

An empirical comparison of the St George's Respiratory Questionnaire (SGRQ) and the Chronic Respiratory Disease Questionnaire (CRQ) in a clinical trial setting

Maureen Rutten-van Mólken, Bianca Roos, J A Van Noord

Abstract

Background—The Chronic Respiratory Questionnaire (CRQ) and the St George's Respiratory Questionnaire (SGRQ) are the two most widely used quality of life questionnaires in chronic obstructive pulmonary disease (COPD). A study was undertaken to compare directly the self-administered version of the CRQ and the SGRQ with respect to feasibility, internal consistency, validity, and sensitivity to changes resulting from bronchodilator therapy.

Methods—One hundred and forty four patients with moderate or severe COPD were randomly assigned to receive three months of treatment with either salmeterol, salmeterol + ipratropium bromide, or placebo. Quality of life was measured at baseline and after 12 weeks of treatment.

Results—The proportions of missing values per patient were low for both questionnaires (0.54% for the CRQ and 2% for the SGRQ). The internal consistency was good for both questionnaires (Cronbach's α coefficients ≥ 0.84 for the CRQ and ≥ 0.76 for the SGRQ). Factor analysis confirmed the original domain structure of the CRQ but not of the SGRQ. Correlations with forced expiratory volume in one second (FEV₁) % predicted and peak expiratory flow rate (PEFR) were low for both questionnaires but better for the SGRQ than for the CRQ. The ability to discriminate between subjects with different levels of FEV₁ was somewhat better for the SGRQ. The correlations with symptom scores were comparable for both questionnaires. Cross sectionally, the scores of the two questionnaires were moderately to highly correlated (coefficients ranged from 0.35 to 0.72). Longitudinally, these correlations were lower (coefficients ranged from 0.17 to 0.54) but were still significant. The CRQ total and emotions score and the SGRQ symptoms score were the most responsive to change. The SGRQ symptoms domain was the only domain where the improvement in patients receiving combination treatment crossed the threshold for clinical relevance.

Conclusions—Since this analysis of reliability, validity, and responsiveness to change did not clearly favour one instrument above the other, the choice between the CRQ and the SGRQ can be based on

other considerations such as the required sample size or the availability of reference values.

(Thorax 1999;54:995–1003)

Keywords: chronic obstructive pulmonary disease; health status assessment; quality of life; St George's Respiratory Questionnaire; Chronic Respiratory Disease Questionnaire

Over the last decade the potential value of incorporating disease specific quality of life questionnaires into clinical trials evaluating medical interventions for chronic obstructive pulmonary disease (COPD) has become well established. Since the associations between spirometric values and disease specific quality of life questionnaires have been shown to be weak, it is generally recognised that these instruments contribute valid additional information about the daily functioning and well being of patients with COPD. Two disease specific quality of life questionnaires are widely used in patients with COPD: the Chronic Respiratory Disease Questionnaire (CRQ) developed by Guyatt *et al* in 1987¹ and the St George's Respiratory Questionnaire (SGRQ) developed by Jones *et al* in 1991.^{2,3} The SGRQ is a self-administered questionnaire whereas the CRQ was developed to be interviewer-administered (at a later stage a self-administered version of the CRQ became available). Both questionnaires have been used to assess the effects of interventions such as drug therapies,^{3–6} oxygen therapy,^{7,8} and pulmonary rehabilitation or training programmes.^{9–12}

Both questionnaires have already been shown to be reliable and valid, but there are few direct comparisons of the two questionnaires in the same patients. One study compared their discriminative properties and reported no substantial differences between the two questionnaires in the correlation with physiological parameters, exercise capacity, dyspnoea, anxiety, and depression.¹³ Another study using the interviewer-administered version of the CRQ in routine clinical practice looked into the responsiveness of both questionnaires by comparing patients who said their health had changed with patients who said their health had not changed.¹⁴ No differences in responsiveness between the two questionnaires were found. To our knowledge there are no papers comparing the performance of both question-

Institute for Medical Technology Assessment, Erasmus University Rotterdam, 3000 DR Rotterdam, The Netherlands
M Rutten-van Mólken
B Roos

Atrium Hospital, Heerlen, The Netherlands
J A Van Noord

Correspondence to:
Dr M Rutten-van Mólken.

Received 15 February 1999
Returned to authors
11 May 1999
Revised manuscript received
30 June 1999
Accepted for publication
12 July 1999

naires in a randomised clinical trial assessing different bronchodilator therapies.

A study was undertaken to compare the SGRQ and CRQ in terms of feasibility, reliability, validity, and sensitivity to change when applied to patients with moderate to severe COPD who participated in a clinical trial comparing three different bronchodilator therapies: (1) salmeterol, (2) salmeterol plus ipratropium bromide, and (3) placebo. We used the self-administered version of the CRQ which has not been validated on a large scale. This direct comparison was deliberately performed within a clinical trial to enable a comparison of the ability to detect change.

Methods

PATIENTS

Patients aged between 40 and 75 years of age diagnosed with COPD according to the ATS criteria were included in the study. The pre-bronchodilator FEV₁ had to be between 40% and 65% of predicted normal and the FEV₁/FVC was required to be ≤60% predicted after inhalation of 200 µg salbutamol. They were required to experience symptoms on mild exertion such as getting dressed or washed on at least four out of seven days of the run in period. Current smokers or previous smokers with a smoking history equivalent to 10 pack years were included. Previous smokers should have stopped smoking for at least six months prior to the run in visit. Patients were excluded if they required regular oxygen therapy, had a history of asthma, allergic rhinitis, or other respiratory diseases, were atopic, had a serious concomitant disease, a respiratory tract infection, or a change in medication within six weeks of the commencement of the study. All maintenance drugs other than the study medication were continued at constant dosage throughout the study.

