Respiratory Medicine (2010) **104**, 1675–1682

available at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/rmed

Reliability and validity of the clinical COPD questionniare and chronic respiratory questionnaire

Ayalu A. Reda [a,b], Daniel Kotz [a,*], Janwillem W.H. Kocks [c], Geertjan Wesseling [d], Constant P. van Schayck [a]

[a] *Department of General Practice, CAPHRI School for Public Health and Primary Care, Maastricht University Medical Centre, P.O.Box 616, 6200MD Maastricht, The Netherlands*
[b] *Department of Epidemiology, CAPHRI School for Public Health and Primary Care, Maastricht University Medical Centre, P.O.Box 616, 6200MD Maastricht, The Netherlands*
[c] *Department of General Practice, University Medical Centre Groningen, Groningen Research Institute for Asthma and COPD (GRIAC), Groningen, The Netherlands*
[d] *Department of Respiratory Medicine, CAPHRI School for Public Health and Primary Care, Maastricht University Medical Centre, P.O.Box 616, 6200MD Maastricht, The Netherlands*

Received 30 May 2009; accepted 23 April 2010
Available online 11 June 2010

KEYWORDS
Chronic obstructive pulmonary disease;
Health-related quality of life;
Questionnaire;
Reliability;
Validity;
Responsiveness

Summary
Background: Questionnaires are often used in assessing health-related quality of life in patients with chronic obstructive pulmonary disease (COPD). It is important that these questionnaires have good reliability, validity, and responsiveness. The aim of this study was to investigate and compare these properties in the disease specific Clinical COPD Questionnaire (CCQ) and the Chronic Respiratory Questionnaire self-reported (CRQ-SR).
Methods: Two hundred ninety six participants with spirometry confirmed mild to moderate COPD were included in a smoking cessation trial. It was assumed that health-related quality of life would improve in participants who stopped smoking. The questionnaires were administered at baseline and at weeks 5, 26, and 52 after the target quit date.
Results: At baseline, 292 (97%) participants returned the CCQ and 296 (100%) the CRQ-SR questionnaire. For both instruments, the internal consistency was good (Cronbach's alpha >70%) as was the convergent validity with each other but not with spirometry. The CCQ was responsive to improvements in respiratory symptoms at both week 26 (−1.02, SD = 0.81) and 52 (−1.04, SD = 0.91) and in the total score at week 26 (−0.54, SD = 0.50) and 52 (−0.43, SD = 0.44).

Abbreviations: COPD, chronic obstructive pulmonary disease; HRQoL, health-related quality of life; CCQ, clinical COPD questionnaire; CRQ, chronic respiratory questionnaire; CRQ-SR, CRQ self reported; CRQ-IL, CRQ interviewer led; GOLD, global initiative for chronic obstructive lung disease; Post-bd., post bronchodilator; FEV_1, forced expiratory volume in 1 s; FVC, forced vital capacity; MCID, minimal clinically important difference.

* Corresponding author. Tel.: +31 43 3882893; fax: +31 43 3619344.
E-mail address: D.kotz@HAG.unimaas.nl (D. Kotz).
URL: http://www.daniel-kotz.de, ResearcherID.com (A-1270-2007)

0954-6111/$ - see front matter
doi:10.1016/j.rmed.2010.04.023

The mastery domain and the total score of the CRQ-SR were responsive at week 26 (1.14, SD = 0.82; 0.67, SD = 0.97 respectively) but not at week 52 (0.04, SD = 0.93; 0.38, SD = 0.57 respectively).

Conclusion: Both the CCQ and CRQ-SR are equally reliable and valid. The long-term responsiveness of the CCQ is better. Both questionnaires can be used in future studies involving patients with mild to moderate COPD. However, when the follow-up exceeds 26 weeks, the CCQ is the recommended alternative.

Netherlands Trial Register: ISRCTN 64481813.

Introduction

Generic health-related quality of life (HRQoL) questionnaires measure general health status, while disease specific HRQoL questionnaires are concerned with specific diseases, such as chronic obstructive pulmonary disease (COPD). The Clinical COPD Questionnaire (CCQ)[1] and Chronic Respiratory Questionnaire (CRQ)[2] are disease specific HRQoL questionnaires for measuring respiratory health status.

