REVIEW

Further Refinement is Required for Patient-Reported Outcome Scales for Respiratory Diseases Based on Traditional Chinese Medicine Theory for Applicability

Peng Zhang [1,2], Jiaming Ren[1], Baichuan Xu [1], Jiajia Wang[2-4], Yang Xie [2-4]

[1]The First Clinical Medical College, Henan University of Chinese Medicine, Zhengzhou, Henan, People's Republic of China; [2]Department of Respiratory Diseases, The First Affiliated Hospital of Henan University of Chinese Medicine, Zhengzhou, Henan, People's Republic of China; [3]Collaborative Innovation Center for Chinese Medicine and Respiratory Diseases Co-Construction by Henan Province & Education Ministry of P.R. China, Henan University of Chinese Medicine, Zhengzhou, People's Republic of China; [4]Henan Key Laboratory of Chinese Medicine for Respiratory Disease, Henan University of Chinese Medicine, Zhengzhou, People's Republic of China

Correspondence: Yang Xie, First Affiliated Hospital of Henan University of Chinese Medicine, No. 19 Renmin Road, Zhengzhou, Henan, 450046, People's Republic of China, Tel +86 371 66248624, Fax +86 371 66248624, Email xieyanghn@163.com

Objective: To summarize the contents and assess the methodological quality and measurement properties of the patient-reported outcome (PRO) scales featured with Traditional Chinese Medicine (TCM) for respiratory diseases based on the guideline of COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN).

Methods: PubMed, Web of Science, Embase, China National Knowledge Infrastructure (CNKI), Wanfang Data, VIP, and China Biology Medicine (CBM) were searched for studies on PRO scales featured with TCM for respiratory diseases from their inception until December 2022. The characteristics of the PRO scales were qualitatively summarized. Following the COSMIN guideline, the risk of bias was assessed according to the checklist, and different measurement properties (content validity, structural validity, internal consistency, reliability, criterion validity, and responsiveness) were evaluated. Finally, the evidence's overall quality was assessed, and the recommendation was formulated using the modified GRADE approach.

Results: A total of 13 scales were included, with 6 for chronic obstructive pulmonary disease (COPD), 3 for lung cancer, 2 for idiopathic pulmonary fibrosis (IPF), 1 for community-acquired pneumonia (CAP), and 1 for bronchiectasis. All 13 scales are disease-specific scales and were developed based on Chinese cultural background to measure the efficacy of TCM. The study did not provide information on measurement error, cross-cultural validity, and hypothesis testing for the construct validity of these measures. No scale was rated as sufficient in content validity and responsiveness. Two scales showed sufficient structural validity, while 11 scales exhibited sufficient internal consistency. Three scales demonstrated sufficient reliability, and 7 scales showed sufficient criterion validity. All 13 scales have a recommendation level of B.

Conclusion: The 13 scales could reflect the clinical efficacy of TCM and are suitable for the Chinese population. Nevertheless, the validation of these scales was not comprehensive enough, and the methodological quality of their studies needs to be further strengthened.

Keywords: patient-reported outcome, quality of life, respiratory disease, traditional Chinese medicine, COSMIN

Introduction

With the change in medical models, the methods of clinical efficacy evaluation have gradually changed from physical and chemical indexes in the laboratory to a comprehensive evaluation of patients' physiological, psychological, and social activity.[1] In certain instances, patient-reported outcomes (PROs) may be the only feasible endpoints, such as fatigue or pain assessments, as improvements in physiologic or clinical endpoint outcomes may not always reflect improvements in the patient's disease state with regard to treatment efficacy.[2] In addition, patients' subjective perceptions of their disease are often inconsistent with their doctor's recognition of the condition.[3] PRO instruments could collect information from multiple dimensions including psychological status, physiological function, social activity, diagnosis

and treatment satisfaction, etc., based on the patient's own perspective, which can assess the patient's current health status or quality of life (QOL).[4,5]

There is a significant percentage of respiratory disease patients in China, especially chronic respiratory diseases that seriously endanger human health by leading to limited daily activity ability, and reduced social activities, thus further affecting patients' emotions and psychology and reducing their QOL.[6–8] The commonly used PRO instruments in respiratory diseases, such as the Clinical COPD Questionnaire (CCQ)[9] and St. George's Respiratory Questionnaire (SGRQ),[10] were developed in Europe. At present, the Chinese version of the Western questionnaire is widely used in the clinical studies of Traditional Chinese medicine (TCM). However, it is problematic to apply the translated questionnaires to clinical studies in China, especially TCM, because health-related measurement instruments developed in Western cultures may not involve concepts that are relevant or essential in Chinese culture, let alone reflecting the issues that patients are more concerned about the treatment process of TCM.[11]