STUDY DESIGN

The study was conducted in three Dutch outpatient clinics and had a double blind, double dummy, parallel group design. Patients were randomly assigned to one of the following three groups, each of which received treatment for 12 weeks: (1) 50 µg salmeterol twice daily plus 40 µg ipratropium bromide four times daily; (2) 50 µg salmeterol twice daily plus placebo matched to ipratropium bromide; or (3) placebo matched to ipratropium bromide plus placebo matched to salmeterol. Randomisation was stratified by current smoker or ex-smoker. Patients were given Ventolin metered dose inhalers or Rotadisks for symptomatic relief. They were scheduled to visit the clinic at run in, at the randomisation visit after two weeks of run in, after four, eight and 12 weeks of treatment, and two weeks after cessation of the treatment. Quality of life measurements were conducted at the end of the run in period and after 12 weeks of treatment. These measurements were conducted before the lung function measurements. During the second administration of the quality of life questionnaires the patients were shown the answers they had given

at baseline.¹⁵ The study was approved by the medical ethics committees of the three participating centres.

SYMPTOMS

Daytime symptoms were recorded in the diary on a six point scale ranging from 0 (no symptoms at rest or on exertion) to 5 (severe symptoms at rest, unable to perform any exertion). Night-time symptom scores ranged from 0 (no symptoms during the night) to 4 (symptoms so severe that they did not sleep at all).

CLINICAL OUTCOME MEASURES

Pre-bronchodilator FEV₁ and peak expiratory flow rate (PEFR) were assessed at each clinic visit using the best of three measurements. Patients were given a diary to record three PEFR measurements every morning and every evening. In this diary they also recorded the number of daytime and night-time puffs of rescue medication and their symptom score. Diary card data for each patient were summarised by calculating means and standard deviations for the two week run in period and the last four weeks of the treatment period.

HEALTH STATUS MEASURES

St George's Respiratory Questionnaire (SGRQ)

The SGRQ is designed to measure the impact of chest disease on health related quality of life and well being.^{2,3} It can be used in COPD as well as in asthma. The responses to its 50 items can be aggregated into an overall score and three subscores for symptoms, activity, and impact. The number of response options per question varies from two to five. Responses are weighted and scores are calculated by dividing the summed weights by the maximum possible weight and expressing the result as a percentage, 0% being the best possible score and 100% the worst. The weights appear to be similar in different countries and languages.¹⁶ The minimally clinically important response to treatment is defined as an improvement of 4% on the separate domains and the total score.² The SGRQ was reported to be valid and reliable in patients with asthma and COPD.^{2,3,17,19}

Chronic Respiratory Disease Questionnaire (CRQ)

The CRQ was the first instrument developed to measure quality of life of patients with COPD.¹ It cannot be used for asthma. It is available in an interview version and a self-administered version. We have used the self-administered version, with a recall period of four weeks. The questionnaire contains 20 questions which can be divided into four domains: dyspnoea (5 questions), fatigue (4 questions), emotional functioning (7 questions), and mastery (4 questions). The dyspnoea domain is "individualised" which means that it is made up of five activities chosen by a patient to cause the greatest shortness of breath. The patient then rates the dyspnoea on these self-selected activities during subsequent administrations of the CRQ. Answers can be scored on a seven point scale ranging from 1

which indicates maximum impairment to 7 which indicates no impairment. The results are expressed as the mean score for each domain and the mean overall score. The minimally important difference was found to be an improvement (or deterioration) of 0.5.^{20, 21} The CRQ was reported to be reliable and valid.^{4, 22}

The SGRQ was originally developed in British English and the CRQ was originally developed in Canadian English. We have used existing standard Dutch versions of the original questionnaires. The translation procedure to obtain these Dutch versions included minimally forward-backward translation, discussion of disagreement over wording by bilingual speakers, and pilot testing in COPD patients.

ANALYSIS

Baseline characteristics

Analysis of differences in baseline characteristics between the treatment groups were based on analysis of variance for continuous data and Kruskal-Wallis tests for ordinal data. Differences in proportions were analysed using χ^2 tests.

Questionnaire content

We performed a qualitative comparison of the content of the domains of the two questionnaires. If the content of two domains was judged to refer to the same underlying concept, they were judged comparable.

Feasibility

A qualitative assessment of the feasibility was performed by looking at the length of the questions, the way they are phrased, the difficulty of and variations in the response options, and the variations in recall periods. Feasibility was further judged by the percentage of missing values per item and the percentage of missing items per patient. In addition, the percentages of multiple responses to items where only one response was required were compared.

Reliability

The internal consistency was expressed in Cronbach's α coefficients.²³ An α coefficient of 0.70 was considered sufficient. Because this study was an intervention study, data on test-retest reliability in stable subjects were not available.

Validity

Two approaches were taken to assess the validity of the two questionnaires in this population. Firstly, we examined cross sectional and longitudinal construct validity by correlating the quality of life scores with each other and with symptom scores and lung function parameters. To assess cross sectional construct validity we correlated absolute scores and to assess longitudinal construct validity we correlated change scores. Conceptually similar domains should be correlated stronger than domains which have less in common. The strength of the correlations was calculated using both Pearson's and Spearman's correlation coefficients. As no major differences were found, Pearson's correlations are shown.