The CCQ has three domains: symptoms, functional state, and mental state.[1] It has good validity, reliability and responsiveness in patients at risk of COPD and patients with COPD.[1,3] The CRQ is composed of four domains: fatigue, dyspnoea, mastery (the patient's feeling of control over their disease), and emotional function.[2,4] It has good convergent validity with most of the commonly used generic and disease specific HRQoL questionnaires.[4–7]

HRQoL and functional status are important aspects of COPD care. The CCQ and CRQ have shown to be adequate instruments in moderate to severe COPD[1,5,8] but evidence is limited in mild to moderate COPD, especially in the long-term follow-up. Yet, the majority of patients with COPD have mild to moderate symptoms and evidence on validity and reliability of CCQ and CRQ is necessary to improve the applicability of these instruments in this group of patients. Furthermore, the question remains which of the two questionnaires performs best in these patients. To date, there are no comparison studies between the CCQ and CRQ. In order to compare and assess the adequacy of these instruments one needs longitudinal data in which the functional status and HRQoL are expected to change clearly within patients with mild to moderate COPD. So we need a situation, such as the improvement in HRQoL and respiratory symptoms which follows quitting smoking,[9,10] in which CCQ and CRQ are measured in a standardized way.

The aim of this study was to assess and compare the medium- and longer-term reliability, validity, and responsiveness of the CCQ and the CRQ in a smoking cessation trial in smokers with mild to moderate airflow limitation.

Methods

The data for this study came from a recently completed randomized controlled trial[11] which aimed to test the efficacy of smoking cessation interventions in smokers with mild to moderate airflow limitation. Prolonged abstinence rate for the whole sample was 23% at week 26 and 10% at week 52. The results of the efficacy of the smoking cessation intervention from this trial have been discussed elsewhere.[12]

A total of 296 participants were included, all with mild (defined as Global Initiative for Chronic Obstructive Lung Disease[13] (GOLD) I) to moderate (GOLD II) airflow limitation, confirmed by spirometry post bronchodilator (post-bd.) forced expiratory volume in 1 s (FEV_1)/forced vital capacity (FVC) of <0.70 in combination with post-bd. $FEV_1 \geq 50\%$ of predicted. For the discriminative validity analysis we included an external sample of 587 smokers with normal lung function who were screened for participation but were found not eligible for the trial. Prolonged abstinence from smoking was defined as self-reported non-smoking at week 5, 26, and 52 after the target quit date, validated by a urine cotinine test (<50 ng/ml).[11] The study was approved by the ethical review committee of Maastricht University Medical Centre and registered at the Netherlands Trial Register (ISRCTN 64481813).

Clinical COPD questionnaire (CCQ)

The CCQ has three domains: symptoms (4 items), functional state (4 items) and mental state (2 items), graded on a 7-point Likert scale from 0 to 6. Lower scores indicate better health status. Analyses were conducted for participants with at least 50% completion rate based on the recommendation of the questionnaire designers.

Chronic respiratory questionnaire self-reported (CRQ-SR)

The underlying structures, content and scoring of the interviewer lead CRQ (CRQ-IL) and the self-reported CRQ (CRQ-SR) are the same.[4,8,14] In the dyspnoea domain (5 items), patients select from a list of activities which make them breathless, and they may also add additional activities.[4] The other three domains are standardized: fatigue (4 items), mastery (4 items), and emotional function (7 items). This provides a 7-point Likert score that enables comparison among the four domains. Lower scores indicate a greater degree of dysfunction. Data of smokers who responded to respectively 2, 2, and 4 or more items of the fatigue, mastery, and emotional function domain were included in the analyses. This gives at least a 53.3% item-completion rate for the three domains: all items of the dyspnoea domain were included in the analysis. All data analyses were conducted for participants with at least 65% item-completion rate.

Participants completed the CCQ and CRQ-SR at home before the baseline measurement (before randomisation) and at weeks 5, 26, and 52 after the target quit date. Mean

scores were calculated by dividing the total score per domain by the corresponding number of items.

Spirometry was conducted at baseline according to the European Respiratory Society/American Thoracic Society guidelines.[15,16] In this report, we used the post-bd. FEV_1/FVC and FEV_1 measurements.[15,16]

Data analysis

'Ceiling effect' (best health score) and 'floor effect' (worst health score) were analysed at baseline. We assessed both minimal clinically important differences (MCID) and statistically significant differences. The MCID for the CCQ and the CRQ-SR were previously estimated to be 0.4[17] and 0.5 respectively.[18] To correct for multiple testing, alpha was set at 0.001. SPSS 15 was used for the analyses.