With the introduction of PRO to China, many researchers were delighted to find that the PRO scale with TCM characteristics, as an instrument for evaluating clinical efficacy, can largely quantify this process and then objectively present the subjective feelings of patients after treatment.[12] TCM has certain therapeutic advantages in treating chronic respiratory diseases,[13–16] and many researchers have developed a lot of PRO scales for respiratory diseases based on TCM theory and Chinese cultures, aiming to explore new means to assess the clinical effectiveness of TCM. However, these scales have hardly been promoted and applied currently. Therefore, this study assessed the methodological quality and measurement properties of the scales featured with TCM for respiratory diseases based on the guidelines of the COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) that points out the problems in the development methods and summarizes their characteristics, with the hope of promoting their clinical application and providing some reference for the correct clinical use.

Methods
Search Strategy
Seven databases, including PubMed, Web of Science, Embase, China National Knowledge Infrastructure (CNKI), Wanfang Data, VIP, and China Biology Medicine (CBM), were searched from their inception until December 2022. The following three groups of search terms in English were used: (1) "traditional Chinese medicine" and "TCM" connected with "OR"; (2) "patient-reported outcome", "PRO", "quality of life", and "QOL" connected with "OR"; (3) "assessment", "measur*" "scal*" "questionnaire*" "instrument", "scor*" and "tool" connected with "OR". Finally, the search terms of the three groups above were connected with the term "AND". We reviewed the references of the original articles for potential relevant studies manually and also attempted to get grey literature from other sources. Relevant search terms in Chinese have also been searched. Examples of the searching strategies used in PubMed (English language) and CNKI (Chinese language) are presented in the supplementary materials: Table S1.

Study Selection
Inclusion Criteria
1. Studies should concern PRO scales, and the aim of the study should be the evaluation of one or more measurement properties and the development of a scale (to rate the content validity).
2. Scales for respiratory system diseases with main lesions in the trachea, bronchi, and lung.
3. The article should clearly state that the scale is developed under the guidance of TCM theory.
4. Scales should be self-reported by patients for measuring their disease condition or the QOL.

Exclusion Criteria
1. Clinical studies where scales are only used as measurement instruments.
2. Scales for diagnosing TCM syndromes or constitution.
3. Scales whose any measurement property has not been validated.
4. Scales mainly targeting pediatric patients.
5. Repeated papers or conference papers.

Data Extraction

Two researchers independently completed the literature screening according to the inclusion and exclusion criteria. The areas of disagreement shall be mutually agreed upon by both parties through consultation, and if consensus cannot be reached, it shall be arbitrated by a third investigator. Data extraction included basic information including study title, first author, study time, study site, target population, and research method, and basic information about the scales, name, type, domain, number of items, rating method, recall period, and completion time.

Evaluation Process

Evaluating the Methodological Quality of the Studies

The methodological quality of the studies was evaluated by using the Risk of Bias checklist of COSMIN.[17] Each study about a certain measurement property of a scale should be evaluated separately. Although this checklist contains nine measurement properties, the researchers only need to complete the evaluation of the studies on measurement properties involved in the article, rather than the whole checklist.[18] The checklist evaluates the bias risk of the studies using a 4-point rating method ("very good", "adequate", "doubtful", and "inadequate"). To assess the overall quality of a study, the lowest rating of any part in the checklist is taken ("the worst score counts" principle).

Evaluating the Measurement Properties of the Scales

The content validity of the scales was evaluated through the users' manual for guidance about how to assess the content validity of PRO measures and the 10 criteria for good content validity.[19,20] It was rated as "sufficient", "insufficient", "inconsistent", or "indeterminate." Other measurement properties were determined as "sufficient", "insufficient", or "indeterminate" through the updated criteria for good measurement properties.[21]

Grading the Quality of Evidence and Formulating Recommendations

The evidence's overall quality was given a grade (high, moderate, low, and very low) by using a modified GRADE approach.[21] It assumed that the evidence has a high level of quality and downgrades it in four aspects, namely risk of bias, inconsistency, imprecision, and indirectness.