In the second approach, common factor analysis with varimax rotation was employed to examine the relationships between the questions of the two health status measures. Factors were extracted using principal component analyses and selected when their own value (the variance explained by each factor) was greater than 1.

Responsiveness

Responsiveness was assessed in three different ways. Firstly, the ability of the questionnaires to detect within-subject changes was compared by means of paired t tests. Secondly, we compared the ability of the questionnaires to detect differences in change between the treatment groups. To test the hypothesis that the mean quality of life scores of the three groups after 12 weeks of treatment were equal, we used analysis of covariance (ANCOVA) with treatment group used as a factor and baseline score used as a covariate. Only when the ANCOVA indicated there were differences between the group means adjusted for baseline score were further ANCOVAs performed to determine which group means were significantly different from each other. Thirdly, we compared the two questionnaires with respect to the proportions of patients in the three groups reaching the minimally important improvement in quality of life.

Results

PATIENTS

One hundred and forty four patients both signed the informed consent form and met the inclusion/exclusion criteria. Due to logistical problems two of them did not complete the quality of life questionnaire at baseline. Nine of the remaining 142 patients failed to complete the second quality of life measurement. Seven patients dropped out because of an exacerbation of COPD (four on placebo, two on salmeterol/ipratropium bromide, and one on salmeterol) and two patients because of an adverse event (both on salmeterol). Hence, the responsiveness of 133 patients was analysed. Drop outs were compared with those who remained in the study with regard to baseline characteristics and quality of life and were not found to be significantly different. Table 1 shows that there were no statistically significant differences between the three groups in baseline characteristics including quality of life.

CLINICAL OUTCOMES

A description and discussion of the physiological and clinical results will be published elsewhere. To facilitate the interpretation of the changes in quality of life it suffices to note here that daytime symptom score, morning and evening PEF, FEV₁, specific airway conductance (sGaw), and night-time use of rescue drugs showed significantly more improvement in patients receiving salmeterol or salmeterol plus ipratropium bromide than in patients receiving placebo. There were statistically significant additional effects of the combination treatment over and above salmeterol treatment for evening PEF and FEV₁ % predicted.

Table 1 Baseline characteristics by treatment group

	Salmeterol+ ipratropium bromide (n=47)	Salmeterol+ placebo (n=47)	Placebo+ placebo (n=50)	p value*
Age (years)	63.0 (7.1)	65.4 (5.8)	63.2 (7.4)	0.165
Years since onset of COPD	11.2 (10.0)	12.4 (10.7)	13.8 (11.0)	0.511
Sex (% male)	87.2%	89.4%	86.0%	0.880
Pre-bronchodilator FEV ₁ (l)	1.4 (0.4)	1.3 (0.4)	1.3 (0.4)	0.649
FEV ₁ (% predicted)	44.7 (11.3)	44.0 (9.6)	42.0 (10.8)	0.447
FVC (l)	3.1 (0.8)	2.9 (0.7)	3.0 (0.7)	0.432
PEFR (l/s)	279 (89)	267 (75)	264 (81)	0.624
Reversibility (% predicted)	6.7 (4.6)	5.7 (4.2)	5.8 (3.1)	0.385
Using inhaled steroids (%)	74	81	76	0.743
On salmeterol before start of the run in (%)	57	57	51	0.791
SGRQ: Total	48.3 (13.9)	50.5 (16.6)	50.2 (15.8)	0.640
Symptoms	54.0 (19.9)	52.7 (25.5)	54.4 (22.6)	0.936
Activities	64.0 (14.8)	65.7 (17.5)	64.5 (17.2)	0.879
Impact	37.2 (16.6)	42.2 (18.6)	40.9 (18.1)	0.387
CRQ: Total	4.44 (0.81)	4.58 (1.03)	4.22 (1.05)	0.241
Dyspnoea	3.21 (0.99)	3.64 (1.19)	3.45 (1.14)	0.331
Fatigue	4.35 (1.06)	4.33 (1.16)	4.11 (1.19)	0.553
Emotions	4.87 (1.00)	5.00 (1.10)	4.45 (1.28)	0.056
Mastery	5.31 (1.22)	5.26 (1.27)	4.92 (1.40)	0.278

Data are presented as mean (SD) unless otherwise stated.

FEV₁ = forced expiratory volume in one second; FVC = forced vital capacity; PEFR = peak expiratory flow rate; SGRQ = St George's Respiratory Questionnaire; CRQ = Chronic Respiratory Disease Questionnaire.

*ANOVA on means of three groups and χ^2 for proportions.

QUESTIONNAIRE CONTENT

Despite the labels, the symptoms domain of the SGRQ and the dyspnoea domain of the CRQ do not cover the same concept. The symptoms domain refers to the frequency in which symptoms such as cough, shortness of breath, and sputum production occur, while the dyspnoea domain measures the extent to which certain activities cause shortness of breath. Conceptually, the dyspnoea domain of the CRQ is more closely related to the activity domain of the SGRQ which asks about activities that cause breathlessness and whether patients slow down or rest while doing certain activities. The CRQ has separate domains for fatigue and mastery, the latter referring to disease control or coping. The SGRQ contains a few questions about fatigue and only one or two questions about coping which are included in the impact domain. The impact of the SGRQ domain further assesses how the chest problems interfere with social life, household activities, paid employment, sports, etc. These items are not addressed in the CRQ. The CRQ contains more questions about emotions such as frustrations, restlessness, embarrassment, worries, depression, and tension than the SGRQ.