Sensitivity analyses were conducted in participants with complete response to all items of the questionnaires that are included in a given analysis in order to examine the validity of our findings. We statistically analysed the influence of background characteristics such as age and sex on completion rates. We used *t*-tests or Mann–Whitney *U* tests for numerical variables and chi-squared test for categorical variables.

Domain specific consistency assessment was conducted to examine reliability. A Cronbach's alpha of more than 0.70 was considered good consistency.[19] We compared the discriminative property of the questionnaires separately between participants with GOLD I ($FEV_1/FVC < 0.70$, $FEV_1 \geq 80\%$) and GOLD II ($FEV_1/FVC < 0.70$, $50\% \leq FEV_1 < 80\%$) COPD; and between participants with COPD (GOLD I or II) and with normal lung function. Convergence between the two questionnaires, and the post-bd. FEV_1 measurements was examined using Spearman's rho.

Responsiveness is the ability of an instrument to measure a clinically meaningful change.[19] Theoretical constructs are commonly employed in validity testing as there are no gold standard tests. The underlying assumption in our responsiveness analysis was that smoking cessation improves HRQoL in smokers in the medium- and long term.[10,20,21] We analysed the mean score changes from baseline for prolonged abstainers at 26 and 52 weeks to test medium and long-term responsiveness, respectively. The 5-week scores were not used because quitting smoking may reduce quality of life in the short term.[9,10,22]

Results

The mean age of participants was 54 years (SD $\pm$ 7.5). Baseline characteristics of participants are shown in Table 1.

At baseline, 292 (97%) participants returned the CCQ and 296 (100%) the CRQ-SR questionnaire. Two hundred sixty eight (92%) and 275 (93%) participants completed all items of the CCQ and CRQ-SR respectively, while 279 (96%) and 296 (100%) completed at least 50% and 65% of the items on the CCQ and CRQ-SR respectively. Baseline characteristics were not statistically associated with item-completion rates in both questionnaires.

One hundred and forty nine (51%) and 146 (49%) participants completed all items of the CCQ and CRQ-SR at baseline and at the follow-up visits in week 5, 26 and 52.

Score distributions

There was no 'floor' effect for both questionnaires. However, there was a modest amount of 'ceiling' effect for some domains (Table 2).

Reliability

The Cronbach's alpha for the symptoms, functional state, and mental state of the CCQ was 0.73, 0.77 and 0.59 respectively. This figure was 0.88, 0.91, 0.75, and 0.89 for the fatigue, dyspnoea, mastery and emotional function domains of the CRQ-SR.

Discriminant validity

There was a statistically significant difference in age between the external sample and study participants where the external sample had a mean age of 48.7 (SD = 8.1) years and trial participants 54.0 (SD = 7.5) ($Z = -9.02$, $p < 0.001$). The pack years of smoking for the external sample was 36.2 (SD = 19.4) and 43.5 (SD = 18.9) for trial participants ($Z = -6.11$, $p < 0.001$). The domains and total score of both questionnaires did not discriminate between participants with COPD (GOLD I and II) and participants with no airflow limitation. Unlike the domains of the CRQ-SR, the total score of the CCQ was able to discriminate statistically between participants with GOLD I and GOLD II COPD (df 277, $t = 3.7$, $p < 0.001$), where the latter group had worse HRQoL (Table 3).

Convergent validity

There was a statistically significant ($p < 0.001$) modest correlation between the baseline scores of the CCQ and CRQ-SR. FEV_1 scores had a very low correlation with both the CCQ and the CRQ-SR (Table 4).

Responsiveness

Both questionnaires detected a clinically meaningful and statistically significant improvement in HRQoL in participants with prolonged abstinence at week 26; this was 1.14 (SD = 0.82) for the mastery domain of the CRQ-SR and −1.02 (SD = 0.81) for the symptom domain of the CCQ (Table 5). In Fig. 1, we standardized (mean/SD) the responsiveness indicators and scaled improvement in HRQoL as positive (and worsening negative) to aid between questionnaire comparisons. As the figure shows, the CCQ performed better.