According to the measurement properties and the overall quality of the evidence, the recommendation was formulated. It is recommended to classify the included scales into three levels: (1) Scales with any level of evidence for sufficient content validity and at least low-quality evidence for sufficient internal consistency were given a level of A; (2) Scales that were not classified as A or C were given a level of B. (3) Scales with high-quality evidence for an insufficient measurement property were given a level of C.[21,22]

Results

Search results

A total of 23,598 articles were obtained after searching Chinese and English databases, and 13 articles[23–35] were finally included after screening, of which 10 were in Chinese and three were in English. The process and results of screening the articles are shown in Figure 1.

Characteristics of the Studies

The 13 articles correspond to 13 scales, respectively. The articles were published from 2007 to 2022, with 6 articles focusing on patients with chronic obstructive pulmonary disease (COPD), 3 articles focusing on patients with lung cancer, 2 articles focusing on patients with idiopathic pulmonary fibrosis (IPF), and 1 article each focusing on patients with community-acquired pneumonia (CAP) and patients with bronchiectasis. All studies were conducted in China. The minimum sample size for these studies was 34, and the maximum was 366. Nine studies[25–27,30–35] reported the average age of the patients, ranging from 53.42 to 68.99. The basic information about the studies is displayed in Table 1.
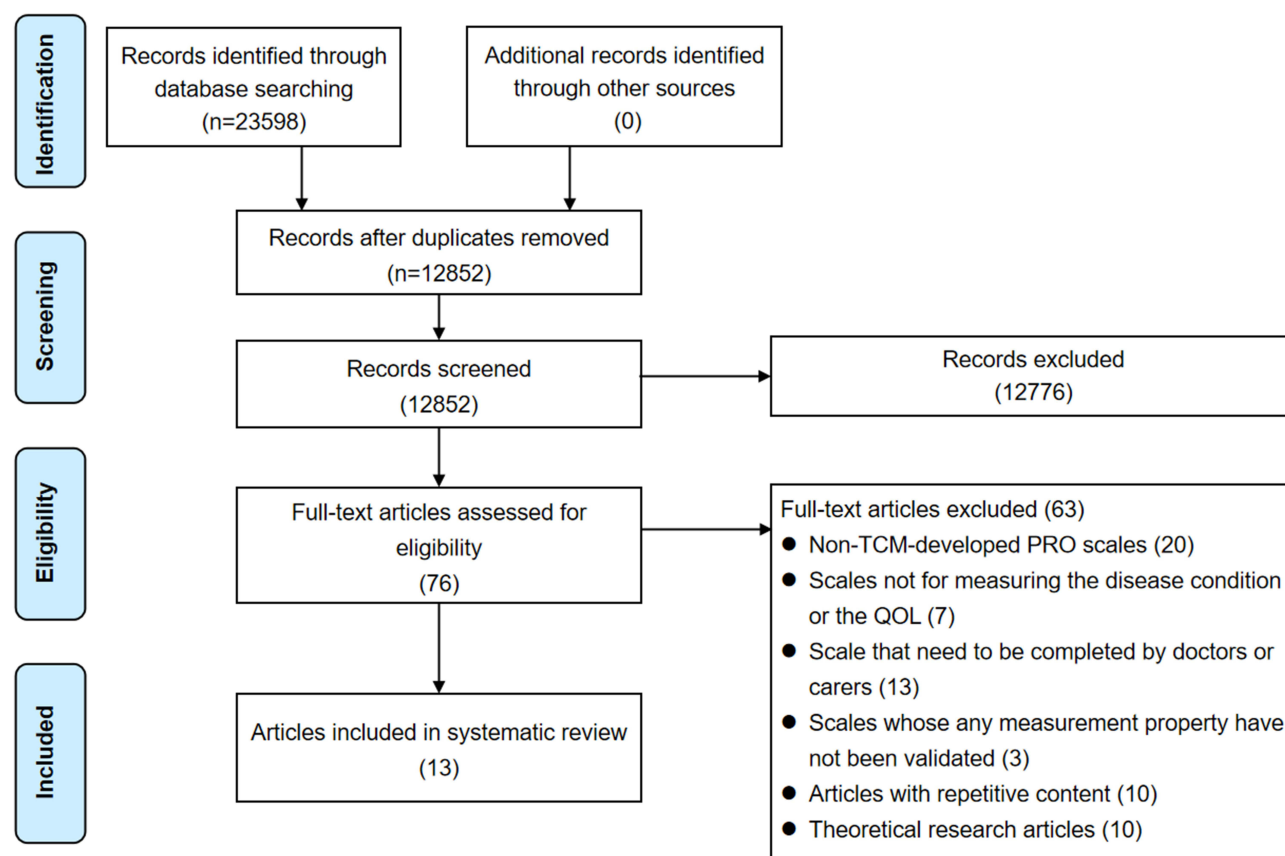
Figure 1 Literature screening flow chart.