FEASIBILITY

Whereas the response options of the CRQ range from 1 to 7 for all questions, the response options of the SGRQ frequently change from 5, 4 or 3 point Likert scales to true/false response categories. The recall period of SGRQ varied from "the last four weeks" for the symptoms domain to "usually" or "these days" for the other domains, whereas the recall period of the CRQ remained four weeks for all questions. As a result, patients indicated that the SGRQ required more careful reading. Both questionnaires contain few negatively worded questions which use words such as "not", "rarely" or "never" and the length of the individual questions was comparable.

Table 2 presents an overview of missing values per question and per patient. Part 1 of the SGRQ contains two questions—"How long

Table 2 Missing values

	SGRQ	CRQ
% of missing values per question (range)	0–58% 0–10%*	0–5.3%
% of missing values per patient	2.0% 0.25%*	0.54%

*When questions 6 and 8 of Part 1 are excluded.

SGRQ = St George's Respiratory Questionnaire; CRQ = Chronic Respiratory Disease Questionnaire.

did the worst attack of chest trouble last?" (no. 6) and "If you have a wheeze, is it worse in the morning?" (no. 8)—which frequently do not apply to patients with COPD. As a result, the answers to these questions were missing in 58% and 27% of the cases, respectively. In the instruction manual of the SGRQ it is advised that these questions be omitted from the score calculation for patients who had not experienced attacks or wheezing. If these questions are not taken into account the number of missing values reduces considerably (table 2). Multiple responses to questions where a single response was required were not observed in any of the questionnaires.

Eliciting the patient's five most important activities is frequently reported as being the most difficult part of the CRQ. However, at baseline all but two patients were able to recall five activities that made them feel short of breath. One patient recalled four activities and one patient recalled only two activities. Of the 133 patients who completed the questionnaire after 12 weeks of treatment only seven patients changed one or more activities listed at baseline. Since this is not allowed, these responses were recoded to missing values.

RELIABILITY

The internal consistency of the CRQ was high with α coefficients ranging from 0.84 to 0.87. The internal consistency of the SGRQ was somewhat lower with α coefficients ranging from 0.76 to 0.77, but none of the α coefficients fell below 0.70 which is usually considered good enough for group comparisons.

Table 3 Cross section construct validity: Pearson's correlation coefficients

	SGRQ				CRQ				
	Symptoms	Activity	Impact	Total	Dyspnoea	Fatigue	Emotions	Mastery	Total
Symptoms									
Daytime score	0.308**	0.194*	0.332**	0.344**	-0.223**	-0.324**	-0.324**	-0.297**	-0.330**
Night-time score	0.138	0.299**	0.264**	0.296**	-0.212*	-0.330**	-0.282**	-0.171*	-0.299**
Lung function									
FEV ₁ (% pred)	0.004	-0.207*	-0.124	-0.144	0.071	0.061	0.023	0.033	0.053
PEFR am	-0.138	-0.252**	-0.084	-0.169*	0.151	0.123	0.097	0.089	0.136
PEFR pm	-0.172*	-0.224**	-0.094	-0.175*	0.124	0.109	0.086	0.085	0.120
SGRQ									
Symptoms					-0.352**	-0.528**	-0.518**	-0.475**	-0.562**
Activities					-0.476**	-0.577**	-0.414**	-0.365**	-0.537**
Impact					-0.367**	-0.616**	-0.633**	-0.590**	-0.664**
Total					-0.469**	-0.700**	-0.645**	-0.602**	-0.725**

SGRQ = St George's Respiratory Questionnaire; CRQ = Chronic Respiratory Diseases Questionnaire; FEV₁ = forced expiratory volume in one second; PEFR = peak expiratory flow rate.
*p <0.05; **p <0.01.

VALIDITY

Cross sectional construct validity

The cross sectional correlation matrix is shown in table 3. The SGRQ activity and total score correlated significantly with morning and evening PEFR and the activity score correlated significantly with FEV₁ % predicted. Nevertheless, the correlation between the questionnaires and lung function measures was weak. Both questionnaires correlated somewhat better with symptom scores. The total scores of both questionnaires correlated best with each other and with the domain scores. The fatigue domain of the CRQ correlated well with all domains of the SGRQ. Among the domains of the SGRQ the impact domain correlated best with the other domains of the CRQ, except dyspnoea. The weakest correlation between the two questionnaires was that between the dyspnoea domain of the CRQ and the symptoms domain of the SGRQ.

Longitudinal construct validity

The longitudinal correlation matrix is shown in table 4. Compared with the CRQ, the change in SGRQ scores showed better correlation with change in lung function. This applies especially to the SGRQ activity score. Compared with the cross sectional correlations, the longitudinal correlations between the quality of life scores and the symptom scores were worse. Again, the SGRQ activity score correlated best. Like the cross sectional correlations, the longitudinal correlations were highest between the total

scores of the two questionnaires. The CRQ fatigue domain correlated best with the SGRQ, the CRQ dyspnoea domain correlated best with the activity domain of the SGRQ, and the CRQ mastery domain correlated best with the SGRQ impact domain. Overall, the longitudinal correlations between the two questionnaires were lower than the cross sectional correlations.

