ORIGINAL ARTICLE

Development and Validation of A Patient Reported Outcome Instrument for Chronic Obstructive Pulmonary Diseases*

LI Jian-sheng (李建生)[1], WANG Ming-hang (王明航)[2], YU Xue-qing (余学庆)[2], LI Su-yun (李素云)[2], and XIE Yang (谢 洋)[1]

ABSTRACT **Objective**: To develop and validate a specific patient reported outcome (PRO) for chronic obstructive pulmonary disease (COPD) patients (COPD-PRO) at a set of standardized procedures. **Methods**: Literature analysis, interview and group discussion were performed to draft an initial model of COPD-PRO. Thereafter, 65 clinicians and experts throughout China reviewed the draft scale. Then cognitive debriefing interviews with 40 patients were conducted to assess respondent comprehension of the scale. After that, the revised scale was validated through pre-testing and field-testing. Finally, the psychometric properties of the COPD-PRO were evaluated by indicators such as validity, reliability and responsiveness based on the data from 230 patients. **Results**: The COPD-PRO contained 17 items in 3 domains: amelioration of clinical symptoms, satisfaction of health condition and satisfaction of treatment effect. The Cronbach's α, Split-half coefficient and test-retest coefficient were 0.806, 0.744, 0.703, respectively; the correlation coefficients between domains and overall scale were 0.835–0.963; 5 factors were extracted according to the conceptual model. The differences of the scale scores before and after treatment were statistically significant (*P*=0.000). **Conclusions**: The COPD-PRO has good validity, reliability and responsiveness. The COPD-PRO could provide patients' response to the treatments and then evaluate the effect of treatment in a standardized way.

KEYWORDS lung disease, obstructive, Chinese medicine, patient reported outcome, reliability, validity

Chronic obstructive pulmonary disease (COPD) is a major public health problem throughout the world. Its prevalence is still on the rise, and costs are substantial. The projection for 2020 indicates that COPD will become the third leading cause of death worldwide.[1,2] In the United States, the total economic costs of COPD were estimated to be $49.9 billion in 2010.[3] COPD patients often have anxiety, depression, or other psychological disturbance over time due to breathing discomfort and limitation in daily activities. The goal for COPD management is to improve mobility and health-related quality of life (HRQOL). Therefore, more and more attention has been paid to developing HRQOL questionnaires and patient reported outcome (PRO) scales. In recent years, through international collaborations between clinicians and method experts, standardized questionnaires or scales are now being used extensively in clinical trials to identify and treat the problems that are most important to them.[4-8]

PRO is any report of the health status that comes from the patient directly, generally including key domains such as symptom, functional limitations and physical, mental and social perspective, which can provide patient's perspective on the effectiveness of treatment. For many diseases the patient is really the only source of health outcome endpoint data.[9] PRO is increasingly viewed as an essential complement to traditional clinical evidence for understanding the impact of treatment on patient function and well-being.[10,11] Measurement of PRO is useful to detect individual differences between patients, and may be an indicator to predict important health outcomes such as hospital admissions or HRQOL.[12]

With remarkable longevity and current popularity, Chinese medicine (CM) has gradually been as a kind of complementary treatment to Western medicine throughout the world. CM and PRO have similar theoretical and clinical underpinnings to a large extent that both focus on

*Supported by 2011 Special Fund for Chinese Medicine-Scientific Research in the Public Interest of Ministry of Finance People's Republic of China and State Administration of Traditional Chinese Medicine (No. 201107002)
1. Geriatric Department, Henan University of Traditional Chinese Medicine, Zhengzhou (450046), China; 2. Department of Respiratory, The First Affiliated Hospital of Henan University of Traditional Chinese Medicine, Zhengzhou (450000), China
DOI: 10.1007/s11655-014-1982-4

Chin J Integr Med 2015 Sep;21(9):667-675

people' physical, psychological, social, and environmental characteristics and their mutual relationship. In recent two or three decades, PRO was mainly used as a secondary outcome in many clinical trials. CM-PRO has developed rapidly, such as the Chinese Quality of Life measures,[13] the Spleen and Stomach Disease PRO scale,[14] the TCM Stroke Scale for QOL,[15] and the Myasthenia Gravis patients of PRO scale.[16] These instruments include generic or disease-specific types, and self or interviewer-administered versions. However, there is a lack of research in PRO for COPD patients according to the concept of lung disease in CM, especially that can reflect the CM concept of COPD.

