



# Self-administration and standardisation of the chronic respiratory questionnaire: a randomised trial in three German-speaking countries

Milo A. Puhan<sup>a,\*</sup>, Michaela Behnke<sup>b</sup>, Marco Laschke<sup>c</sup>, Alfred Lichtenschopf<sup>d</sup>, Otto Brändli<sup>e</sup>, Gordon H. Guyatt<sup>f</sup>, Holger J. Schünemann<sup>f,g</sup>

<sup>a</sup>Horten Centre, University of Zurich, Postfach Nord, CH-8091 Zurich, Switzerland

<sup>b</sup>Pulmoresearch Institute, Grosshansdorf, Germany

<sup>c</sup>Klinik Barmelweid, Barmelweid, Switzerland

<sup>d</sup>Rehabilitationszentrum, Weyer/Enns, Austria

<sup>e</sup>Zuercher Hoehenklinik Wald, Faltigberg-Wald, Switzerland

<sup>f</sup>Department of Clinical Epidemiology and Biostatistics, McMaster University, Hamilton, Ont., Canada

<sup>g</sup>Departments of Medicine and of Social and Preventive Medicine, University at Buffalo, Buffalo, NY, USA

Received 22 May 2003; accepted 7 October 2003

## KEYWORDS

COPD;  
Health related quality of life;  
Respiratory questionnaires;  
Standardisation;  
Self-administration

**Summary** The chronic respiratory questionnaire (CRQ) has demonstrated excellent measurement properties in patients with chronic obstructive pulmonary disease (COPD), but in its original form it is limited by the requirement for interviewer-administration and the individualised dyspnoea questions. The objective of this randomised trial was to examine the evaluative properties of the interviewer and self-administered German CRQ as well as of a standardised CRQ dyspnoea domain. In a multinational trial we randomly allocated 71 patients with COPD to complete the interviewer administered CRQ (CRQ-IA) or the self-administered CRQ (CRQ-SA) and other validation measures at the beginning and end of a respiratory rehabilitation program. We assessed and compared responsiveness and longitudinal validity of the CRQ. The change scores of all CRQ domains were above the minimal clinically important difference of 0.5. Responsiveness of the fatigue domain was higher for the CRQ-SA compared to CRQ-IA ( $P = 0.02$ ), but there was no difference in responsiveness on the other domains. Compared to the standardised dyspnoea domain the individualised dyspnoea questions tended to show greater responsiveness for both the CRQ-IA ( $P = 0.07$ ) and CRQ-SA ( $P = 0.10$ ). We found better longitudinal validity for the CRQ-SA represented by larger correlations between CRQ change scores and those of other validation instruments. Taken these results into consideration, researchers in COPD, in particular those in German-language countries can utilise any one of four CRQ formats that have proved both valid and responsive.

© 2003 Elsevier Ltd. All rights reserved.

**Abbreviations:** COPD, chronic obstructive pulmonary disease; CRQ, chronic respiratory questionnaire; CRQ-IA, chronic respiratory questionnaire interviewer-administered; CRQ-SA, chronic respiratory questionnaire self-administered; FT, feeling thermometer; HRQL, health related quality of life; 6MWT, Six minutes walk test; *sd*, standard deviation

\*Corresponding author. Tel.: +41-1-2558709; fax: +41-1-2559720.

E-mail address: milo.puhan@evimed.ch (M.A. Puhan).

## Introduction

Projections suggest that COPD is the only cause of death from chronic disease that will increase worldwide until 2020.<sup>1</sup> In Switzerland, Germany and Austria the smoking prevalence is above 30%<sup>2</sup> and the decline in smoking prevalence seen in North-America and Great-Britain is absent.<sup>3-5</sup> In 1998, there were about 3–5 million patients with COPD in German-speaking countries.<sup>6</sup>

Health-related quality of life (HRQL) in patients with COPD is important in determining response to treatment.<sup>7,8</sup> Although there are German translations of COPD specific quality of life instruments, these translations have not been properly validated in clinical trials.