The CCQ detected a clinically meaningful and statistically significant improvement in HRQoL in participants with prolonged abstinence at week 52; this was −1.04 (SD = 0.97) for the symptom domain (Table 5). The CRQ-SR domains and total score were unable to detect a clinically meaningful and statistically significant improvement at week 52 except its dyspnoea domain which indicated a clinically meaningful

Table 1 Baseline characteristics of the study population ($N = 296$).[a]

	Sex	
Characteristic (Mean, SD)	Female (185, 62.5%)	Male (111, 37.5%)
Age (years)	54.6 (7.7)	53.0 (7.2)
BMI at baseline (kg/m^2)[b]	25.6 (4.0)	24.1 (4.0)
Number of cigarettes	23.9 (10.0)	22.4 (7.8)
Pack years at baseline[c]	45.3 (20.1)	40.5 (16.4)
Fagerström Test for Nicotine Dependence[d]	4.6 (1.6)	4.4 (1.4)
FEV_1 post-bd. %predicted baseline	80.9 (14.2)	82.5 (17.3)
FVC post-bd. %predicted baseline	102.9 (14.8)	110.1 (16.9)
FEV_1/FVC post-bd.[e]	62.3 (6.2)	63.0 (6.0)
GOLD[f] COPD classification, N (%)		
GOLD I (mild COPD)	103 (64.4%)	57 (35.6%)
GOLD II (moderate COPD)	82 (60.3%)	54 (39.7%)

[a] Mean (SD) unless indicated otherwise.
[b] Body Mass Index.
[c] 1 pack year = number of cigarettes smoked per day × number of years smoking/20.
[d] Fagerström Test for Nicotine Dependence: ranges from 0 (lowest nicotine dependence) to 10 (highest nicotine dependence).
[e] Post bronchodilator forced expiratory volume in 1 s/forced vital capacity.
[f] Global Initiative for Chronic Obstructive Lung Disease.

improvement. In Fig. 2, we standardized the responsiveness indicators. Also here, the CCQ performed better.

As a sensitivity analysis, we re-ran all analyses using data of smokers with 100% item-completion rate to each questionnaire. There was no difference in the validity and reliability estimates for both questionnaires.

Discussion

We examined and compared the validity, reliability, and responsiveness of the Clinical COPD Questionnaire (CCQ) and the Chronic Respiratory Questionnaire self-reported (CRQ-SR) in smokers with mild to moderate airflow limitation. Both questionnaires showed good internal consistency and good convergent validity with each other but not with spirometry. The CRQ-SR showed good responsiveness for the mastery domain and total score after 26 weeks follow-up, and the CCQ showed good responsiveness for the symptom domain and total score after both medium- (26 weeks) and longer term (52 weeks) follow-up.

We detected no difference in baseline characteristics between participants with item-completion rates below and above 50% and 65% of items of CCQ and CRQ, respectively. This could indicate that the items were approximately randomly missing with regard to the baseline characteristics such as age and pack years smoking and hence may not have biased our analyses significantly. This rate is similar or equal to previous studies for both questionnaires.[1,6]

Because of the larger number of items and the longer time needed for completion of the CRQ-SR, the CCQ seems to be favoured in terms of feasibility. This has important implication for both clinical practice and research.

Although the CCQ performed somewhat better, both questionnaires discriminated poorly between GOLD I and II COPD. This could be because the questionnaires were not designed for discriminative purposes. The inclusion of mild

Table 2 Scores of participants at baseline assessment.[a]

	N	Mean (SD)	Median	'Floor' effects, % n (worst health score)	'Ceiling' scores, % n (best health score)
CRQ − SR					
Fatigue	296	4.7 (1.2)	4.8	0.3 (1)	1.0 (3)
Emotional function	296	4.8 (1.1)	4.7	0.0 (0)	1.7 (5)
Mastery	296	5.2 (1.3)	5.3	0.0 (0)	12.9 (49)
Dyspnoea	205	3.8 (1.2)	3.7	0.5 (1)	1 (2)
Total	296	4.7(0.9)	5.0	0.0(0)	0.0(0)
CCQ					
Symptom	294	2.0 (1.1)	1.8	0.3 (1)	0.7 (2)
Functional state	286	0.9 (0.8)	0.8	0.0 (0)	14.7(42)
Mental state	284	0.9 (1.0)	0.5	0.0 (0)	37.0(105)
Total	279	1.2 (0.8)	1.0	0.0 (0)	0.7(2)

[a] The CRQ-SR was rated from 1 'worst health condition' to 7 'best health condition'. The CCQ is rated from 0 'best health condition' to 6 'worst health condition'.