Content of the Scales

Totally, 13 scales are all disease-specific scales (6 for COPD, 3 for lung cancer, 2 for IPF, 1 for CAP, and 1 for bronchiectasis). The characteristics of these scales are shown in Table 2.

There are 6 scales targeting the population of COPD among the scales included in this study. The self-reported scale for patients with chronic obstructive pulmonary disease in TCM (TCMPRO-COPD) combined TCM theory with the concept of QOL in its development and is mainly used to evaluate the QOL of COPD patients. There are two TCM syndrome domains in its conceptual framework, namely "syndrome elements" and "seven modes of emotions of TCM". The conceptual framework of the patient-reported outcome instrument for chronic obstructive pulmonary disease with characteristics of TCM (PRO-COPD) is in line with the concept of health made by the World Health Organization (WHO), but it also contains a TCM syndrome domain, which includes items on sleep, diet, sweating, stool, urine, and other TCM characteristics. In addition, it contains items for an overall assessment of health status and treatment satisfaction, which provides a more comprehensive report on the disease status of COPD patients. The cold and fluid syndrome-COPD-patient report outcome (CFS-COPD-PRO) contains two domains, both of which are TCM-related, and was used to evaluate COPD patients who fit the "cold and fluid syndrome" of TCM, while the TCMPRO-COPD and PRO-COPD specifically set up the domains of TCM in addition to disease-specific domains, which is an important manifestation of the combination of disease and syndrome and is also a difference from the Western measures.

The PRO instrument for chronic obstructive pulmonary diseases (COPD-PRO) was entirely based on the physiological and pathological foundations of TCM to construct the conceptual framework, such as "the lung is the master of Qi and respiration" and "the lung dominates the body fluid flow, unclogging and regulating the channels." Therefore, its measurement content mainly focuses on clinical symptoms and health status, and it is mainly used to assess the clinical efficacy of TCM in stable COPD patients. The modified PRO scale for chronic obstructive pulmonary disease (mCOPD-PRO) is a further revision

Table 1 Characteristics of the Included Studies

Reference	Instrument	Population	Disease Stage	Sample Size (% Female)	Age	Country	Number of Locations	Completion Rate
Ren et al, 2011[23]	TCMPRO-COPD	COPD	AECOPD	210[a]	18–70[b]	China	NR	100%
Zhou et al, 2016[24]	PRO-COPD	COPD	Stable COPD	60[a]	18–80[b]	China	1	100%
Li et al, 2011[25]	COPD-PRO	COPD	Stable COPD	215 (30.0)	63.9 ± 9.2[c]	China	1	93.48%
Li et al, 2020[26]	mcopd-PRO	COPD	Stable COPD	366 (23.77)	66.0	China	6	99.5%
Zhu et al, 2011[27]	scopd-PRO	COPD	Stable COPD	319 (28.21)	62.78 ± 9.42[c]	China	6	98.46%
Wei et al, 2021[28]	CFS-COPD-PRO	COPD	Cold and fluid syndrome (TCM syndrome)	300[a]	≥30[b]	China	1	92.31%
Zang et al, 2016[29]	IPF-TQ32	IPF	—	50[a]	30–70[b]	China	1	NR
Liang et al, 2016[30]	QOL-IPF	IPF	—	34 (38.24)	61 ± 14[c]	China	1	NR
Li et al, 2016[31]	CAP-PRO	CAP	—	240[a]	68.99 ± 11.48[c]	China	4	98.3%
Diao et al, 2007[32]	QOL-AL	Lung cancer	Advanced Lung cancer	100 (51.00)	63.89 ± 9.95[c]	China	1	NR
Yang et al, 2011[33]	QLASTCM-Lu	Lung cancer	—	240 (65)	60.3 ± 10.2[c]	China	NR	NR
Wang et al, 2019[34]	QLASTCM-Lu (modified)	Lung cancer	—	100 (45)	61.2 ± 10.1[c]	China	5	92%
Guan et al, 2022[35]	BE-PRO	Bronchiectasis	—	160 (61.3)	53.42 ± 13.4[c]	China	6	100%