Two widely used COPD-specific questionnaires are available in English: the interviewer-administered "Chronic Respiratory Questionnaire" (CRQ)<sup>9</sup> and the self-administered "St George's Respiratory Questionnaire".<sup>10</sup> The CRQ is a valid and reliable instrument.<sup>11,12</sup> Compared to the St. George's Respiratory Questionnaire and generic questionnaires, the CRQ is more responsive to small but important changes of HRQL.<sup>13,14</sup> Another advantage of the CRQ is the simple analysis and interpretability supported by a significant body of evidence.<sup>15-17</sup>

Despite its established psychometric properties, investigators have criticised the CRQ because it is interviewer-administered and the individualised dyspnoea domain makes comparisons between patients with different limitations in their daily activities difficult.<sup>11,18</sup> Although a self-administered<sup>19</sup> and a standardised<sup>20</sup> version of the CRQ exist in English, further validation and description of the measurement properties associated with self-administration and standardisation is important before widespread use of these simplified administration formats.

The objective of this study was to assess and compare the measurement properties of the self-administered and the interviewer-administered version of the German CRQ including standardised dyspnoea questions.

## Methods

### Patients

We included patients with COPD defined as FEV<sub>1</sub>/FVC < 70% predicted and postbronchodilator FEV<sub>1</sub> < 80% predicted, German as the first or daily language, age > 40 years and ability to complete

the CRQ within one session. We excluded patients with inability to read or write, with cognitive difficulties (due to dementia, substance abuse or neurological disorders), with cancer or lung diseases other than COPD. We recruited patients consecutively from August to November 2002. All patients followed a multidisciplinary program with emphasis on physical exercise. Additional elements were educational sessions, psychological support and relaxation therapies.

### Study design

We conducted a multinational clinical trial to concurrently validate a German translation of the interviewer-administered CRQ (CRQ-IA) and the self-administered CRQ (CRQ-SA) including a standardised dyspnoea domain. In order to allow comparisons between these administration formats we randomised patients to complete either the CRQ-IA or the CRQ-SA. Within each group (CRQ-IA or CRQ-SA) patients completed both the individualised and standardised dyspnoea questions. To eliminate order effects we also randomised in a 2 × 2 factorial design the order of receiving the individualised or standardised dyspnoea questions. We produced a computer-generated randomisation list with blocks of four per centre. To ensure concealment of allocation, the centre's research leaders who were not aware of block size called the study coordinator (MP) to obtain group assignment for each participant. We assessed all patients within 3 days after enrolment (baseline) and after at least two weeks of intense respiratory rehabilitation (follow-up) in any of the programs (total duration of rehabilitation 2–3 weeks). We recruited patients from four rehabilitation centres (three inpatient and one outpatient programme) in Switzerland (Zuercher Hoehenklinik Wald and Klinik Barmelweid, Barmelweid), Germany (Pulmoresearch Institute, Grosshansdorf) and Austria (Rehabilitationszentrum, Weyer/Enns). All interviewers completed standardised training session on the use and administration of the CRQ from one of the investigators. All local ethics committees approved the study protocol and all patients provided informed consent prior to participation in the study.

### The German version of the chronic respiratory questionnaire

All questions of the German interviewer and self-administered version are identical and correspond to those of the English versions.<sup>9,20</sup> The CRQ

includes questions across the four domains of fatigue, emotional function, mastery and dyspnoea. Questions covering the domains of fatigue, emotions and mastery are standardised and patients answer each question on a seven-point scale to express the degree of disability from 1 (maximum impairment) to 7 (no impairment). For the individualised dyspnoea domain, respondents select activities from a list of 26 items that cause the greatest degree of shortness of breath in daily life or to list important activities not on the list. Subsequently, patients select the five activities of greatest importance to them and indicate the breathlessness associated with these activities. The standardised dyspnoea domain comprises five questions concerning activities that cause shortness of breath in most patients with COPD.<sup>20</sup> These five activities are: (a) taking care of your basic needs (bathing, showering, eating or dressing), (b) walking, (c) feeling emotional such as angry or upset, (d) performing chores (such as housework, shopping, groceries), (e) participating in social activities. We used the informed version of the CRQ that allows patients at follow-up to see their prior responses.<sup>21</sup>

For translation and cultural adaptation of the CRQ in the German-speaking countries, we used a standard approach for translation, back translation, pilot testing and evaluation of instrument test-retest reliability and internal consistency reliability.<sup>22</sup> We will describe these and other details of the development process as well as the discriminative properties of the CRQ versions in a separate report.