Table 3 CCQ and CRQ-SR mean (range) scores at baseline within subgroups.[a]

	Normal lung function (GOLD 0)	Mild COPD (GOLD I)	Moderate COPD (GOLD II)
CRQ-SR			
Fatigue	4.6 (1.0—7.0)	4.7 (1.0—7.0)	4.6 (1.2—7.0)
Emotional function	4.7 (1.3—7.0)	4.8 (1.6—7.0)	4.7 (1.7—7.0)
Mastery	5.0 (3.0—7.0)	5.0 (2.7—7.0)	4.9 (1.7—7.0)
Dyspnoea	3.8 (1.0—7.0)	3.7 (1.0—6.0)	3.8 (1.6—7.0)
Total	4.6 (1.9—7.0)	4.7 (2.3—6.8)	4.6 (2.5—6.6)
CCQ			
Symptom	1.8 (0.0—5.5)	1.8 (0.0—6.0)	2.2[b,c] (0.3—5.3)
Functional state	0.9 (0.0—5.5)	0.8 (0.0—5.0)	1.0 (0.0—4.0)
Mental state	0.9 (0.0—6.0)	0.8 (0.0—5.0)	1.0 (0.0—4.0)
Total	1.3 (0.0—4.0)	1.1 (0.0—4.3)	1.5[c] (0.2—4.7)

[a] The CCQ is rated from 0 'best health state' to 6 'worst health state' and CRQ-SR from 1 'worst health state' to 7 'best health state'.
[b] significant difference between GOLD 0 and GOLD II participant scores.
[c] significant difference between GOLD I and GOLD II participant scores.

to moderate COPD patients only may have also influenced the performance of the questionnaires as compared to the situation in which a diverse case mix is available, especially with more severe cases of COPD.

The internal consistency of the CCQ was lower than that of the CRQ-SR which reached 0.90 for the emotional and fatigue domains. The higher alpha for the CRQ-SR could be due the higher number of items for the domains such as emotional function. The internal consistency of the CRQ-SR in this report is similar to[8] or higher[6] than earlier reports. Even though previous findings indicated a low internal consistency of the dyspnoea domain in the range of 0.53—0.64,[6,8] we found a very high figure (0.91). Interestingly, our finding is similar to that Rutten-van Mölken et al. reported (0.86).[7] In general, both questionnaires had good internal consistency, supporting the assumption that the respective questionnaire items are related to each other.

The statistically significant albeit modest correlation between the two questionnaires could indicate that they may be measuring the same target clinical condition. We could not compare our findings with that of other investigators since this study was the first to compare the CCQ and CRQ-SR. There was very poor linear correlation between the CCQ and CRQ-SR with FEV_1 post-bd measurements at baseline. This finding is similar to other studies on the CRQ-SR[7] and CRQ -IL[8] but not for the CCQ in which a Spearman's rho of −0.49 was reported.[1] So far no explanation exists for the low correlation of these instruments with lung function.

The total score of the CCQ and its symptom domain, and the total score and mastery domain of the CRQ-SR indicated statistically significant and clinically meaningful differences in the medium term. However, the responsiveness indicators of the CRQ-SR diminished at 52 weeks unlike the CCQ, which remained stable. Particularly, the mastery domain made a noticeable decline from a statistically and a clinically meaning meaningful difference of 1.14 to a non significant difference of 0.04 units. The mental state domain of the CCQ was not responsive probably because of ceiling effect (37%). The attenuation of the responsiveness we detected for the CRQ-SR has been reported for the interviewer-led version previously.[6] This could not be explained by regression to the mean as the CCQ detected statistically and clinically significant changes at both 26 and 52 weeks in line with our expectation. Therefore, as the CRQ is unstable, its long-term responsiveness in clinical trials involving smokers with mild to moderate COPD is questionable.

We used a p-value of 0.001 to guard against false positive findings as other investigators did.[7] We chose this figure

Table 4 Correlation between the CCQ, CRQ-SR and FEV_1 post-bd. baseline measurements.[a]

	CCQ domains				
	Total	Symptom	Functional state	Mental state	FEV_1
CRQ-SR domains					
Fatigue	−0.58*	−0.35*	−0.53*	−0.57*	0.05
Emotional function	−0.61*	−0.42*	−0.43*	−0.65*	0.06
Mastery	−0.56*	−0.43*	−0.50*	−0.46*	0.16
Dyspnoea	−0.42*	−0.35*	−0.43*	−0.20*	−0.06
Total	−0.72*	−0.50*	−0.62*	−0.64*	0.08
FEV_1post-bd baseline[b]	−0.19	−0.20	−0.18	−0.12	1.00

*$p < 0.001$.
[a] Spearman's rank correlation (ρ).
[b] forced expiratory volume in 1 s.