Notes: [a]Number of males or females in the sample not reported; [b]Only the age range was reported; [c]Mean ± standard deviation; NR, Not reported.
Abbreviations: TCMPRO-COPD, Self-reported scale for patients with chronic obstructive pulmonary disease in TCM; PRO-COPD, PRO instrument for chronic obstructive pulmonary disease with characters of TCM; COPD-PRO, PRO instrument for chronic obstructive pulmonary diseases; mCOPD-PRO, Modified PRO scale for chronic obstructive pulmonary disease; sCOPD-PRO, PRO scale for patients with stable chronic obstructive pulmonary disease; CFS-COPD-PRO, Cold and fluid syndrome-COPD-PRO; IPF-TQ32, Health-related QOL of TCM scale for patients with idiopathic pulmonary fibrosis; QOL-IPF, QOL scale for patients with idiopathic pulmonary fibrosis; CAP-PRO, PRO instrument for community-acquired pneumonia; QOL-AL, QOL scale in patients with advanced lung cancer; QLASTCM-Lu, QOL assessment instrument for lung cancer patients based on TCM; QLASTCM-Lu (modified), Modified version of QOL assessment instrument for lung cancer patients based on TCM; BE-PRO, the scale combination of disease and syndrome of PRO with bronchiectasis; COPD, Chronic obstructive pulmonary disease; AECOPD, Acute exacerbation of chronic obstructive pulmonary disease; TCM, Traditional Chinese medicine; IPF, Idiopathic pulmonary fibrosis; CAP, Community-acquired pneumonia.

based on the COPD-PRO, which adds some items to make the content more comprehensive. The patient-reported outcome scale for patients with stable chronic obstructive pulmonary disease (sCOPD-PRO) has the same measurement purpose as the COPD-PRO, but its development is based on the concept of "soma and spirit harmonization." The measurement content of lung and kidney deficiency symptoms, spleen deficiency symptoms, and functional activities corresponded to "soma", while emotional impact corresponded to "spirit." The COPD-PRO, mCOPD-PRO, and sCOPD-PRO are different from TCMPRO-COPD and PRO-COPD in that they are not simply adding a TCM syndrome domain or certain TCM characteristic symptoms to the established physiological or psychological domains. The construction of their conceptual framework was entirely guided by TCM theory, which can fundamentally explain the measurement concept, as the construction of the conceptual framework is the first step in the development of PRO measures.

Three scales targeting lung cancer were included. The QOL scale in patients with advanced lung cancer (QOL-AL) covers a wide range of domains, involving physiological status, social and family status, emotional status, functional status, and other related issues. It can not only reflect the common aspects of QOL in patients with advanced lung cancer but also some typical symptoms that patients are concerned about in the treatment process of TCM. The QOL assessment instrument for lung cancer patients based on TCM (QLASTCM-Lu) combined common modules for various cancers with specific modules for lung cancer. The modified version of the QOL assessment instrument for lung cancer patients based on TCM (QLASTCM-Lu (modified)) concretized the conceptual framework, including five domains: physical status, mental condition, social function, unity of man and nature, and symptoms. Both the QLASTCM-Lu and the QLASTCM-Lu (modified) could be used to measure the QOL of lung cancer patients.