### Validation instruments to assess longitudinal validity of the CRQ

We used the self-administered SF-36 Health Survey, a generic instrument for assessment of the health-related quality of life of patients. The SF-36 assesses 8 subscales of health-related quality of life. The SF-36 has been translated into German and validated as part of the IQOLA project.<sup>23</sup> Other investigators used the SF-36 in trials with COPD patients participating in respiratory rehabilitation.<sup>24</sup>

The FT is an anchor-based visual analogue scale from 0 to 100 where 0 (dead) represents the worst and 100 (full health) the best health state. Accumulating evidence suggests that the FT works well as an evaluative instrument in various groups, including patients with chronic obstructive pulmonary disease (COPD).<sup>25</sup>

We used the six-minute walking test (6MWT) to assess functional exercise capacity. It measures

the distance a patient can walk within 6 min. In addition, we used a modified Borg scale in German<sup>26</sup> to assess the intensity of perceived dyspnoea at the end of the 6MWT. The Borg scale consisted of a vertical line labelled 0–10 and with verbal descriptors. Zero represented “no dyspnoea at all” and 10 “maximal dyspnoea”.

### Statistical analysis

We calculated the mean score for each CRQ domain by summing the scores for each domain question and dividing it by the number of scored questions. We used parametric tests because the distributions on the seven points Likert-type scale were normal.

We calculated the mean change between baseline and follow-up domain scores and the corresponding 95% confidence intervals. We used paired *t*-tests to determine if baseline and follow-up scores differed significantly. To compare the responsiveness of the CRQ-IA and CRQ-SA we used independent *t*-tests and, finally, we used paired *t*-tests to compare the individualised and standardised dyspnoea domain.

To assess longitudinal construct validity, we calculated Pearson's correlation coefficients between change scores on the CRQ domains and the FT, SF-36, 6MWT and Borg scale.

We were interested in comparing the number of missing items between the different formats of administration. Our main interest was in the comparison of missing or not applicable items resulting from standardisation of the dyspnoea domain.

For calculating the required sample size, we used the following formula for paired observations:  $n = [(Z\alpha + Z\beta)^* SD/\Delta]^2$ .<sup>27</sup> where *n* represents the number of patients required per group, *Z* $\alpha$  and *Z* $\beta$  the terms to set the level of significance and the power, *SD* the standard deviation of the instrument and  $\Delta$  the desired difference between baseline and follow-up. The minimal sample size required per group was<sup>27</sup> to show effects for all four dimensions of 0.5 (the minimal important difference that patients or physicians judge as important<sup>15</sup>) points assuming a standard deviation of 0.8,  $\alpha = 0.05$  and  $\beta = 0.10$ . With an estimated dropout rate of 20%, the required sample size increased to 32 per group.

All statistical analyses were performed with SPSS for Windows version 10.0 (SPSS Inc, Chicago, IL, USA).

## Results

We recruited patients from the Zuercher Hoehenklinik Wald ( $n = 20$ ), the Klinik Barmelweid ( $n = 20$ ), the Rehabilitationszentrum Weyer/Enns ( $n = 12$ ) and the Pulmoresearch Institute Hamburg ( $n = 28$ ). We randomly assigned them to either the CRQ-IA group ( $n = 40$ ) or CRQ-SA group ( $n = 40$ ). Nine patients did not complete the study for the following reasons: five withdrew for non-specified reasons, two did not complete the rehabilitation program and upon review two patients did not meet the a priori inclusion criteria. Table 1 shows that patients randomised to the CRQ-IA or CRQ-SA had similar baseline characteristics.

In Table 2, we present the baseline and follow-up scores of the CRQ domains by randomisation group. Compared to the CRQ-IA, baseline scores on the CRQ-SA were consistently lower for each of the domains. However, only for the individualised dyspnoea domain was the difference significant ( $P = 0.02$ ). The differences between the CRQ-IA and CRQ-SA in the follow up scores were small and not statistically significant.