Table 5 Mean change in baseline scores[a] among participants with prolonged abstinence from smoking at 26 weeks and 52 weeks after the target quit date.

Questionnaire	26 weeks		52 weeks	
CRQ-SR domains	Mean difference (SD)	N	Mean difference (SD)	N
Fatigue	0.27 (1.23)	69	0.18 (0.81)	30
Emotional function	0.36 (1.07)	69	0.40 (0.67)	30
Mastery	Δ1.14 (0.82)*	68	0.04 (0.93)	30
Dyspnoea	Δ0.73 (1.20)	25	Δ0.60 (1.18)	16
Total	Δ0.67(0.92)*	69	0.38(0.57)	30
CCQ domains				
Total	Δ0.54 (0.50)*	62	Δ0.43(0.44)*	28
Symptom	Δ1.02 (0.81)*	67	Δ1.04 (0.97)*	30
Functional state	0.36 (0.54)	65	0.19 (0.52)	28
Mental state	0.25 (1.06)	63	0.11 (0.57)	28

*$p < 0.001$.
Wilcoxon Signed ranks test or related samples *t*-test were used to assess statistical significance. Δ change in unstandardized scores greater than MCIDs, 0.4 unit for the CCQ and 0.5 unit for the CRQ. The CCQ is rated from 0 'best health state' to 6 'worst health state' and CRQ-SR from 1 'worst health state' to 7 'best health state'.
[a] Positve differences were coded to indicate improvement, negative worsening.

as corrections such as Bonferonni could have provided an extremely conservative *p*-value due to the large number of tests conducted. As we also used MCID as a significance indicator, the actual significance level is higher than 0.001.

One of the limitations of our study is that it is based on participants in the setting of an intervention trial. Our participants were all current heavy smokers at study entry who were motivated to quit smoking under the controlled conditions of a trial. Hence, our findings are not representative of the whole population of mild to moderate COPD patients that are seen under routine primary care conditions. Furthermore, as we used data from a clinical trial and because the scales were administered by more than the recommended time of two weeks for conducting test-retest reliability, we were not able to calculate intra-class correlation coefficients. As a result, reliability analysis did not proceed beyond internal consistency. Even though we had good response rates at baseline as the participants were motivated to participate, respondents failed to complete all

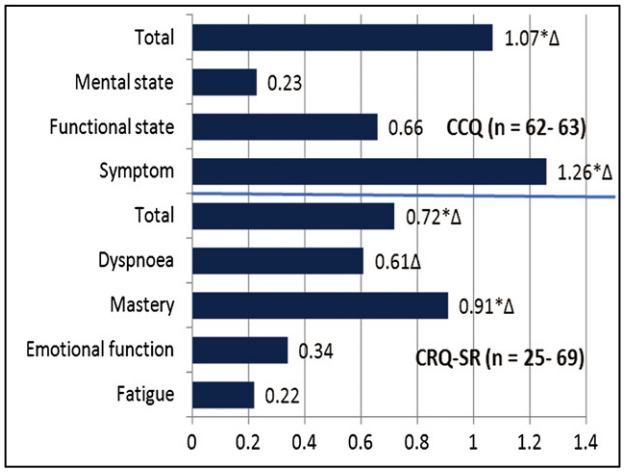

Figure 1 The standardized mean difference (SMD) in scores between baseling and 26 weeks follow-up provides an insight into the difference in score changes between baseline and six months among continous abstainers while overcoming the difference in scale of the questionnaires. (SMD = mean difference divided by the within group SD. Positive differences were scaled to indicate improvement, negative otherwise for both instruments. Δ change in unstandardized scores greater than respective MCIDs − 0.4 for the CCQ, 0.5 for the CRQ. The CCQ was rated from 0 'best health state' to 6 'worst health state' and CRQ-SR from 1 'worst health state' to 7 'best health state'. *$p < 0.001$).