The purpose of this study was to develop and validate a specific PRO scale for COPD patients (COPD-PRO) based on the concept of Lung (Fei) in CM and a set of standardized procedures that was commonly recognized and accepted in Western medicine.

METHODS

Development of the Initial COPD-PRO

Research Group

It was composed by 25 multi-disciplinary and multi-level staffs that are clinicians and nurses, scale researchers, respiratory disease experts, linguists, statisticians, psychologist, sociologists, public health scholars, COPD patients, other diseases patients and health persons. The research group was divided into a core group and a discussion group. The core group was responsible for decision-making, organization and scale evaluation. The other group was responsible for putting forward questions and participating in the discussion for the procedure of COPD-PRO development and evaluation.

Establishment of the Initial Structure

According to the results of literature review on relevant PRO instruments for COPD patients, e.g. HRQOL questionnaire, physical activity scale and dyspnea rating scale, and integrated with characteristics of CM, the core team put forward the initial model of COPD-PRO, which also took into account China's actual conditions and cultural identity. Then an expert panel meeting was held. The discussion group was asked to comment on the initial model and provide suggestion on items generation.

Items Generation

Through literature analysis, interview and focus group discussion, potentially relevant items for the COPD-PRO were collected. To begin with, the item pool was derived from publications on relevant clinical research of COPD home and abroad. Then 45 COPD patient interviews were conducted to collect their opinions on the items. Moreover, focus group discussions were performed to define and analyze the relevant items response to each domain that drafted by experts. To ensure that all possible items were included, we also reviewed other COPD questionnaires and consulted a number of clinicians involved in the treatment of COPD patients. Finally, an initial item pool was generated by the team.

Pre-evaluation of the Initial COPD-PRO

Review of the Drafted Items by Clinicians and Experts

Totally 65 clinicians and experts in COPD management throughout China were invited to participate in reviewing the drafted items and the domain structure of the COPD-PRO. They were asked to give their comments whether they agreed on the domains and the items, and whether the draft items should be improved or new items should be added. With their help, a set of revised items were developed for the cognitive debriefing interviews.

Cognitive Debriefing Interviews

Forty COPD patients who were receiving CM treatment were interviewed to assess respondent comprehension of the revised scale. They were asked to comment on the linguistic and semantic clarity of the items as well as the improvement in the item wordings. Then the revised item pool was prepared in the pre-evaluation study.

Response Scales Selection and Domain Scores

Referred to the 5-point response scales in the World Health Organization Quality of Life Assessment (WHOQOL)-100,[17] a typical format of a Likert five-level response[18] was adopted in the COPD-PRO, which ranges from 1 to 5. For each item, score 1 referred to the highest QOL and 5 referred to the lowest QOL, each domain score was got by summing every included item score. The lower score of the scale, the better QOL of the patients had.

Pre-Testing Survey

To validate the domains with its items of the initial COPD-PRO, the pre-testing survey was

adopted. One hundred and twenty patients were surveyed in Zhengzhou, China. After that, item pool was revised again and prepared for the field testing.

Development of the Final COPD-PRO and Its Evaluation

Field Test

A total of 230 COPD patients were recruited to examine psychometric properties of the COPD-PRO. In the field test, all patients were asked to fill in the field test version of COPD-PRO, and comment on their own feelings by answering a single question on self-perceived health status, i.e., rate their cough status as "never" or "seldom" or "often" or "quite often" or "always", comment on their disease condition whether aggravated by weather changes by responding "never" or "seldom" or "often" or "quite often" or "always". Assistance was provided to those who had difficulties in reading and writing due to various reasons.

Methods of Screening Item

Five common methods as well as the principle of certainty of screening item were adopted for the development COPD-PRO. (1) The experts grading method. From the perspective of importance, experts were invited to give their score on each item, ranging from 1 to 100, 1 referred to not important, and 100 referred to the very important. If the mean score <90, the item would be reduced. (2) Discrete trend method, from the perspective of sensitivity to select items. Because the objective dimension of each item is the same, the standard deviation (SD) can be used to reflect the discrete trend to a large extent. Therefore, if the SD <0.7, the item would be reduced. However, as for the domain with less items (1–2 items), the item with minimum SD would be reduced. (3) Correlation coefficient method, from the perspective of representative and independence to select items. Through calculation the Pearson correlation coefficients between the items and its domain, if the correlation coefficient <0.6, the item would be reduced. (4) Cronbach's α method, from the perspective of internal consistency to select items. Calculating the Cronbach's α of any domain, then deleting one item in any domain and calculating the α again. If the Cronbach's α of any domain increased significantly when one item deleted, the deleted item influenced the internal consistency and would be reduced. (5) Factor analysis, from the perspective of representative to select items. According to the correlation matrix of standardization item, the factor analysis was made and the variance was went maximum rotation. Then selecting the item by factor loading coefficient, if the item loading coefficient <0.4 in one factor, the item would be reduced.