Both the interviewer and self-administered administration demonstrated statistically and clinically significant improvement over the course of rehabilitation (Table 3). The lower limits of the confidence intervals were above the minimum important difference of 0.5 for all domains. The

**Table 1** Demographic information for patients randomised to the CRQ-IA or CRQ-SA.

Patient characteristic	Interviewer administration ( $n = 38$ )	Self-administration ( $n = 33$ )
Male (%)	24 (63.2)	23 (69.7)
Female (%)	14 (36.8)	10 (30.3)
Age*	67.4 (8.7)	67.7 (8.3)
FEV <sub>1</sub> in % predicted*	45.1 (15.6)	45.1 (15.5)
FEV <sub>1</sub> /FVC in % predicted*	48.5 (13.3)	49.9 (10.5)
Pack years*	44.9 (26.1)	46.8 (27.6)
Six minute walking test at baseline*	359 (111)	361 (134)
Social status		
Living alone (%)	8 (21.1)	12 (36.4)
Working (%)	6 (15.8)	5 (15.2)
Not working (%)	10 (26.3)	5 (15.2)
Retired (%)	22 (57.9)	23 (69.7)
Duration of rehabilitation (days)*	16.5 (3.0)	16.9 (3.8)

\*Mean (standard deviation).

**Table 2** Baseline and follow-up scores of the interviewer and self-administered formats of the CRQ.

Domain		CRQ-IA ( $n = 38$ )	CRQ-SA ( $n = 33$ )	$\Delta$ (95% confidence interval, $P$ value*)
Dyspnoea individualised	Baseline	3.17 $\pm$ 0.66	2.72 $\pm$ 0.91	0.45 (0.08; 0.83; $P = 0.02$ )
	Follow-up	4.19 $\pm$ 0.93	3.97 $\pm$ 1.37	0.22 (-0.32; 0.77; $P = 0.41$ )
Dyspnoea standardised	Baseline	4.00 $\pm$ 1.01	3.55 $\pm$ 1.33	0.45 (-0.11; 1.01; $P = 0.11$ )
	Follow-up	4.78 $\pm$ 0.95	4.59 $\pm$ 1.31	0.19 (-0.34; 0.73; $P = 0.48$ )
Fatigue	Baseline	4.19 $\pm$ 1.20	3.81 $\pm$ 1.13	0.38 (-0.17; 0.94; $P = 0.17$ )
	Follow-up	5.02 $\pm$ 0.97	5.20 $\pm$ 1.03	-0.17 (-0.64; 0.30; $P = 0.47$ )
Emotional function	Baseline	4.38 $\pm$ 0.99	4.12 $\pm$ 1.22	0.26 (-0.26; 0.79; $P = 0.32$ )
	Follow-up	5.41 $\pm$ 0.84	5.24 $\pm$ 0.95	0.17 (-0.25; 0.60; $P = 0.42$ )
Mastery	Baseline	4.37 $\pm$ 1.24	4.25 $\pm$ 1.42	0.12 (-0.50; 0.76; $P = 0.69$ )
	Follow-up	5.26 $\pm$ 1.03	5.46 $\pm$ 0.89	-0.21 (-0.66; 0.25; $P = 0.38$ )

CRQ domain scores with values from 1 (largest impairment) to 7 (no impairment); CRQ-IA = Interviewer administered German chronic respiratory questionnaire; CRQ-SA = self-administered German chronic respiratory questionnaire; Values are mean  $\pm$  standard deviation;  $\Delta$  = difference between the CRQ-IA and CRQ-SA.

\* $P$ -values for independent  $t$ -tests to compare the CRQ-IA and CRQ-SA.

**Table 3** Change scores of the interviewer and self-administered format of the CRQ.

Domain	CRQ-IA (n = 38)	P value*	CRQ-SA (n = 33)	P value*	P value†
Individualised dyspnoea	1.02 (0.76–1.28)	<0.001	1.25 (0.86–1.64)	<0.001	0.32
Standardised dyspnoea	0.79 (0.52–1.06)	<0.001	1.04 (0.63–1.46)	<0.001	0.28
Fatigue	0.83 (0.55–1.12)	<0.001	1.39 (1.03–1.75)	<0.001	0.02
Emotional function	1.03 (0.78–1.28)	<0.001	1.12 (0.73–1.51)	<0.001	0.73
Mastery	0.88 (0.61–1.16)	<0.001	1.22 (0.77–1.66)	<0.001	0.21

CRQ domain scores with values from 1 (largest impairment) to 7 (no impairment); CRQ-IA = interviewer administered German chronic respiratory questionnaire; CRQ-SA = self-administered German chronic respiratory questionnaire; Values are mean (95% confidence interval).