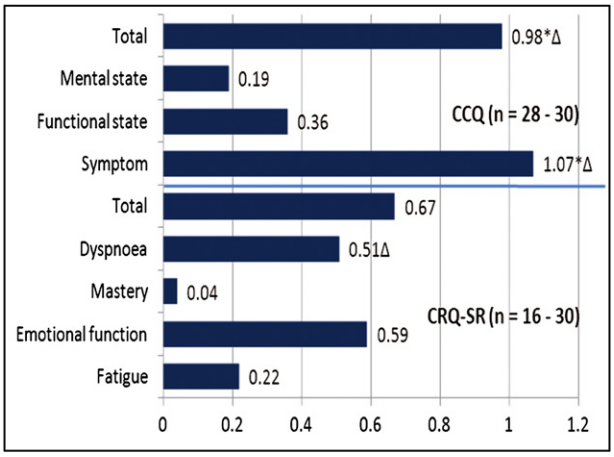

Figure 2 The standardized mean difference (SMD) in scores between baseline and 52 weeks follow-up provides an insight into the difference in score changes between baseline and six months among continous abstainers while overcoming the difference in scale of the questionnaires. (SMD = mean difference divided by the within group SD. Positive differences were scaled to indicate improvement, negative otherwise for both instruments. Δ change in unstandardized scores greater than respective MCIDs − 0.4 unit for the CCQ, 0.5 unit for the CRQ. The CCQ was rated from 0 'best health state' to 6 'worst health state' and CRQ-SR from 1 'worst health state' to 7 'best health state'. *$p < 0.001$).

items as the questionnaires were self-administered. Even though item-completion rate was not statistically associated with important baseline characteristics we may not rule out the possibility of bias. This fact, and partly, the low number of quitters, has reduced the effective sample size available for analysis. As a result we were not able to conduct validity and reliability tests by stratifying the sample by sex, age etc. The absence of evidence based guidelines on item-level missing value replacement is also a limitation. However, to examine the influence of missing values we conducted sensitivity analyses using 100% item-completion rates. The analyses showed limited variation of the key reliability and validity indicators.

This study is the first to examine and compare the CCQ and CRQ-SR questionnaires. Our report has one of the longest follow-up periods compared to previous studies. It is also the first to examine the longer term responsiveness of the CCQ. We have also effectively used current knowledge on the importance of smoking cessation in improving patients' HRQoL[10,20,21] unlike previous studies that employed specific medical symptoms and signs to examine COPD progression.

We conclude that the CCQ is a valid, reliable and responsive instrument in cases with mild to moderate COPD for both medium and longer term follow-up. The CRQ-SR has good indicators for the medium term but its responsiveness declines in the longer term. There is poor discriminative validity for both questionnaires. Both questionnaires have good convergent validity with each other, but not with FEV_1 post-bd. measurements. Thus, the two questionnaires are suitable for prospective monitoring of self-reported respiratory health status but are not suitable for diagnostic purposes. Future longitudinal studies involving mild to moderate cases of COPD could use both the CCQ and CRQ-SR if their time frame is limited to medium term time frames such as 6 months. However, when the follow-up exceeds 6 months, the CCQ is the recommended alternative. This is also considering the favourable feasibility of the CCQ.

Acknowledgement

The authors gratefully acknowledge the contributions of Kitty van der Meer (telephonic screening, logistics), Arja van de Voorde, Ellen de Goeij (spirometry), Mischa Aussems (validation spirometric test results), Willem de Goeij (database construction), Ger Driessen and Paula Rinkens (data cleaning) and Professor Thys van der Molen (for providing us information on the CCQ).

This study was funded by grants from the Dutch Asthma Foundation, Partners in Care Solutions for COPD (an initiative of CAPHRI School for Primary Care and Public Health, Boehringer Ingelheim, and Pfizer), and Maastricht University Medical Centre. The authors' work was independent of the funders.

Contribution of the authors

AAR and DK conceived the study. CPS, GW and DK were involved in its planning and implementation. AAR took the lead in data analyses and writing of the report. All co-authors critically commented on several drafts of the manuscript.

Conflict of interest statement

The authors of this manuscript declare not to have any conflict of interest regarding this manuscript. None of the authors have any financial interests with any commercial entity that has interest in the subject- or outcome of this manuscript including consultancy, stock ownership, paid expert consultancy, or honoraria, patent application, as well as other forms of conflict of interest, including personal and academic issues. The authors to the best of their knowledge conducted the study and reported the conclusions independently without any interference from partial or full funding sources or other entities.