Table 2 Characteristics of the Included Scales

Instrument	Domain	Number of Items	Response Options	Scoring Direction	Scoring Algorithm	Recall Period	Completion Time
TCMPRO-COPD	5 domains: disease characters, syndrome elements, seven modes of emotions of TCM, limitation of activity and social activities	27	NR	W–B	NR	NR	8.85 min [a]
PRO-COPD	5 domains: TCM syndrome, physical domain, mental domain, individual life domain, remedial satisfaction	35	5-point (1–5)	(W–B) R	NR	AECOPD: 3 days; Stable COPD: 1 week	14.48 ± 5.26 min [c]
COPD-PRO	3 domains: amelioration of clinical symptom, satisfaction of health conditions, satisfaction of treatment effect	17	5-point (0–4)	B–W	Sum score (range, 0–68)	NR	6.7 ± 1.9 min [c]
mcopd-PRO	3 domains: physiological domain, psychological domain and environmental domain	27	5-point (0–4)	(B–W) R	Average score of items (range, 0–4)	2 weeks	5 min (4–11 min) [b]
scopd-PRO	4 domains: lung and kidney deficiency symptom, spleen deficiency symptom, functional activities and emotional impact	23	5-point (1–5)	B–W	Proportional score (range, 0–100)	NR	13.55 ± 6.32 min [c]
CFS-COPD-PRO	2 domains: external cold pattern and phlegm-drink pattern	27	5-point (0–4)	B–W	NA	NR	9.7 min [a]
IPF-TQ32	4 domains: symptoms of integration of TCM and western medicine, physiological function, mental function and social adaptability	32	5-point (1–5)	(W–B) R	Sum score (range, 32–160)	NR	NR
QOL-IPF	3 domains: clinical symptom, activity ability, influence of diseases on daily life	38	2-point (0–1); 3-point (0–2); 4-point (0–3); 5-point (0–4)	B–W	Proportional score (range, 0–100)	3 months	NR
CAP-PRO	3 domains: symptom, health satisfaction and efficacy satisfaction	17	5-point (1–5)	B–W	Sum score (range, 17–85)	3 days	9.85 ± 1.32 min [c]
QOL - AL	5 domains: physiological status, social and family status, emotional status, functional status, and other related issues	37	4-point	(W–B) R	Proportional score (range, 0–100)	NR	NR
QLASTCM-Lu	3 domains: correspondence between nature and man, integration of body and spirit, specific module	46	5-point	(B–W) R	Proportional score (range, 0–100)	NR	NR
QLASTCM-Lu (modified)	5 domains: physical status, mental condition, social function, unity of man and nature, and symptoms	49	5-point	(W–B) R	Sum score	NR	NR
BE-PRO	5 domains: physiology, psychology, environment, society, and satisfaction.	16	5-point (0–4)	B–W	Sum score (0–64)	2 weeks	9.07 ± 4.35 min [c]

Notes: W–B, As the scale score increases, the better the level of health; B–W, As the scale score increases, the health level deteriorates; R, The scale contains reverse scoring items; [a]Standard deviation not reported; [b]Median (Interquartile Range, IQR); [c]Mean ± standard deviation; NR, Not reported.
Abbreviations: TCMPRO-COPD, Self-reported scale for patients with chronic obstructive pulmonary disease in TCM; PRO-COPD, PRO instrument for chronic obstructive pulmonary disease with characters of TCM; COPD-PRO, PRO instrument for chronic obstructive pulmonary diseases; mCOPD-PRO, Modified PRO scale for chronic obstructive pulmonary disease; sCOPD-PRO, PRO scale for patients with stable chronic obstructive pulmonary disease; CFS-COPD-PRO, Cold and fluid syndrome-COPD-PRO; IPF-TQ32, Health-related QOL of TCM scale for patients with idiopathic pulmonary fibrosis; QOL-IPF, QOL scale for patients with idiopathic pulmonary fibrosis; CAP-PRO, PRO instrument for community-acquired pneumonia; QOL-AL, QOL scale in patients with advanced lung cancer; QLASTCM-Lu, QOL assessment instrument for lung cancer patients based on TCM; QLASTCM-Lu (modified), Modified version of QOL assessment instrument for lung cancer patients based on TCM; BE-PRO, the scale combination of disease and syndrome of PRO with bronchiectasis; COPD, Chronic obstructive pulmonary disease; AECOPD, Acute exacerbation of chronic obstructive pulmonary disease.

There are 2 scales used to measure the IPF patients in this study. The health-related quality of life of TCM scale for patients with idiopathic pulmonary fibrosis (IPF-TQ32) is similar to the TCMPRO-COPD and PRO-COPD, because it has a special domain for symptoms of integration of TCM and western medicine, which contains characteristic symptoms of TCM, such as "How is the sleep quality" and "whether vitality is strong or not." The QOL scale for patients with idiopathic pulmonary fibrosis (QOL-IPF) could fully reflect the connotation of IPF patients' QOL, including clinical symptoms, activity ability, and the influence of diseases on daily life. The QOL-IPF was not specifically designed for the domains of TCM syndrome, instead of integrating TCM characteristic symptoms into the domains of clinical symptoms and activity ability.

