," "fear of cancer," and "smoking habits," already described in the literature as consequences of participating in lung cancer screening, were also identified in this study. Two dimensions were developed to measure the theme about smoking habits—"harm of smoking" in part I and "regretful of still smoking" in part II. The two single items included in part II (Table 1), "anxiety about lung cancer" and "belief in not having lung cancer," cover the aspects about risk and fear of lung cancer, which were also found in the Pittsburgh study [28].

The two new dimensions generated for part II (Table 3), "impulsivity" and "empathy," and the three dimensions from part II of the core-questionnaire COS (Table 1) all showed the profound changes in attitude toward life that may occur after a person has been declared free from suspicion of cancer. These dimensions are all proxy measurements for how much fear and worry the persons having abnormal screening results experienced during the critical period from the abnormal screening result to the final true- or false-positive result.

The response options in part II allow for replies in opposite directions (Fig. 2). This corresponds with the findings in a survey where patients with a diagnosis of, and treated for, cancer experienced positive as well as adverse consequences [43]. It also

corresponds with the results of the Rasch analyses of part II. All seven dimensions fitted the Rasch model in the data set that was recoded as "laterally reversed." Here, responses in any of the two opposing directions were regarded as long-term psychosocial consequences. The fit to the two other data sets recoded as "missing" or as "the same as before" was less convincing, confirming that long-term psychosocial consequences should not only be regarded as changes in one direction. Furthermore, the ordering of the items also changed for some of the items in two of the dimensions, "existential values" and "impulsivity," again due to changes in the items' response after the recoding. Therefore, the recoding "missing" and "the same as before" should only be used in dimensions where all the items' response patterns are skewed. Furthermore, if future data collected with part II are not skewed, then only the recoding as "laterally reversed" can be used for Rasch analysis.

Qualitative studies have shown that patients with a diagnosis of lung cancer feel stigmatized and blame themselves for their cancer disease if they still smoke [61]. The group interviews in the present study revealed that these themes were also important for persons with abnormal screening results that were later confirmed to be false positive. Furthermore, the Rasch analysis also confirmed that these themes were distinct dimensions that were measurable.

The authors of the Belgian-Dutch study on HRQoL effects on participation in lung cancer screening argue for the necessity of using generic instruments. They claim to be able to equate and calibrate against other adverse health outcomes and screening programs. Furthermore, they claim to be able to balance the positive and negative effects of screening against costs by using generic measures [29]. Nevertheless, besides the problem of irrelevance described in the introduction of the present article, such assessments demand invariant psychometric properties across different populations and different screening programs. In the present study, several items were identified as possessing DIF. This would most likely also be the case with the items in the generic measures. Before comparison can be done between different screening programs or with other health-related outcome measures, data collected using generic measures must be merged and tested for DIF. If DIF is identified, then it means that the comparisons between the screening programs and the cost estimations are all biased. Nevertheless, it is possible to overcome these biases if adjustment of the mean scores of the dimensions' sum scores is done [55].

Conclusion

In conclusion, the reliability and dimensionality of a condition-specific measure with high content validity for persons with an abnormal and false-positive lung cancer screening results have been demonstrated. The new questionnaire, COS-LC, covers in two parts the psychosocial experiences of lung cancer screening.

Part I: "anxiety," "behavior," "dejection," "sleep," "self-blame," "focus on airway symptoms," "stigmatization," "introvert," and "harm of smoking."

Part II: "calm/relax," "social network," "existential values," "impulsivity," "empathy," and "regretful of still smoking."

In addition, the measure includes several "poor" or single items where future studies on the actual psychosocial impact of lung cancer screening must decide whether these items should be regarded as single items or deleted from the COS-LC.

Perspectives

The COS-LC is currently in use in two surveys in Denmark [26] and The Netherlands [25]. It is planned to translate and adapt the COS-LC into British English and German to be used in surveys in the UK and Germany, respectively. Future analyses on data from these surveys will make it possible to analyze the psychometric properties of each single language version of the COS-LC and, also, to compare the psychosocial consequences of lung cancer screening between the four countries. These analyses will contribute to the final construction of the COS-LC.

The items and the dimensions listed in Table 1 comprising the core-questionnaire COS have been shown to be relevant and valid in breast cancer screening and lung cancer screening. The common core-questionnaire COS provides the opportunity for Rasch analysis to be used to compare the psychosocial consequences of the two screening programs. Nevertheless, future Rasch analysis on DIF has to be conducted to explore the consistency of the COS.

References

1 Dean C, Roberts MM, French K, Robinson S. Psychiatric morbidity after screening for breast cancer. J Epidemiol Community Health 1986;40:71–5.

2 Eardley A, Elkind A. Breast and cervical screening: looking at screening and its problems. Nurs Stand 1990;4:35–7.

3 Fallowfield LJ, Rodway A, Baum M. What are the psychological factors influencing attendance, non-attendance and re-attendance at a breast screening centre? J R Soc Med 1990;83:547–51.