Statistical Analysis

Feasibility was calculated by qualified answering rate and completion time of COPD-PRO. Internal consistency reliability was measured by Cronbach's α coefficient.[19] Test-retest reliability was estimated by intraclass correlations (ICC) between the baseline and after treatment. The ICC was estimated using a fixed-effects analysis of variance (ANOVA) model.[20] Content validity was evaluated by Pearson's correlation between domains and overall scale, and correlation between the domains and its items. Structural validity was measured by exploratory factor analysis and maximum variance rotation, standard for choosing characteristic factors was the root >1. Responsiveness was assessed by comparing the difference of the scale scores before and after treatment as well as computing the standardized response mean (SRM).[21] A *P* value <0.05 was considered statistically significant. Data analysis was performed by SPSS 19.0 (License No. 6f1d84c801f1e6010dc).

RESULTS

Theoretical Framework of COPD-PRO in CM

Ancient and modern literature on the concept of Lung in CM was analyzed to establish the theoretical basis. Then, the theoretical framework of COPD-PRO was deducted from our understanding of the core concepts of Lung in CM. The Lung situates in the chest, connects with the throat and opens into the nose. Its main physiological functions and indicators are: (1) dominating qi and controlling respiration; (2) dominating the dispersion and descent; (3) dominating the skin and hair; (4) regulating water passages; and (5) opening into the nose. If the Lung function is abnormal, or pathogenic phlegm obstructs the Lung, it will bring about an abnormal flowing of Lung qi and further result in pathological changes such as cough, fullness of chest, and gurgling with sputum. If there is a deficiency of Lung qi, there will be feeble respiration, uneven breathing, weak speech, fatigue, etc.

A half day expert panel meeting was held. Based on the discussion, the research group revised the

Chin J Integr Med 2015 Sep;21(9):667-675

theoretical framework of COPD-PRO. The framework consisted of 3 domains to reflect the endpoints of amelioration of clinical symptoms, the satisfaction of health conditions and the satisfaction of treatment effect.

Development of Items

According to the revised domain, many items were drafted by the group. The number of items in the amelioration of clinical symptom domain was 34, and that in the satisfaction of health conditions domain was 3, in the satisfaction of treatment effect domain was 2 (Table 1). The process of screening item including 4 steps: firstly, review of the drafted items by clinicians and experts; secondly, cognitive debriefing interviews with patients; thirdly, pre-testing survey; and finally field test.

To begin with, 65 clinicians and experts in COPD management throughout China were invited to give their comments on the domains and the items. In this step, the experts grading method and discrete trend method were adopted. Therefore, 13 items in the amelioration of clinical symptoms domain were deleted due to low importance, low correlation and reduplication, which were hemoptysis, hoarseness, hot flash, tinnitus, diarrhea, night sweating, feverishness in palms and soles, dry mouth and nose, chest pain, cold limbs, soreness and weakness of waist and knees, abdominal distention, irritable. As a result, 26 revised items were developed for the cognitive debriefing step (Table 1).

In addition, cognitive debriefing was performed with 40 COPD patients who were receiving CM treatment. The demographics of the sample were shown in Table 2. They were asked to answer the 26 questions and comment on the linguistic and semantic clarity of the items as well as the improvement in the item wordings. In this session, correlation coefficient method and Cronbach's α method were adopted. Therefore, 6 items were removed in the end. Specifically, 5 items, muscle pain, chill and cold, emaciation, upset and dry throat in the amelioration of clinical symptoms domain were deleted, and 1 item, satisfaction of health services in the satisfaction of health conditions domain was deleted. As a result, 20 items were developed for pre-testing survey step (Table 1).