Factor analysis

Based on the criterion of an eigenvalue of >1, the factor analysis on the data of the CRQ revealed three relatively independent factors, explaining almost 65% of the total variance in the original measure. The factor loadings after varimax rotation are shown in table 5. No factor loading fell below 0.5. The first factor extracted could be labelled "energy and coping". It showed positive correlation with three of the four fatigue items, two of the four mastery items, and three of the seven emotions items. The second factor appears to reflect a mental health dimension. It showed positive correlation with four of the seven emotions items and two of the four mastery items. The third factor could be labelled "shortness of breath" since it showed positive correlation with all dyspnoea items and one fatigue item.

Factor analysis of the SGRQ data yielded 18 different factors with an eigenvalue of >1, explaining almost 73% of the variance. These factors were very difficult to interpret. When the factor analysis was forced to extract three

Table 4 Longitudinal construct validity: Pearson's correlation coefficients

	SGRQ				CRQ				
	Symptoms	Activity	Impact	Total	Dyspnoea	Fatigue	Emotions	Mastery	Total
Symptoms									
Daytime score	0.167	0.234**	0.056	0.186*	-0.105	-0.133	-0.053	0.010	-0.083
Night-time score	0.014	0.103	0.001	0.041	-0.076	-0.097	-0.124	-0.032	-0.105
Lung function									
FEV ₁	-0.298*	-0.332**	+0.038	-0.195*	0.183*	0.258**	0.040	0.075	0.155
PEFR am	-0.111	-0.256**	-0.137	-0.224*	0.072	0.178*	0.050	0.080	0.107
PEFR pm	-0.009	-0.203*	-0.009	-0.126*	0.044	0.074	0.068	0.059	0.075
SGRQ									
Symptoms					-0.221*	-0.457**	-0.380**	-0.346**	-0.412**
Activities					-0.338**	-0.444**	-0.324**	-0.291**	-0.414**
Impact					-0.165	-0.303**	-0.361**	-0.379**	-0.361**
Total					-0.315**	-0.528**	-0.497**	-0.486**	-0.543**

SGRQ = St George's Respiratory Questionnaire; CRQ = Chronic Respiratory Diseases Questionnaire; FEV₁ = forced expiratory volume in one second; PEFR = peak expiratory flow rate.
*p <0.05; **p <0.01.

Table 5 Factor analysis with varimax rotation of CRQ

	Loading	Percentage of explained variance	
Factor 1			
Worn out or sluggish (f)	0.785	45.9	
Relaxed and free of tension (e)	0.765		
Low in energy (f)	0.742		
Happy, satisfied, pleased (e)	0.717		
Discouraged or down in the dumps (e)	0.714		
Complete control of breathing problems (m)	0.705		
How much energy (f)	0.647		
Confident and sure (m)	0.574		
Factor 2			
Fear or panic (m)	0.846		13.1
Upset or scared (m)	0.774		
Frustrated or impatient (e)	0.666		
Embarrassed by coughing (e)	0.621		
Restless, tense or uptight (e)	0.616		
Upset, worried or depressed (e)	0.609		
Factor 3			
Item 4 (d)	0.794	5.8	
Item 5 (d)	0.792		
Item 3 (d)	0.786		
Item 2 (d)	0.759		
Item 1 (d)	0.712		
How tired have you felt? (f)	0.571		

d = dyspnoea; f = fatigue; e = emotions; m = mastery.

factors only 28% of the variance was explained, but the three factors were more closely related to the original domains. The first factor showed positive correlation with 10 of the 16 activity items, 11 of the 26 impact items, and two of the eight symptoms items. We labelled it “physical activity” because many of the impact items were asking about exercise. The second factor showed positive correlation with seven of the eight symptoms items, seven of the 26 impact items (most of which were asking about coughing or chest trouble), and four of the 16 activity items. The factor could therefore be labelled “symptoms”. The third factor showed positive correlation with eight impact items and one activity item and was therefore labelled “impact”. Twenty three items had a factor loading below 0.4. When these items were excluded the resemblance between the three factors and the three original domains increased.

RESPONSIVENESS

The SGRQ symptom score and the CRQ fatigue and total scores showed statistically sig-

Table 6 Responsiveness to changes within subjects and differences in changes between salmeterol + ipratropium bromide (n = 45), salmeterol + placebo (n = 43), and placebo + placebo (n = 45)

	Change in quality of life			p value [†] (overall)
	Salmeterol+ ipratropium bromide	Salmeterol + placebo	Placebo+ placebo	
SGRQ				
Symptoms	-2.4 (11.0)	-0.01 (7.1)	+0.5 (9.2)	0.232
Activities	-8.1 (15.3)**	+1.4 (16.3)	+0.7 (20.0)	0.019
Impact	-2.6 (14.5)	+0.5 (8.1)	+1.2 (10.9)	0.227
CRQ				
Symptoms	-0.5 (12.1)	-0.5 (10.9)	-0.1 (13.3)	0.944
Dyspnoea	0.19 (0.6)*	-0.21 (0.8)	<0.01 (0.7)	0.036
Fatigue	0.21 (0.8)	-0.08 (0.9)	0.07 (1.0)	0.368
Emotions	0.25 (0.8)*	-0.13 (0.9)	<0.01 (0.9)	0.085
Mastery	0.17 (0.8)	-0.34 (0.9)*	0.04 (0.8)	0.009
	0.17 (0.8)	-0.20 (1.0)	<0.01 (0.7)	0.130

SGRQ = St George's Respiratory Questionnaire; CRQ = Chronic Respiratory Disease Questionnaire.

†Covariance analysis on change scores with treatment groups as a factor and baseline score as a covariate. Improvement is indicated by a minus for the SGRQ and a plus for the CRQ.

Within-patient changes: *p<0.05; **p<0.001.