In addition, there is 1 scale for CAP and 1 scale for bronchiectasis in this study. The PRO instrument for community-acquired pneumonia (CAP-PRO) includes three domains of symptoms, health satisfaction, and efficacy satisfaction, mainly focusing on patients' satisfaction with the disease's health status and treatment effectiveness. The CAP-PRO also integrated TCM characteristic symptoms into the established domain. For example, the TCM characteristic items in the CAP-PRO, such as the impacts of weather on diseases and the impact of emotions on diseases, are sensitive to the evaluation of TCM efficacy. The scale combination of disease and syndrome of PRO with bronchiectasis (BE-PRO) can be used to assess the efficacy of TCM in treating bronchiectasis, covering five domains: physiology, psychology, environment, society, and satisfaction.

Methodological Quality and Measurement Properties

According to the measurement properties that have been reported for these scales, the content validity, structural validity, internal consistency, reliability, criterion validity, and responsiveness of the scales were evaluated. The validation studies of the measurement properties for each scale were presented in the same article and were all single. The evaluation results of the studies' methodological quality are illustrated in Table 3. Meanwhile, the evaluation results of the scales'

Table 3 Methodological Quality of the Included Studies

Instrument	Content Validity			Structural Validity	Internal Consistency	Reliability	Criterion Validity	Responsiveness
	Relevance	Comprehensiveness	Comprehensibility					
TCMPRO-COPD	D[a,b]	NR	NR	A	V	D	NR	D
PRO-COPD	D[a]	NR	D[b]	I	V	NR	I	V
COPD-PRO	D[b]	NR	NR	A	V	D	NR	D
mCOPD-PRO	D[a,b]	NR	D[b]	V	V	NR	D	V
sCOPD-PRO23	NR	NR	D[b]	A	V	NR	V	V
CFS-COPD-PRO	NR	NR	NR	A	V	D	NR	D
IPF-TQ32	D[a,b]	NR	NR	NR	D	NR	D	NR
IPF-PRO	D[a,b]	NR	NR	I	I	NR	V	NR
CAP-PRO	D[b]	NR	NR	A	V	NR	NR	D
QOL - AL	D[b]	NR	NR	I	V	NR	V	V
QLASTCM-Lu	D[a]	NR	D[b]	A	V	D	I	D
QLASTCM-Lu (modified)	NR	NR	NR	I	V	NR	I	D
BE-PRO	D[a]	NR	NR	V	V	NR	V	D

Notes: V, very good; A, adequate; D, doubtful; I, inadequate; [a]consulting experts; [b]consulting patients; NR, Not reported.
Abbreviations: TCMPRO-COPD, Self-reported scale for patients with chronic obstructive pulmonary disease in TCM; PRO-COPD, Patient-reported outcome instrument for chronic obstructive pulmonary disease with characters of TCM; COPD-PRO, Patient-reported outcome instrument for chronic obstructive pulmonary diseases; mCOPD-PRO, Modified patient-reported outcome scale for chronic obstructive pulmonary disease; sCOPD-PRO, Patient-reported outcome scale for patients with stable chronic obstructive pulmonary disease; CFS-COPD-PRO, Cold and fluid syndrome-COPD-patient report outcome; IPF-TQ32, Health-related quality of life of traditional Chinese medicine scale for patients with idiopathic pulmonary fibrosis; QOL-IPF, Quality of life scale for patients with idiopathic pulmonary fibrosis; CAP-PRO, Patient-reported outcome instrument for community-acquired pneumonia; QOL-AL, Quality of life scale in patients with advanced lung cancer; QLASTCM-Lu, Quality of life assessment instrument for lung cancer patients based on traditional Chinese medicine; QLASTCM-Lu (modified), Modified version of quality of life assessment instrument for lung cancer patients based on traditional Chinese medicine; BE-PRO, the scale combination of disease and syndrome of patient-reported outcomes with bronchiectasis.

measurement properties are demonstrated in Table 4. A summary of the validation status for the scales is presented in supplementary materials: Table S2.

Content Validity

A total of 11 studies[23–27,29–32,34,35] reported content validity: 7 studies[23,24,26,29,30,33,35] conducted the expert evaluation and 6 studies[23,24,26,29,30,33] conducted the expert and patient evaluation. There are more studies on the comprehensibility and relevance of the items by patients or experts, but no studies on the comprehensiveness of the items. For the relevance evaluations, 7 studies[25,26,29–32,35] only used quantitative surveys to evaluate the content validity, of which 3[25,31,32] were patient surveys, 1[35] was an expert survey, and 3[26,29,30] were surveyed patients and experts. In addition, 2 studies[24,33] used qualitative studies with expert consultation, and 1 study[23] used a quantitative patient survey and qualitative expert consultation. For comprehensibility evaluations, 4 studies[24,26,27,33] were performed, and all were assessed through qualitative patient consultation. However, the research process and statistical methods in the quantitative or qualitative studies were not described in detail. Consequently, the methodological quality of all 11 content validity studies was doubtful. According to the criteria of content validity, 11 scales were evaluated as either insufficient or indeterminate.