Moreover, the pre-testing survey was adopted to validate the appropriateness and structure fitness of the revised COPD-PRO. One hundred and twenty patients were surveyed in Zhengzhou, China. The demographics of 120 samples were shown in Table 2. In this session, the Cronbach's α method was adopted. Two items in the amelioration of clinical symptoms domain were removed due to low correlation, which are palpitation and edema. As a result, 18 items were developed for field test step (Table 1).

Finally, 230 COPD patients were recruited to test the psychometric properties of the COPD-PRO. The demographics of 230 samples were shown in Table 2. Through exploratory factor analysis and confirmatory factor analysis, 1 item, sweating in the amelioration of clinical symptoms domain was excluded with factor loading less than 0.4. Eventually, the number of the final items was 17, and the domain and its item of the final model were shown in Table 1. The COPD-PRO in the original Chinese and English translation versions were listed in Appendix 1 (supplementary material available at http://bbs.etjournals.com/showtopic-n.aspx).

Psychometric Properties of COPD-PRO

Feasibility

Two hundred and thirty scales were sent to the 230 patients who were recruited for the field test, and 215 qualified scales were got in the end. Therefore, the qualified answering rate was 93.48%. Meanwhile, the completion time of COPD-PRO was about 6.6 min.

Table 1. Number of Items in Various Stages of Development

Domains	Number of item				
	Initial items drafted by the research team	Review by clinicians and experts	Cognitive debriefing interviews	Pre-testing survey	Field test
Amelioration of clinical symptom	34	21	16	14	13
Satisfaction of health conditions	3	3	2	2	2
Satisfaction of treatment effect	2	2	2	2	2
Overall COPD-QOL	39	26	20	18	17

Table 2. Demographic Characteristics for COPD Samples

Characteristics	Cognitive debriefing interview sample (*n*=40)	Pre-testing survey sample (*n*=120)	Field test sample (*n*=230)
Age (Year)			
Mean (SD)	61.2 (10.7)	62.7 (9.8)	63.9 (9.2)
Range	40–80	40–80	40–80
Course of disease (Month)			
Mean (SD)	158.2 (143.4)	183.2 (337.3)	159.0 (255.2)
Range	18–456	12–360	12–414
Gender [Case (%)]			
Male	28 (70.0)	86 (71.7)	161 (70.0)
Female	12 (30.0)	34 (28.3)	69 (30.0)
Education level [Case (%)]			
Primary school	2 (5.0)	12 (10.0)	40 (17.4)
Middle school	15 (37.5)	29 (24.2)	56 (24.3)
High school	15 (37.5)	35 (29.2)	64 (27.8)
College diploma	4 (10.0)	27 (22.5)	49 (21.3)
University degree or above	1 (2.5)	11 (9.2)	19 (8.3)
Missing	3 (7.5)	6 (5.0)	2 (0.9)
Smoking status [Case (%)]			
Current smoking	14 (35.0)	49 (40.8)	95 (41.3)
None smoking	26 (65.0)	71 (59.2)	135 (58.7)
Gold stage [Case (%)]			
Mild	0 (0)	7 (5.9)	9 (3.9)
Moderate	32 (80.0)	52 (43.3)	97 (42.2)
Severe	8 (20.0)	61 (50.8)	124 (53.9)

Distribution of Item Score

The score in each item was normally distributed. The mean item score ranged from 2.13 to 3.12 in the amelioration of clinical symptoms domain. The mean score in the satisfaction of health condition domain was 2.81, 2.79, and that in the satisfaction of treatment effect domain was 2.49, and 2.50. No ceiling and floor effects were noted (Table 3).

Reliabilities

The reliability of COPD-PRO was evaluated by 2 aspects: internal consistency and test-retest. The internal consistencies of the domains were all good. Cronbach's α was 0.806 for the overall score and that in each domain was 0.809, 0.979, 0.954, respectively. Split-half coefficient was 0.744 for the overall score and that in each domain was 0.768, 0.979, 0954, respectively (Table 3). Test-retest reliability study was conducted on 100 patients within 1 week. Test-retest reliability was good, with ICC values all above 0.70. The ICC value for each domain and the overall COPD-PRO were 0.720, 0.834, 0.783, 0.703, respectively (Table 4).