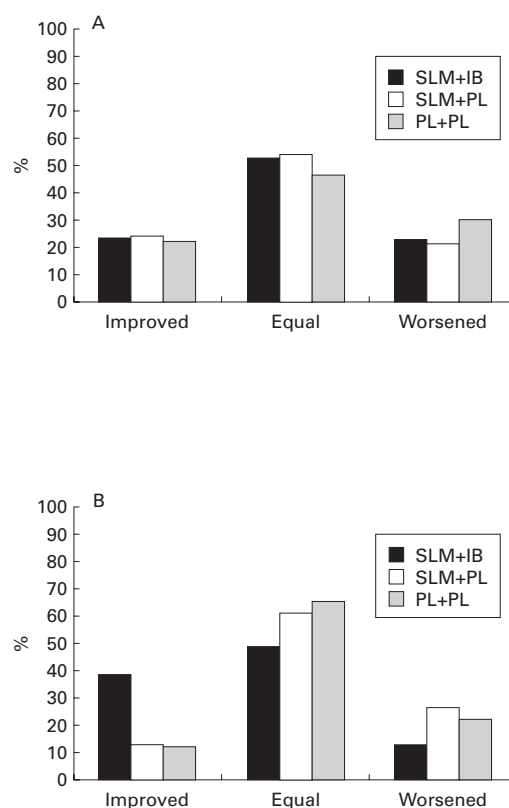


Figure 1 Percentage of patients showing a clinically relevant change in quality of life according to (A) the SGRQ ($p = 0.743$, Kruskal-Wallis test) and (B) the CRQ ($p = 0.005$, Kruskal-Wallis test). The threshold for a minimum clinically relevant change is 4 for the SGRQ and 0.5 for the CRQ. SLM = salmeterol; IB = ipratropium; PL = placebo.

nificant within-patient improvements in patients receiving salmeterol plus ipratropium bromide. Except for a deterioration on the CRQ emotions domain in the salmeterol group, the other domains did not show statistically significant within-patient changes (table 6). Table 6 also shows that the changes in the symptoms domain of the SGRQ and in the total and emotions domains of the CRQ differed significantly between the three groups. Further pairwise analyses revealed that the symptoms score showed a significantly greater improvement on treatment with salmeterol and ipratropium bromide than on salmeterol alone (mean difference 9.4; 95% CI 2.7 to 16.1) or placebo (mean difference 8.7; 95% CI 1.3 to 16.2), neither of which showed a significant change. The average differences exceeded the minimally important difference of 4%. A significantly greater improvement on combination treatment than on salmeterol treatment was also found for the overall CRQ score (mean difference 0.4; 95% CI 0.11 to 0.69) and the CRQ emotions domain (mean difference 0.51; 95% CI 0.17 to 0.83). These differences were due to a worsening of scores with salmeterol and an improvement of scores with salmeterol plus ipratropium bromide. However, the changes on the CRQ remained below what is considered the minimally clinically important difference. Neither the SGRQ nor the CRQ showed any significant differences between salmeterol and placebo.

Figure 1 shows the proportion of patients reaching at least a minimum clinically relevant change in the overall score of the two quality of life questionnaires. Almost 40% of the patients on combination treatment reached a clinically relevant improvement in quality of life as measured by the CRQ compared with 13% and 12% of the patients on salmeterol or placebo, respectively. These proportions were significantly different. Of the patients on combination treatment, 23% reached the minimally important improvement on the SGRQ compared with 24% of the patients on salmeterol and 23% of the patients on placebo. These proportions were not significantly different. Of all the domain scores, the SGRQ symptoms and the CRQ emotions domains showed a significantly higher proportion of patients who reached a clinically relevant improvement on combination treatment (symptoms: 53% on salmeterol and ipratropium bromide, 26% on salmeterol, 33% on placebo; Kruskal-Wallis test, $p = 0.024$; emotions: 32% on salmeterol and ipratropium bromide, 9% on salmeterol, 22% on placebo; Kruskal-Wallis test, $p = 0.006$).

Discussion

The results of this direct comparison of the self-administered version of the CRQ and the SGRQ in a clinical trial have confirmed their feasibility for use in patients with moderate to severe COPD. Completion rates were good for both questionnaires, with that for the self-administered version of the CRQ being even better than the completion rate for the interviewer-administered version of the CRQ reported by Harper *et al*,¹⁴ probably because our measurements were part of a clinical trial whereas their measurements were done in routine clinical practice.

Although the internal consistency was somewhat better for the CRQ, both questionnaires showed good internal consistency with α coefficients above 0.75. We could not confirm the findings of Wijkstra *et al*²⁵ and Guell *et al*²² who reported considerably lower internal consistency values for the CRQ dyspnoea domain than for the other CRQ domains of 0.53 and 0.51, respectively, compared with our value of 0.86. As all psychometric properties may depend on patient characteristics, it is important to point out that the studies by Wijkstra *et al* and Guell *et al* were conducted in patients with more severe illness than in our trial.