4 Austoker J, Ong G. Written information needs of women who are recalled for further investigation of breast screening: results of a multicentre study. J Med Screen 1994;1:238–44.

5 Cullen J, Schwartz MD, Lawrence WF, et al. Short-term impact of cancer prevention and screening activities on quality of life. J Clin Oncol 2004;22:943–52.

6 Brett J, Bankhead C, Henderson B, et al. The psychological impact of mammographic screening. A systematic review. Psychooncology 2005;14:917–38.

7 Brodersen J. Measuring Psychosocial Consequences of False-Positive Screening Results—Breast Cancer As An Example. Copenhagen: Department of General Practice, Institute of Public Health, Faculty of Health Sciences, University of Copenhagen: Månedsskrift for Praktisk Lægegerning, 2006. ISBN: 87-88638-36-7.

8 Brodersen J, Thorsen H, Kreiner S. Validation of a condition-specific measure for women having an abnormal screening mammography. Value Health 2007;10:294–304.

9 Brodersen J, Thorsen H, Cockburn J. Validity of short-term consequences of cancer prevention and screening activities? J Clin Oncol 2005;23:244–5.

10 Brodersen J, Thorsen H, Cockburn J. The adequacy of measurement of short and long-term consequences of false-positive screening mammography. J Med Screen 2004;11:39–44.

11 McGovern PM, Gross CR, Krueger RA, et al. False-positive cancer screens and health-related quality of life. Cancer Nurs 2004;27:347–52.

12 Hobart JC, Williams LS, Moran K, Thompson AJ. Quality of life measurement after stroke: uses and abuses of the SF-36. Stroke 2002;33:1348–56.

13 Comins J, Brodersen J, Krogsgaard M, Beyer N. Rasch analysis of the Knee injury and Osteoarthritis Outcome Score (KOOS): a statistical re-evaluation. Scand J Med Sci Sports 2008;18:336–45.

14 Hagell P, Tornqvist AL, Hobart J. Testing the SF-36 in Parkinson's disease. Implications for reporting rating scale data. J Neurol 2008;255:246–54.

15 Hagell P, Relmer J, Nyberg P. Whose quality of life? Ethical implications in patient-reported health outcome measurement. Value Health 2009;12:613–17.

16 Wiebe S, Guyatt G, Weaver B, et al. Comparative responsiveness of generic and specific quality-of-life instruments. J Clin Epidemiol 2003;56:52–60.

17 Brodersen J, McKenna SP, Doward LC, Thorsen H. Measuring the psychosocial consequences of screening. Health Qual Life Outcomes 2007;5:3.

18 Bach PB, Jett JR, Pastorino U, et al. Computed tomography screening and lung cancer outcomes. JAMA 2007;297:953–61.

19 Henschke CI, McCauley DI, Yankelevitz DF, et al. Early Lung Cancer Action Project: overall design and findings from baseline screening. Lancet 1999;354:99–105.

20 Swensen SJ, Jett JR, Sloan JA, et al. Screening for lung cancer with low-dose spiral computed tomography. Am J Respir Crit Care Med 2002;165:508–13.

21 Pastorino U, Bellomi M, Landoni C, et al. Early lung-cancer detection with spiral CT and positron emission tomography in heavy smokers: 2-year results. Lancet 2003;362:593–7.

22 Welch HG, Woloshin S, Schwartz LM, et al. Overstating the evidence for lung cancer screening: the International Early Lung Cancer Action Program (I-ELCAP) study. Arch Intern Med 2007;167:2289–95.

23 The National Lung Screening Trial. Available from: http://www.cancer.gov/nlst [Accessed April 28, 2009].

24 Mascalchi M, Belli G, Zappa M, et al. Risk-benefit analysis of X-ray exposure associated with lung cancer screening in the Italung-CT trial. AJR Am J Roentgenol 2006;187:421–9.

25 van Iersel CA, de Koning HJ, Draisma G, et al. Risk-based selection from the general population in a screening trial: selection criteria, recruitment and power for the Dutch-Belgian randomised lung cancer multi-slice CT screening trial (NELSON). Int J Cancer 2007;120:868–74.

26 Pedersen JH, Ashraf H, Dirksen A, et al. The Danish randomized lung cancer CT screening trial—overall design and results of the prevalence round. J Thorac Oncol 2009;4:608–14.

27 Becker N, Delorme S, Kauczor HU. The German component of the European trial on the efficacy of multislice-CT for the early detection of lung cancer. Onkologie 2008;130(Suppl. 1):130. Ref Type: Abstract.

28 Byrne MM, Weissfeld J, Roberts MS. Anxiety, fear of cancer, and perceived risk of cancer following lung cancer screening. Med Decis Making 2008;28:917–25.

29 van den Bergh KA, Essink-Bot ML, Bunge EM, et al. Impact of computed tomography screening for lung cancer on participants in a randomized controlled trial (NELSON trial). Cancer 2008;113:396–404.