References

1. van der Molen T, Willemse BW, Schokker S, et al. Development, validity and responsiveness of the Clinical COPD Questionnaire. *Health Qual Life Outcomes* 2003;**1**:13.
2. Guyatt GH, Berman LB, Townsend M, et al. A measure of quality of life for clinical trials in chronic lung disease. *Thorax* 1987;**42**:773–8.
3. Jones PW, Quirk FH, Baveystock CM, et al. A self-complete measure of health status for chronic airflow limitation. The St. George's Respiratory Questionnaire. *Am Rev Respir Dis* 1992; **145**:1321–7.
4. Williams JE, Singh SJ, Sewell L, et al. Development of a self-reported chronic respiratory questionnaire (CRQ-SR). *Thorax* 2001;**56**:954–9.
5. Singh SJ, Sodergren SC, Hyland ME, et al. A comparison of three disease-specific and two generic health-status measures to evaluate the outcome of pulmonary rehabilitation in COPD. *Respir Med* 2001;**95**:71–7.
6. Harper R, Brazier JE, Waterhouse JC, et al. Comparison of outcome measures for patients with chronic obstructive pulmonary disease (COPD) in an outpatient setting. *Thorax* 1997;**52**:879–87.
7. Rutten-van Molken M, Roos B, Van Noord JA. An empirical comparison of the St George's respiratory questionnaire (SGRQ) and the chronic respiratory disease questionnaire (CRQ) in a clinical trial setting. *Thorax* 1999;**54**:995–1003.
8. Wijkstra PJ, TenVergert EM, Van Altena R, et al. Reliability and validity of the chronic respiratory questionnaire (CRQ). *Thorax* 1994;**49**:465–7.
9. *Diagnostic and statistical manual of mental disorders*. APA; 1994. DSM-IV:243–245.
10. Kanner RE, Anthonisen NR, Connett JE. Lower respiratory illnesses promote FEV(1) decline in current smokers but not ex-smokers with mild chronic obstructive pulmonary disease: results from the lung health study. *Am J Respir Crit Care Med* 2001;**164**:358–64.
11. Kotz D, Wesseling G, Huibers MJ, et al. Efficacy of confrontational counselling for smoking cessation in smokers with previously undiagnosed mild to moderate airflow limitation: study protocol of a randomized controlled trial. *BMC Public Health* 2007;**7**:332.
12. Kotz D, Wesseling G, Huibers MJ, et al. Efficacy of confronting smokers with airflow limitation for smoking cessation. *Eur Respir J* 2009;**7**:7.
13. Rabe KF, Hurd S, Anzueto A, et al. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease: GOLD executive summary. *Am J Respir Crit Care Med* 2007;**176**:532–55.
14. de Torres JP, Pinto-Plata V, Ingenito E, et al. Power of outcome measurements to detect clinically significant changes in

pulmonary rehabilitation of patients with COPD. *Chest* 2002; **121**:1092—8.
15. Miller MR, Crapo R, Hankinson J, et al. General considerations for lung function testing. *Eur Respir J* 2005;**26**:153—61.
16. Miller MR, Hankinson J, Brusasco V, et al. Standardisation of spirometry General considerations for lung function testing. *Eur Respir J* 2005;**26**:319—38.
17. Kocks JW, Tuinenga MG, Uil SM, et al. Health status measurement in COPD: the minimal clinically important difference of the clinical COPD questionnaire. *Respir Res* 2006;**7**:62.
18. Redelmeier DA, Guyatt GH, Goldstein RS. Assessing the minimal important difference in symptoms: a comparison of two techniques. *J Clin Epidemiol* 1996;**49**:1215—9.
19. Streiner DL, Norman RG. *Health measurement scales: a practical guide to their development and use*. 3rd ed. Oxford University Press; 2003.
20. Anthonisen NR, Lindgren PG, Tashkin DP, et al. Bronchodilator response in the lung health study over 11 yrs. *Eur Respir J* 2005;**26**:45—51.
21. van Schayck OC, Pinnock H, Ostrem A, et al. IPCRG Consensus statement: tackling the smoking epidemic - practical guidance for primary care. *Prim Care Respir J* 2008;**17**:185—93.
22. Foulds J. The neurobiological basis for partial agonist treatment of nicotine dependence: varenicline. *Int J Clin Pract* 2006;**60**:571—6.