Table 3. Distributions of the COPD-QOL Item Score (*n*=215)

Items	Mean	SD	Floor effect (%)	Ceiling effect (%)
A1 Cough	2.66	0.808	5.6	0.9
A2 Sputum	2.82	0.878	9.3	0.9
A3 Phlegm easily expectorated	3.03	0.785	1.8	1.3
A4 Tightness	2.95	0.892	7.9	0.4
A5 Pant	2.27	1.108	4.7	2.8
A6 Fatigue	2.72	0.872	6.5	1.3
A7 Cyanosis	2.59	0.946	6.9	0.4
A8 Common cold	2.41	1.009	6.1	0.9
A9 Appetite	2.13	1.069	8.4	0.4
A10 Sleep	2.61	0.835	4.7	1.9
A11 Overtired	2.90	0.883	4.6	3.3
A12 Weather change	3.12	0.924	2.8	4.7
A13 Mood swings	2.51	0.869	8.8	2.3
B1 Satisfaction of health	2.81	0.674	4.7	0.4
B2 Satisfaction of capability	2.79	0.671	5.1	0.4
C1 Satisfaction of disease amelioration	2.49	0.674	4.1	0.4
C2 Satisfaction of treatment effect	2.50	0.632	3.2	0.4

Content Validity

Content validity was evaluated by Pearson's correlations between domains to overall scale, and item correlations with its domain. The domain to overall COPD-PRO correlation ranged from 0.835 to 0.963 (Table 5). Moreover, there were good correlations among each domain score, which ranged from 0.837 to 0.889 (Table 5). For item correlations with its domains, each item correlated highly with its own domain, the Pearson's correlation ranged from 0.457 to 0.702 in the amelioration of clinical symptoms domain, that in the satisfaction of health condition domain was 0.989 and 0.989, and that in the satisfaction of treatment effect domain was 0.980 and 0.977 (Table 6).

Construct Validity

Construct validity was measured by exploratory factor analysis method and maximum variance rotation. Based on the results of Kaiser-Meyer-Olkin measure of sampling (KMO, *P*=0.704) and Bartlett's test (Chi-Square=1683.941, *P*=0.000), the partial correlation is very weak, and exploratory factor analysis method was suitable.

Through the exploratory factor analysis on the 17

Table 4. Internal Consistency and Test-Retest Reliability of COPD-PRO

Domains	Cronbach's α (*n*=215)	Split-half coefficient (*n*=215)	ICC (*n*=100)
Amelioration of clinical symptoms	0.809	0.744	0.720
Satisfaction of health conditions	0.979	0.768	0.834
Satisfaction of treatment effect	0.954	0.979	0.783
Overall COPD-QOL	0.806	0.954	0.703

Table 5. Correlation Coefficients between Domains and Overall COPD-PRO (*n*=215)

Domains	Amelioration of clinical symptoms	Satisfaction of health conditions	Satisfaction of treatment effect	Overall COPD-QOL
Amelioration of clinical symptoms	1			0.963**
Satisfaction of health conditions	0.877**	1		0.877**
Satisfaction of treatment effect	0.837**	0.889**	1	0.835**

Table 6. Pearson Correlation Coefficients between Items and Domains of COPD-PRO (*n*=215)

Items	Amelioration of clinical symptoms	Satisfaction of health conditions	Satisfaction of treatment effect
A1 Cough	0.640**	0.135	0.063
A2 Sputum	0.459**	0.102*	0.037
A3 Phlegm easily expectorated	0.508**	0.078	0.093
A4 Tightness	0.580**	0.163*	0.127
A5 Pant	0.666**	0.082	0.137*
A6 Fatigue	0.447**	0.079	0.070
A7 Cyanosis	0.606**	0.123	0.084
A8 Common cold	0.702**	0.098	0.057
A9 Appetite	0.457**	0.070	0.012
A10 Sleep	0.605**	0.075	0.108
A11 Overtired	0.532**	0.125	0.070
A12 Weather change	0.584**	0.066	0.093
A13 Mood swings	0.380**	0.081	0.047
B1 Satisfaction of health	0.180**	0.989**	0.197**
B2 Satisfaction of capability	0.171*	0.989**	0.177**
C1 Satisfaction of disease amelioration	0.126	0.222**	0.980**
C2 Satisfaction of treatment effect	0.143*	0.146*	0.977**

Note: Correlation is significant at the 0.05 level (2-tailed)* or 0.01 level (2-tailed)**