\*P-values for paired *t*-tests to compare baseline and follow up scores of the CRQ-IA and CRQ-SA.

†P-values for independent *t*-tests to compare the change scores of the CRQ-IA and CRQ-SA.

distribution of the change scores was wider for the CRQ-SA compared to the CRQ-IA. When we compared the change scores of the CRQ-IA and CRQ-SA, we found a significant difference between the fatigue domains ( $P = 0.02$ ). The change scores of the dyspnoea, emotional function and mastery domains did not differ significantly. For both the CRQ-IA and CRQ-SA responsiveness tended to be higher for the individualised dyspnoea questions compared to the standardised dyspnoea questions ( $P = 0.07$  for the CRQ-IA and  $P = 0.10$  for the CRQ-SA).

We show the longitudinal validity represented by the correlations between CRQ change scores and change scores of the FT, SF-36, 6MWT and Borg scale in Tables 4 and 5. Most of the correlation coefficients for the change scores were higher in the groups randomised to the CRQ-SA.

Time required completing the CRQ-IA and the CRQ-SA with individualised dyspnoea questions was for both administration formats 16 min at baseline and 10 min at follow-up. Administration time for CRQ-IA and the CRQ-SA with standardised dyspnoea questions was for both formats 8 min at baseline and 8 (CRQ-IA) and 7 (CRQ-SA) at follow-up. Thus, patients needed less time to complete the standardised CRQ-SA while the mode of administration (self versus interviewer) did not affect the time required for completion.

There were more unanswered items in the standardised than in the individualised dyspnoea domain. 90% of the respondents completed five, 5% four and 5% three individual dyspnoea questions at baseline. For the standardised questions, 66% of patients completed five, 26% four and 8% three dyspnoea questions. At follow-up, 66% of the respondents completed five, 11% four and 18% three individual dyspnoea questions. 50% of the patients completed five, 18% four and 26% three standardised dyspnoea questions at follow-up.

Patients frequently responded “not applicable” to those questions that included activities, which patients did not perform during the rehabilitation program. All patients completed all items of the fatigue, emotion and mastery domain at baseline. At follow up, 100% of the patients completed all items of the mastery domain, 97% all items of the fatigue domain and 92% (CRQ-IA) or 97% (CRQ-SA), respectively, all items of the emotion domain.

## Discussion

We validated four formats of the German CRQ concurrently in a randomised trial. All four formats were highly responsive to HRQL changes resulting from respiratory rehabilitation. Longitudinal validity was better for the CRQ-SA.

In general, baseline scores were lower for the CRQ-SA (Table 2). Although lower baseline scores could result in larger improvements, we did observe higher responsiveness only for the fatigue domain of the CRQ-SA compared to the CRQ-IA (Table 3). We found differences between the individualised and standardised dyspnoea domains for both the CRQ-IA and CRQ-SA that failed to reach conventional levels of statistical significance. A likely explanation for increased responsiveness is that the individualised dyspnoea questions focus on important activities and patients may perceive larger benefits from effective interventions such as respiratory rehabilitation compared to standardised questions. Another explanation could be that baseline scores of the individualised questions were lower for both the CRQ-IA and CRQ-SA because patients chose activities causing the most severe dyspnoea.

The difference in responsiveness between the standardised and the individualised dyspnoea

**Table 4** Longitudinal validity for the individualised and standardised dyspnoea domains. Correlations for change scores\*.