The validity of the two quality of life questionnaires is supported by the correlations between their total scores and their domain scores. Almost all cross sectional and longitudinal correlations are highly statistically significant ($p < 0.01$). Cross sectionally, the scores of the two questionnaires were moderately to highly correlated (coefficients of 0.35–0.72). The correlations between change scores were lower, although the correlation between the changes in total scores of the two questionnaires was still 0.54. The overall patterns of correlations were consistent with expectations. The CRQ domain fatigue correlated best with the domains of the SGRQ. This fatigue domain contains general items about tiredness and

energy which one would expect to be associated with symptoms, level of activity, and the impact of the disease on daily life. The impact domain of the SGRQ correlated best with the emotions and mastery domains of the CRQ. The impact domain of the SGRQ is a very broad domain covering a wide variety of concepts. Among the weakest correlations found was the correlation between the dyspnoea domain of the CRQ and the symptoms domain of the SGRQ. The CRQ dyspnoea domain correlates best with the activity domain of the SGRQ. This observation confirms our qualitative comparison of domain content. Despite the similarity of the domain labels, the dyspnoea domain of the CRQ measures the extent to which certain activities cause breathlessness, whereas the symptoms domain of the SGRQ measures the frequency of symptom occurrence.

When looking at the correlations with lung function and symptom scores, the SGRQ showed somewhat better validity than the CRQ. Its correlation with *change* in lung function was better, as was its ability to discriminate between patients with more or less severely impaired lung function. Nevertheless, the correlation between quality of life and lung function and the ability of both questionnaires to discriminate between patients with more or less severely impaired lung function were rather limited. This is not surprising given the limited variation in lung function. Only patients with a pre-bronchodilator FEV₁ between 40% and 65% were eligible. The small variation in symptom scores might also explain the relatively low longitudinal correlation between quality of life and symptom scores.

Although we realise that it is difficult to confirm factor structures in different populations, especially when the ratio of subjects to items is not very large, we tried to compare the robustness of the factor structure of the two questionnaires. The factor analysis on data of the CRQ proposed a three dimensional model instead of the current four domains. In particular, items of the mastery domain were distributed over the first two factors. The construction of the CRQ was not based on psychometric approaches like factor analysis but the items were selected using the “impact method”.²⁶ This method selects items that are most frequently perceived as important by the patients. The “impact method” and the psychometric approach represent two different schools of thought. However, in reanalysing the item pool generated for the construction of the Asthma Quality of Life Questionnaire, Juniper *et al* showed that the two methods generated similar instruments.²⁶ This was partly confirmed by our analysis of the CRQ. Little is published about the item generation of the SGRQ, but factor analysis did not seem very suitable. It presupposed an 18-factor model that could not be interpreted. When a three-factor model was designed the proportion of the variance that was explained by the model fell from 73% to 28%. Almost 50% of the items of the SGRQ had a factor loading below 0.4, indicating a lack of homogeneity.

Our study is one of the first clinical trials that enables a direct comparison of the response of the SGRQ and the CRQ to changes in health resulting from bronchodilator therapy. The symptoms domain was the most responsive of the SGRQ. This domain not only showed a statistically significant improvement of more than 8% in patients receiving combination treatment, but the improvement also exceeded the 4% threshold for clinical relevance. In an earlier study Jones and Lasserson found a change of about 8% in patients who considered their medication to be moderately efficacious.¹⁷ Furthermore, the symptoms domain was the only domain of the SGRQ which showed that a significantly higher proportion of patients reached a clinically relevant improvement on combination treatment than on the other treatments.

The total score of the CRQ and its fatigue and emotions domains were also able to detect changes in health. The total score and the emotions domain showed that the changes in health were statistically significantly different between the combination treatment and treatment with salmeterol alone. Although the changes in health were statistically significant, they did not cross the threshold value of minimally clinically important change. Nevertheless, both the CRQ total and emotions scores were able to show that a significantly higher proportion of the patients receiving combination treatment reached a minimally important improvement.

The minimally important differences of 0.5 for the CRQ and 4% for the SGRQ were established by the designers of the questionnaires. Such differences may vary according to type of patients, study setting, and interventions.²⁷ For the CRQ the minimally important difference was established in patients with COPD undergoing inpatient rehabilitation treatment using two different patient oriented approaches, one within-patient and one between-patient.²¹ For the SGRQ the minimally important difference was also established using two different approaches, one physician based² and one patient based¹⁶ in patients receiving nedocromil sodium. Little is known about cross cultural differences in estimates of the minimally important difference. Dutch studies of the CRQ and SGRQ aimed at confirming the values of minimally important difference have not yet been conducted.

Because the ratio between the minimally important difference and the standard deviation of the scores is larger, the CRQ generally requires a smaller sample size than the SGRQ to detect the minimally important difference. The power of this trial to detect a difference of 0.5 between two groups in the CRQ total score was 89%, while the power to detect a 4% difference in the SGRQ total score was 63%.

One of the limitations of this study is its relatively short duration. A three month treatment period might be too short for patients with COPD to recognise fully the impact of an effective treatment. Improvement in breathlessness or exercise tolerance might not immediately attract the patients' attention because

they have adopted a life style in which they avoid strenuous daily and leisure activities.

Another limitation is that we have selected relatively stable patients, most of whom were already well treated with bronchodilator and anti-inflammatory medication. As a result, most changes in quality of life were rather small. More than 75% of the patients were using inhaled corticosteroids before the start of the trial. More than 50% of the patients were already using salmeterol but they were required to stop this during the two week run in period. We believe that small improvements due to new medications are typical for the current medical practice with respect to COPD in The Netherlands where the condition of many patients is already well controlled with existing medications. The small changes in health were not due to ceiling effects since the percentage of patients with a baseline score above 80% of the best possible score was only 1.4% on the SGRQ and 10.9% on the CRQ. Hence, there was enough room for improvement.