Table 7. Explanation of the Total Variance of Factor Analysis

Factor	Initial eigenvalues			Extraction sums of squared loadings			Rotation sums of squared loadings		
	Value	Variance (%)	Cumulative (%)	Value	Variance (%)	Cumulative (%)	Value	Variance (%)	Cumulative (%)
1	4.280	25.175	25.175	4.280	25.175	25.175	2.768	16.280	16.280
2	2.127	12.513	37.688	2.127	12.513	37.688	2.412	14.185	30.465
3	1.666	9.798	47.485	1.666	9.798	47.485	1.969	11.584	42.049
4	1.583	9.311	56.796	1.583	9.311	56.796	1.951	11.475	53.525
5	1.119	6.584	63.379	1.119	6.584	63.379	1.675	9.855	63.379

items, 5 factors were generated that explained about 63.379% of cumulative variance of the data in the data set. Factor 1 consisted of 8 items in the amelioration of clinical symptoms domain, which explained about 25.1175% of the total variance of the data. Factor 2 consisted of 5 items in the amelioration of clinical symptoms domain, which explained about 12.513% of the total variance. Factor 3 consisted of all the 2 items in the satisfaction of health condition domain, which explained 9.798% of the total variance. Factor 4 consisted of all the 2 items in the satisfaction of treatment effect domain, which explained 9.311% of the total variance. Factor 5 consisted of 2 items in the amelioration of clinical symptoms domain, which explained about 6.584% of the total variance. Factor 1, 3 and 4 made up mostly of items of the COPD-PRO, which explain more than 2/3 of the total variance (Tables 7 and 8).

Table 8. Exploratory Factor Analysis on the 17 Items of COPD-PRO

Items	Component				
	1	2	3	4	5
A1 Cough	0.456				0.742
A2 Sputum					0.887
A3 Phlegm easily expectorated	0.417				
A4 Tightness		0.440	0.469		
A5 Pant	0.655				
A6 Fatigue	0.599				
A7 Cyanosis	0.579				
A8 Common cold	0.701				
A9 Appetite		0.631			
A10 Sleep		0.436			
A11 Overtired		0.811			
A12 Weather change		0.835			
A13 Mood swings		0.562			
B1 Satisfaction of health			0.977		
B2 Satisfaction of capability			0.978		
C1 Satisfaction of disease amelioration				0.961	
C2 Satisfaction of treatment effect				0.969	

Notes: Extraction method: principal component analysis; Eigen value >1; rotation method: varimax rotation with Kaiser normalization. Sorted by factor names; factor loading <0.5 suppressed

Responsiveness

The paired-sampled T test or Wilcoxon U test according to sample distribution was used to test the responsiveness. For the three domains and the overall COPD-PRO, the differences between the scale scores before and after treatment were statistically significant (P=0.000), as well as the reasonably higher SRM (Table 9).

DISCUSSION

According to the U.S. Food and Drug Administration guidance for industry PRO measures[11] and CM theory of Lung, the COPD-PRO was developed by the programmed decision procedures:[22-24] Development of the initial scale, preliminary evaluation of the initial scale, development of the final scale and its evaluation. As for the COPD-PRO, we cared more about the effect of the interventions, especially in

Table 9. Responsiveness to Change of COPD-PRO Domain Scores before and after Treatment (n=215)

Domains	Before treatment		After treatment		t/Z	P	SRM
	Mean	SD	Mean	SD			
Amelioration of clinical symptoms	34.73	6.58	27.93	4.47	15.060	0.000	1.03
Satisfaction of health conditions	5.59	1.33	3.42	1.12	19.992	0.000	1.63
Satisfaction of treatment effect	4.99	1.28	2.76	0.99	18.396	0.000	1.74
Overall COPD-QOL	45.31	7.26	34.11	5.55	20.707	0.000	1.54

symptom control. Hence, the items about symptoms and satisfaction were set as an independent domain other than going to the physical domain. Therefore, COPD-PRO consists of 3 domains: the amelioration of clinical symptoms, the satisfaction of health conditions and the satisfaction of treatment effect.

The item selection was based on not only qualitative analysis such as focus group discussions and interviews but also special CM concepts of health. In this study, we developed some items that reflected the characteristic of CM, for example, 'Does the overtired or mood swings can aggravate your disease?' and 'How is your appetite or sleep?' The item pool consisting of 39-item was formed. And the numbers of items in the final version were effectively reduced from 39 to 17. Item reduction was carried out by four steps and five methods. Meanwhile, a practical PRO scale should be evaluated at least 3 aspects: reliability, validity and responsiveness. Results showed that the psychometric properties of the COPD-PRO were good, which is valid, reliable and promises to be responsive to changes in patients.