Instrument and domain	CRQ-IA dyspnoea domains		CRQ-SA dyspnoea domains	
	Individualised*	Standardised*	Individualised*	Standardised*
Feeling thermometer	0.38 (0.23;0.53)	0.21 <sup>†</sup> (0.05;0.37)	0.42 (0.27;0.57)	0.53 <sup>†</sup> (0.39;0.67)
SF-36-General health perception index	0.16 (0.00;0.32)	-0.02 <sup>†</sup> (-0.19;0.15)	0.41 (0.26;0.56)	0.46 <sup>†</sup> (0.31;0.61)
SF-36-Physical functioning index	0.20 (0.04;0.36)	-0.04 <sup>†</sup> (-0.21;0.13)	0.16 (0.00;0.32)	0.30 <sup>†</sup> (0.14;0.46)
Mental health index	0.06 (-0.11;0.23)	0.05 <sup>†</sup> (-0.12;0.22)	0.25 (0.09;0.41)	0.40 <sup>†</sup> (0.25;0.55)
SF-36-Vitality index	-0.26 <sup>‡</sup> (-0.42;-0.10)	0.09 <sup>‡</sup> (-0.08;0.26)	-0.16 (-0.32;0.00)	0.03 (-0.14;0.20)
Six minutes walk test	0.39 <sup>‡</sup> (0.24;0.54)	0.06 <sup>‡</sup> (-0.11;0.23)	-0.01 (-0.18;0.16)	-0.02 (-0.19;0.15)
Borg scale	-0.50 (-0.64;0.36)	-0.33 (-0.49;-0.17)	-0.46 (-0.61;-0.31)	-0.54 (-0.68;-0.40)

CRQ-IA = Interviewer administered German chronic respiratory questionnaire; CRQ-SA = self-administered German chronic respiratory questionnaire.

\*Pearson correlation coefficient (95% confidence intervals);  $r > 0.28$  significant at  $P < 0.05$ .

<sup>†</sup>Indicate significant differences between the domains of the CRQ-IA and CRQ-SA.

<sup>‡</sup>Indicate significant differences between the individualised and standardised dyspnoea domains.

domain may have important consequences for clinical trials because it affects sample size. Given the same power and level of statistical significance to detect an identical change in HRQL the required sample size in a trial using the standardised dyspnoea domain increases two-fold compared to using the individualised domain. The effect of decreased responsiveness with the standardised dyspnoea domain may be relevant for studies with sample size limitations. In these studies, investigators could use the individualised CRQ dyspnoea domain. It is important to note that changing to standardised items will not reduce the correlation between change in the CRQ dyspnoea domain and change in other HRQL measures. In fact, the correlations between change scores were generally higher in the group randomised to the CRQ-SA. Therefore, if large sample sizes are feasible and investigators feel the greater ease of the standardised items is important in their study, they can choose the standardised approach without fearing they are compromising the CRQs validity as a measure of change over time or cross-sectionally (manuscript in preparation).

For both the interviewer and self-administered CRQ, there were few missing items for the fatigue, emotional function and mastery domain. However,

we observed more missing items for the standardised than for the individualised dyspnoea questions. The latter finding is not surprising because on the individualised domain patients can choose important activities themselves. Therefore, patients are more likely to perform these activities compared with activities provided in the standardised dyspnoea domain. Nevertheless, 94% of all patients completed at least three items at baseline and follow up for both the individualised and standardised domain. If less than three items would have been completed by a large number of patients, validity would have suffered whereas responsiveness is not much affected if there are at least two completed items per domain.<sup>28</sup>

The strengths of our trial include the rigorous methodological approach and the validation of the questionnaire using an intervention of known effectiveness.<sup>7</sup> We trained all interviewers alike to ensure equal adherence to the study protocol. We conducted a multicentre trial in three countries to validate a single CRQ that is applicable in the German-speaking countries in Europe. Finally, we used random allocation, which allowed us to validate four different formats of the CRQ concurrently and to make direct comparisons between them.