30 Ashraf H, Tonnesen P, Pedersen JH, et al. Smoking habits were unaffected by CT screening at 1-year follow-up in the Danish Lung Cancer Screening Trial (DLCST). Thorax 2009;64:388–92.

31 Spielberger CD, Gorsuch RL, Lushene RE, et al. Manual for the State-Trait Anxiety Inventory (Form Y). Palo Alto, CA: Consulting Psychologists Press, 1983.

32 Ong G, Austoker J, Brett J. Breast screening: adverse psychological consequences one month after placing women on early recall because of a diagnostic uncertainty. A multicentre study. J Med Screen 1997;4:158–68.

33 Brett J, Austoker J, Ong G. Do women who undergo further investigation for breast screening suffer adverse psychological consequences? A multi-centre follow-up study comparing different breast screening result groups five months after their last breast screening appointment. J Public Health Med 1998;20:396–403.

34 Tenenbaum G, Furst D, Weingarten G. A statistical reevaluation of the STAI anxiety questionnaire. J Clin Psychol 1985;41:239–44.

35 Brodersen J, Thorsen H. Consequences of screening in breast cancer (COS-BC): development of a questionnaire. Scand J Prim Health Care 2008;26:251–6.

36 Elsass P. Sundhedspsykologi. København: Gyldendal, 1992.

37 Iversen L, Kristensen TS, Holstein B, Due P. Medicinsk Sociologi. København: Munksgaard Danmark, 2002.

38 Engel GL. The clinical application of the biopsychosocial model. Am J Psychiatry 1980;137:535–44.

39 Engel GL. The need for a new medical model: a challenge for biomedicine. Science 1977;196:129–36.

40 Willis GB. Cognitive interviewing. A "How to" guide. Available from: http://appliedresearch.cancer.gov/areas/cognitive/interview. pdf [Accessed December 4, 2009].

41 Collins D. Pretesting survey instruments: an overview of cognitive methods. Qual Life Res 2003;12:229–38.

42 Jacobsen B, Jorgensen SD, Jorgensen SE. Kræft og eksistens—om at leve med kræft. Dansk Psykologisk Forlag 1998.

43 Carver CS, Antoni MH. Finding benefit in breast cancer during the year after diagnosis predicts better adjustment 5–8 years after diagnosis. Health Psychol 2004;23:595–8.

44 Antonovsky A. Health, Stress and Coping. San Francisco: Jossey-Bass Inc Pub, 1979.

45 Antonovsky A. Unraveling the Mystery of Health: How People Manage Stress and Stay Well. San Francisco: Jossey-Bass Inc Pub, 1987.

46 Gram IT, Lund E, Slenker SE. Quality of life following a false positive mammogram. Br J Cancer 1990;62:1018–22.

47 McDowell I, Newell C. Measuring Health. A Guide to Rating Scales and Questionnaires (2nd ed.). Oxford: Oxford University Press, 1996.

48 Streiner DL, Norman GR. Health Measurement Scales: A Practical Guide to Their Development and Use (4th ed.). New York: Oxford University Press, 2008.

49 Andrich D, Luo G. Conditional pairwise estimation in the Rasch model for ordered response categories using principal components. J Appl Meas 2003;4:205–21.

50 RUMM 2020 [computer program]. Version 4.0 for Windows. RUMM Laboratory Pty Ltd. Available from: http://www.rummlab.com/ [Accessed December 4, 2009].

51 Wright BD, Panchapakesan N. A procedure for sample-free item analysis. Educ Psychol Meas 1969;29:23–48.

52 Rasch G. An informal report on a theory of objectivity in comparisons. In: Van der Kamp LJTH, Vlek CAJ, eds. Psychological Measurement Theory. Leyden: University of Leyden, 1967.

53 Andersen EB. Sufficient statistics and latent trait models. Psychometrika 1977;42:69–81.

54 Hagquist C, Andrich D. Is the sense of coherence-instrument applicable on adolescents? A latent trait analysis using Rasch modelling. Pers Individ Dif 2004;36:955–68.

55 Brodersen J, Meads DM, Kreiner S, et al. Methodological aspects of differential item functioning in the Rasch model. J Med Econ 2007;10:309–24.

56 Hagquist C. Psychometric properties of the psychosomatic problems scale: a Rasch analysis on adolescent data. Soc Indic Res 2008;86:511–23.

57 Cronbach LJ. Coeffcient alpha and the internal structure of tests. Psychometrika 1951;16:297–334.

58 Cronbach LJ. Internal consistency of tests: analyses old and new. Psychometrika 1988;53:63–70.

59 Smith EV, Jr. Evidence for the reliability of measures and validity of measure interpretation: a Rasch measurement perspective. In: Smith EV, Jr, Smith RM, eds. Introduction to Rasch Measurement. Theory, Models and Applications (1st ed.). Maple Grove, MN: JAM Press, 2004.

60 Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc B Stat Methodol 1995;57:289–300.

61 Chapple A, Ziebland S, McPherson A. Stigma, shame, and blame experienced by patients with lung cancer: qualitative study. BMJ 2004;328:1470–4.