Reliability refers to the reproducibility or consistency of scores from one assessment to another.[25] Internal consistency reliability (Cronbach's α) and test-retest reliability (ICC) are the most frequently used indicators. Cronbach's α with values >0.70 stands for larger reliability. From Table 3, the Cronbach's α of each domain and overall COPD-PRO were all above 0.80. The ICC values were all above 0.70. Thus, the COPD-PRO had good reliability.

Validity is the extent to which an instrument captures what it purports to measure. Validity was divided into content validity and construct validity. The former was evaluated by Pearson's correlations between domains to overall scale, and item correlations with its domain.[25] The latter was measured by exploratory factor analysis and maximum variance rotation. From Tables 5 and 6, the correlation between items to domain, and domains to overall COPD-PRO were all high. From Tables 7 and 8, 5 factors were generated and the factor 1, 3 and 4 could make up mostly of items of the COPD-PRO and explain more than 2/3 of the total variance. COPD-PRO had good content validity and construct validity.

Responsiveness refers to an instrument's ability to detect change, which is important in clinical applications.[26] We focused on the internal responsiveness under the hypothesis that the sensitive instrument could reflect changes after treatment. The paired-sampled T test was used to compare the responses before and after treatment. Moreover, the SRM value was also used, with its value of 0.20, 0.50, and 0.80 standing for small, moderate, and large responsiveness, respectively. From Table 9, it can be seen that the changes of each domain and the overall COPD-PRO, after treatment can be identified, resulting in SRM from 1.03 to 1.74. Before and after treatment, there were significant differences between the mean scores. The COPD-PRO had good responsiveness.

In summary, the COPD-PRO showed strong internal consistency, test-retest reliability, content and construct validity, and responsiveness through evaluation. However, there were some limitations of this study. To begin with, the COPD-PRO was established in form of a questionnaire and this could only capture information, i.e. signs, symptoms, emotions and satisfaction, that could be consciously aware by the subjects. There might be some other important information, e.g. the pulse, which could only be assessed by CM practitioners not included in the scale. Moreover, comparison between COPD-PRO and other standard scale was also the limitation. In addition, we relied very much on the exploratory factor analysis in establishing and testing of the structure of the COPD-PRO, and the sample size was relatively small to ensure a robust structure. Therefore, further studies on the construct validity of the scale are necessary.

Acknowledgement

The authors acknowledge Prof. LIU Feng-bin, First Affiliated Hospital of Guangzhou University of Chinese Medicine; Prof. WAN Chong-hua, School of Public Health, Kunming Medical University; Prof. HAO Yuan-tao, School of Public Health, Sun Yat-Sen University; Dr. ZHAO Li, Hong Kong Baptist University; Profs. LIU Bao-yan, XIE Yan-ming, HU Jing-qing, HE Li-yun, China Academy of Chinese Medical Sciences; Prof. SUN Zi-kai, Jiangsu Province Traditional Chinese Medicine Hospital; Chief physician MA Li-jun, Henan Province People's Hospital; Chief physician TANG Bing-xiang, Henan Provincial Chest Hospital; Prof. ZHANG Guo-jun, The First Affiliated Hospital of Zhengzhou University; Chief physician QU Bing, Henan Province Chinese Medicine Research Institute; Prof. XU Jingsheng, Profs. FAN Yu-hong, CUI Ying-min, Dr. XU Zhen-

guo, Henan University of Traditional Chinese Medicine, for their great assistance in the preparation of methodology, psychology and linguistics of the scale, the basic theory of CM, clinical respiratory disease and their suggestion. The authors are also grateful to Ministry of Finance People's Republic of China and State Administration of Traditional Chinese Medicine for funding this study.