**Table 5** Longitudinal validity\* for the fatigue, emotion and mastery domains.

Instrument and domain	CRQ-IA Domains			CRQ-SA Domains		
	Fatigue*	Emotion*	Mastery*	Fatigue*	Emotion*	Mastery*
Feeling thermometer	0.22 (0.06;0.38)	0.27 (0.11;0.43)	0.33 (0.17;0.49)	0.47 (0.32;0.62)	0.42 (0.27;0.57)	0.55 (0.41;0.69)
SF-36-General health perception index	0.34 (0.18;0.50)	0.42 (0.27;0.57)	0.29 (0.13;0.45)	0.56 (0.42;0.70)	0.44 (0.29;0.59)	0.39 (0.24;0.54)
SF-36-Physical functioning index	0.18 (0.02;0.34)	0.02 <sup>†</sup> (-0.15;0.19)	0.10 (-0.07;0.27)	0.41 (0.26;0.56)	0.38 <sup>†</sup> (0.23;0.53)	0.38 (0.23;0.53)
SF-36-Mental health index	0.23 (0.07;0.39)	0.41 (0.26;0.56)	0.33 (0.17;0.49)	0.40 (0.25;0.55)	0.47 (0.32;0.62)	0.39 (0.24;0.54)
SF-36-Vitality index	0.37 (0.22;0.52)	0.30 (0.14;0.46)	0.31 (0.15;0.47)	0.49 (0.34;0.64)	0.41 (0.26;0.56)	0.38 (0.23;0.53)
Six minutes walk test	-0.15 (-0.31;0.01)	-0.10 (-0.27;0.07)	-0.07 (-0.24;0.10)	-0.18 (-0.34;-0.02)	-0.09 (-0.26;0.08)	-0.05 (-0.22;0.12)
Borg scale	-0.48 (-0.63;-0.33)	-0.35 (-0.51;0.19)	-0.35 <sup>†</sup> (-0.51;-0.19)	-0.57 (-0.71;-0.43)	-0.64 (-0.77;-0.51)	-0.68 <sup>†</sup> (-0.80;-0.56)

Correlations for change scores.

CRQ-IA = interviewer administered German chronic respiratory questionnaire; CRQ-SA = self-administered German chronic respiratory questionnaire.

\*Pearson correlation coefficient (95% confidence intervals);  $r > 0.28$  significant at  $P < 0.05$ .

<sup>†</sup>Indicate significant differences between the domains of the CRQ-IA and CRQ-SA.

The trial had the following weaknesses. Follow-up evaluation was performed at the end of the rehabilitation program rather than upon return to the home environment. In addition, due to the absence of a validated German disease specific HRQL instrument and, therefore, we were limited to the use of generic instruments for validation. However, we demonstrated similar improvement of HRQL compared to other respiratory rehabilitation trials.<sup>24,29</sup>

In conclusion, we validated four responsive and valid versions of the CRQ. The resource and time consuming interviewer administration using individualised dyspnoea questions are not any longer a hindrance for the use of the CRQ. If one accepts the presumably lower responsiveness of the standardised dyspnoea questions, the CRQ-SA is an easily applicable instrument.

**Funding:** GlaxoSmithKline Switzerland and the Swiss Lung League funded this study with grants to the Horten Centre. The sponsors were not involved in the study design, conduction of the trial, analysis of data and manuscript writing. MB joined GlaxoSmithKline Germany this year for six months. This engagement was not in relation to this research project.

## Acknowledgements

The CRQ-IA and CRQ-SA are copyrighted by McMaster University; Principal Authors Dr. Gordon Guyatt and Dr. Holger Schünemann. Use of the instrument requires licensing.

We would like to thank Cornelia Flamann (Zuercher Hoehenklinik Wald), Dr. Martin Frey and Dr. Thomas Grueter (both Klinik Barmelweid) for data collection in their centres.