In conclusion, this analysis of psychometric properties does not clearly favour one instrument above the other. The completion rate of the self-administered version of the CRQ was good and comparable to the SGRQ. The internal consistency was somewhat better for the CRQ but the longitudinal validity and discriminative validity were somewhat better for the SGRQ. The homogeneity of the domains of the SGRQ can be improved. The symptoms domain was the most responsive of the SGRQ and the total and emotions scores were the most responsive scores of the CRQ. Since the psychometric performances of the CRQ and the SGRQ are largely comparable, the choice between these two instruments can be based on other considerations such as the size of the required sample which is smaller for the CRQ than for the SGRQ, or the availability of reference values which favours the more widely used SGRQ.

This study was financially supported by Glaxo Wellcome.

- 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.
- Jones PW, Quirk FH, Baveystock CM. The St George's Respiratory Questionnaire. *Respir Med* 1991;85(Suppl B):25-31.
- 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.
- Guyatt GH, Townsend M, Keller J, et al. Measuring functional status in chronic lung disease: conclusions from a randomized control trial. *Respir Med* 1989;83:293-7.
- Jaeschke R, Guyatt GH, Willan A, et al. Effect of increasing doses of beta-agonists on spirometric parameters, exercise capacity, and quality of life in patients with chronic airflow limitation. *Thorax* 1994;49:479-84.
- Jones PW, Bosh TKL. Quality of life changes in COPD patients treated with salmeterol. *Am J Respir Crit Care Med* 1997;155:1283-9.
- Okubadejo AA, O'Shea L, Jones PW, et al. Home assessment of activities of daily living in patients with severe chronic obstructive pulmonary disease on long-term oxygen therapy. *Eur Respir J* 1997;10:1572-5.
- Andersson A, Strom K, Brodin H, et al. Domiciliary liquid oxygen versus concentrator treatment in chronic hypoxaemia: a cost-utility analysis. *Eur Respir J* 1998;12:1284-9.
- Goldstein RS, Gort EH, Stubbing D, et al. Randomised controlled trial of respiratory rehabilitation. *Lancet* 1994;344:1394-7.
- Wijkstra PJ, Ten Vergert EM, Van Altena R, et al. Long term benefits of rehabilitation at home on quality of life and

- exercise tolerance in patients with chronic obstructive pulmonary disease. *Thorax* 1995;**50**:824–8.
- 11 Lacasse Y, Wond E, Guyatt GH, *et al.* Meta-analysis of respiratory rehabilitation in chronic obstructive pulmonary disease. *Lancet* 1996;**348**:1115–9.
 - 12 Ketelaars CA, Abu-Saad HH, Schlösser MA, *et al.* Long-term outcome of pulmonary rehabilitation in patients with COPD. *Chest* 1997;**112**:363–9.
 - 13 Hajiro R, Nishimura K, Tsukino M, *et al.* Comparison of discriminative properties among disease-specific questionnaires for measuring health-related quality of life in patients with chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 1998;**157**:785–90.
 - 14 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.
 - 15 Guyatt GH, Townsend M, Keller JL, *et al.* Should study subjects see their previous responses: data from a randomized control trial. *J Clin Epidemiol* 1989;**42**:913–20.
 - 16 Quirk FH, Baveystock CM, Wilson RC, *et al.* Influence of demographic and disease related factors on the degree of distress associated with symptoms and restriction on daily living due to asthma in six countries. *Eur Respir J* 1991;**4**:167–71.
 - 17 Jones PW, Lasserson D. Relationship between change in St George's Respiratory Questionnaire score and patients' perception of treatment efficacy after one year of therapy with nedocromil sodium. *Am Rev Respir Crit Care Med* 1994;**149**:A211.
 - 18 Ferrer M, Alonso J, Prieto L, *et al.* Validity and reliability of the St George's Respiratory Questionnaire after adaptation to a different language and culture: the Spanish example. *Eur Respir J* 1996;**9**:1160–6.
 - 19 Bouchet C, Guillemin F, Hoang Thi TH, *et al.* Validation of the St George's Respiratory Questionnaire for measuring quality of life in patients with chronic obstructive pulmonary disease. *Rev Mal Respir* 1996;**13**:43–6.
 - 20 Jaeschke R, Singer J, Guyatt GH. Measurement of health status: ascertaining the minimal clinically important difference. *Control Clin Trials* 1989;**10**:407–15.
 - 21 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.
 - 22 Guell R, Casan P, Sanges M, *et al.* The Spanish translation and evaluation of a quality of life questionnaire in patients with chronic obstructive pulmonary disease. *Arch Bronconeumol* 1995;**31**:202–10.
 - 23 Streiner DL, Norman GR. *Health measurement scales. A practical guide to their development and use.* New York: Oxford University Press, 1989.
 - 24 Siafakas NM, Vermeire P, Pride NB, *et al.*, on behalf of the Task Force Optimal Assessment and Management of Chronic Obstructive Pulmonary Disease (COPD). A consensus statement of the European Respiratory Society (ERS). *Eur Respir J* 1995;**8**:1398–420.
 - 25 Wijkstra PJ, Ten Vergert EM, Van Altna R, *et al.* Reliability and validity of the chronic respiratory questionnaire. *Thorax* 1994;**49**:465–7.
 - 26 Juniper EF, Guyatt GH, Streiner DL, *et al.* Clinical impact versus factor analysis for quality of life questionnaire constructing. *J Clin Epidemiol* 1997;**50**:233–8.
 - 27 Wright JG. The minimal important difference: who's to say what is important? *J Clin Epidemiol* 1996;**49**:1221–2.