REFERENCES

1. Murray CJ, Lopez AD. Alternative projections of mortality and disability by cause 1990–2020: global burden of disease study. Lancet 1997;349:1498-1504.
2. Raherison C, Girodet PO. Epidemiology of COPD. Eur Respir Rev 2009;18:213-221.
3. National Heart Lung and Blood Institute. Morbidity and mortality: 2009 chart book on cardiovascular, lung, and blood diseases. 1st ed. American: National Heart, Lung, and Blood Institute;2009:62-66.
4. 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;10:773-778.
5. Jones PW, Quirk FH, Baveystock CM. The St George's respiratory questionnaire. Respir Med 1991;85:25-31.
6. Meguro M, Barley EA, Spencer S, Jones PW. Development and validation of an improved, COPD-specific version of the St. George respiratory questionnaire. Chest 2007;2:456-463.
7. van der Molen T, Willemse BW, Schokker S, ten Hacken NH, Postma DS, Juniper EF. Development, validity and responsiveness of the clinical COPD questionnaire. Health Qual Life Outcomes 2003;1:13.
8. Jones PW, Harding G, Berry P, Wiklund I, Chen WH, Kline Leidy N. Development and first validation of the COPD assessment test. Eur Respir J 2009;3:648-654.
9. Revicki DA, Osoba D, Fairclough D, Barofsky I, Berzon R, Leidy NK, et al. Recommendations on health-related quality of life research to support labeling and promotional claims in the United States. Qual Life Res 2000;8:887-900.
10. Acquadro C, Berzon R, Dubois D, Leidy NK, Marquis P, Revicki D, et al. Incorporating the patient's perspective into drug development and communication: an ad hoc task force report of the Patient-Reported Outcomes (PRO) Harmonization Group meeting at the Food and Drug Administration, February 16, 2001. Value Health 2003;5:522-531.
11. U.S Department of Health and Human Services Food and Drug Administration Guidance for Industry. Patient-reported outcome measures: use in medical product development to support labeling claims. U.S. FDA, clinical/medical (Internet). 2009 Dec (cited 2011 Jan 5). Available from: http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM193282.pdf
12. Chen W, Liu JP. Application of patient-reported outcomes in clinical research on therapeutic efficacy. Chin J Integr Tradit West Med (Chin) 2009;8:746-749.
13. Leung KF, Liu FB, Zhao L, Fang JQ, Chan K, Lin LZ. Development and validation of the Chinese quality of life instrument. Health Qual Life Outcomes 2005;3:26.
14. Liu FB, Wang WQ. Establishment of the spleen-stomach patients reported outcomes scale in Chinese medicine and the corresponding item selection. World Sci Technol Mod Tradit Chin Med (Chin) 2009;4:527-530.
15. Li H, Liang WX. Study and development of TCM Stroke Scale for quality of life measurement (1)—establishment of the scale. Liaoning J Tradit Chin Med (Chin) 2008;3:376-378.
16. Chen XL, Liu FB, Guo L. Development of patient-reported outcome scale for myasthenia gravis: a psychometric test. J Chin Integr Med (Chin) 2010;2:121-125.
17. The WHOQOL Group. The World Health Organization Quality of Life assessment (WHOQOL): position paper from the World Health Organization. Soc Sci Med 1995;10:1403-1409.
18. Jamieson S. Likert scales: how to (ab) use them. Med Educ 2004;12:1217-1218.
19. Deyo RA, Diehr P, Patrick DL. Reproducibility and responsiveness of health status measures. Statistics and strategies for evaluation. Control Clin Trials 1991;4(Suppl):142S-158S.
20. Cronbach LJ. Coefficient alpha and the internal structure of tests. Psychometrika 1951;3:297-334.
21. Terwee CB, Dekker FW, Wiersinga WM, Prummel MF, Bossuyt PM. On assessing responsiveness of health-related quality of life instruments: guidelines for instrument evaluation. Qual Life Res 2003;4:349-363.
22. Donald L, Patrick. Patient-reported outcomes (PROs). An organizing tool for concepts, measures and applications. QoL Newsletter 2003;31:1-6.
23. Wyrwich KW, Bullinger M, Aaronson N, Hays RD, Patrick DL. Estimating clinically significant differences in quality of life outcomes. Qual Life Res 2005;2:285-295.
24. Coons SJ, Gwaltney CJ, Hays RD, Lundy JJ, Sloan JA, Revicki DA, et al. Recommendations on evidence needed to support measurement equivalence between electronic and paper-based patient-reported outcome (PRO) measures: ISPOR ePRO Good Research Practices Task Force report. Value Health 2009;4:419-429.
25. Fang JQ, ed. The methods and application of quality of life measurement. 1st ed. Beijing: Beijing Medical University Press; 2000:263-266.
26. Wright JG, Young NL. A comparison of different indices of responsiveness. J Clin Epidemiol 1997;3:239-246.

(Received August 5, 2013)
Edited by YUAN Lin