## References

- World Health Report. World Health Organization. [www.who.int/whr/2000/en/statistics.htm](http://www.who.int/whr/2000/en/statistics.htm). Geneva: 2000.
- WHO. The Tobacco Free Initiative, Geneva: The Tobacco Atlas; 2002.
- Forey BA, Lee PN, Fry JS. Updating UK estimates of age, sex and period specific cumulative constant tar cigarette consumption per adult. *Thorax* 1998;53:875–8.
- Cigarette smoking among adults—United States, 2000. *MMWR Morb Mortal Wkly Rep* 2002; 51: 642–45.
- Swiss Federal Statistical Office. Neuchâtel: Swiss Health Survey; 1998.
- Leuenberger P, Kunzli N, Ackermann-Liebrich U, et al. Swiss Study on Air Pollution and Lung Diseases in Adults (SAPALDIA). *Schweiz Med Wochenschr* 1998;128:150–61.
- Lacasse Y, Brosseau L, Milne S, et al. Pulmonary rehabilitation for chronic obstructive pulmonary disease. *Cochrane Database Syst Rev* 2002;CD003793.
- Guyatt GH, Feeny DH, Patrick DL. Measuring health-related quality of life. *Ann Intern Med* 1993;118:622–9.
- Guyatt GH, Berman LB, Townsend M, Pugsley SO, Chambers LW. A measure of quality of life for clinical trials in chronic lung disease. *Thorax* 1987;42:773–8.
- Jones PW, Quirk FH, Baveystock CM, Littlejohns P. A self-complete measure of health status for chronic airflow limitation. The St. George's Respiratory Questionnaire. *Am Rev Respir Dis* 1992;145:1321–7.
- Wijkstra PJ, TenVergert EM, van Altena R, et al. Reliability and validity of the chronic respiratory questionnaire (CRQ). *Thorax* 1994;49:465–7.
- Larson JL, Covey MK, Berry JK, Wirtz S, Kim MJ. Reliability and validity of the chronic respiratory disease questionnaire. *Am Rev Respir Dis* 1993;147:A530.
- 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.
- Singh SJ, Sodergren SC, Hyland ME, Williams J, Morgan MD. 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.
- Jaeschke R, Singer J, Guyatt GH. Measurement of health status. Ascertaining the minimal clinically important difference. *Control Clin Trials* 1989;10:407–15.
- 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.
- Wyrwich KW, Tierney WM, Wolinsky FD. Further evidence supporting an SEM-based criterion for identifying meaningful intra-individual changes in health-related quality of life. *J Clin Epidemiol* 1999;52:861–73.
- Jones PW. Issues concerning health-related quality of life in COPD. *Chest* 1995;107:187S–93S.
- Williams JE, Singh SJ, Sewell L, Guyatt GH, Morgan MD. Development of a self-reported chronic respiratory questionnaire (CRQ-SR). *Thorax* 2001;56:954–9.
- Schunemann HJ, Griffith L, Jaeschke R, et al. A comparison of the original chronic respiratory questionnaire with a standardized version. *Chest* 2003;124:1421–9.
- Guyatt GH, Townsend M, Keller JL, Singer J. Should study subjects see their previous responses: data from a randomized control trial. *J Clin Epidemiol* 1989;42:913–20.
- Bullinger M. Creating and evaluation cross-cultural instruments. In: Spilker B, editor. *quality of life and pharmacoeconomics in clinical trials*. Philadelphia: Lippincott; 1996. p. 659–68.
- Bullinger M. German translation and psychometric testing of the SF-36 Health Survey: preliminary results from the Iqola project. International quality of life assessment. *Soc Sci Med* 1995;41:1359–66.
- Griffiths TL, Burr ML, Campbell IA, et al. Results at 1 year of outpatient multidisciplinary pulmonary rehabilitation: a randomised controlled trial. *Lancet* 2000;355:362–8.
- Schunemann HJ, Griffith L, Jaeschke R, Goldstein R, Stubbing D, Guyatt GH. Evaluation of the Minimal Important Difference for the Feeling Thermometer and the St. George's Respiratory Questionnaire in Patients with Chronic Airflow Obstruction. *J Clin Epidemiol* 2003;56:1170–6.

26. Kirsten DK, Taube C, Lehnigk B, Jorres RA, Magnussen H. Exercise training improves recovery in patients with COPD after an acute exacerbation. *Respir Med* 1998;**92**: 1191–8.
27. Guyatt G, Walter S, Norman G. Measuring change over time: assessing the usefulness of evaluative instruments. *J Chronic Dis* 1987;**40**:171–8.
28. Moran LA, Guyatt GH, Norman GR. Establishing the minimal number of items for a responsive, valid, health-related quality of life instrument. *J Clin Epidemiol* 2001;**54**:571–9.
29. Bendstrup KE, Ingemann JJ, Holm S, Bengtsson B. Out-patient rehabilitation improves activities of daily living, quality of life and exercise tolerance in chronic obstructive pulmonary disease. *Eur Respir J* 1997;**10**:2801–6.