Structural Validity

Totally, 12 studies[23–28,30–35] were conducted to verify the structural validity. The structural validity studies of the mCOPD-PRO[26] and BE-PRO[35] were given a very good rating for methodological quality because confirmatory factor analysis (CFA) was employed with a sufficient sample size, and they performed well in terms of structural validity due to the reported comparative fit index (CFI) that met the criteria. Ten studies[23–25,27,28,30–34] used exploratory factor analysis (EFA), and 6 of them[23–25,27,28,31,33] had adequate sample sizes and no other methodological flaws, so these 6 studies were rated as adequate for methodological quality. Meanwhile, the methodological quality of the other 4[24,30,32,34] was all rated as inadequate because their sample size did not meet the requirement of at least five times the number of items. According to the criteria of structural validity, the 10 scales were evaluated as indeterminate.

Internal Consistency

Thirteen studies[23–35] were all conducted to evaluate the internal consistency. As the structural validity of the instrument was not tested to determine whether it met the requirement of one-dimensionality in the internal consistency study of the IPF-TQ32,[29] the study's methodological quality was assessed as doubtful. The internal consistency study of the IPF-PRO[30] did not report Cronbach's alpha coefficients for each domain, and thus the study's methodological quality was inadequate. The remaining 11 studies[23–28,31–35] reported the results on the one-dimensionality for the scales and the Cronbach's alpha coefficients of each domain, thus resulting in their satisfactory methodological quality. Among the 13 scales, the internal consistency of TCMPRO-COPD[23] was evaluated as insufficient because the Cronbach's alpha coefficients of some domains of them were less than 0.7. The IPF-PRO[30] has not yet been reported with the internal consistency coefficients of each domain, so its internal consistency was indeterminate, while the other 11 scales were evaluated as sufficient in terms of internal consistency.

Reliability

Only 4 studies[23,25,28,33] evaluated the test–retest reliability, but none of them indicated whether the composition of patients in the two surveys was stable, whether the setting and mode of the surveys were similar, or whether the interval time between surveys was appropriate, so these studies' methodological quality was all rated as doubtful. The intraclass correlation coefficient (ICC) between two measurements was not provided in the study of the TCMPRO-COPD,[23] and then, its reliability was evaluated as indeterminate. In the studies of the COPD-PRO,[25] CFS-COPD-PRO[28] and QLASTCM-LU,[33] the ICC values greater than 0.70 indicated that the reliability of the 2 scales was sufficient.

Criterion Validity

Nine studies[24,26,27,29,30,32–35] reported the criterion validity. The universal scale was used as the gold standard rather than specific scale in the study of the PRO-COPD,[24] which did not meet the COSMIN guidelines, so the methodological quality of its study was rated as inadequate, and the criterion validity of the PRO-COPD was indeterminate. The mCOPD-PRO[26]

Table 4 The Measurement Properties, Quality of Evidence, and Recommendation of the Scales

PROM	Content Validity		Structural Validity		Internal Consistency		Reliability		Criterion Validity		Responsiveness		Recommendation
	M	E	M	E	M	E	M	E	M	E	M	E	
TCMPRO-COPD	–	Low	?	Moderate	–	Moderate	?	Low	NR	NR	?	Low	B
PRO-COPD	?	Moderate	?	Moderate	+	Moderate	NR	NR	?	Low	?	Moderate	B
COPD-PRO	–	Moderate	?	High	+	High	+	Moderate	NR	NR	?	Moderate	B
mCOPD-PRO	?	Moderate	+	High	+	High	NR	NR	+	Moderate	?	High	B
sCOPD-PRO23	–	Moderate	?	High	+	High	NR	NR	+	High	?	High	B
CFS-COPD-PRO	NR	NR	?	Moderate	+	Moderate	+	Moderate	NR	NR	?	Low	B
IPF-TQ32	–	Moderate	NR	NR	+	Moderate	NR	NR	+	Low	NR	NR	B
IPF-PRO	–	Moderate	?	Moderate	?	Very low	NR	NR	+	Low	NR	NR	B
CAP-PRO	–	Moderate	?	High	+	High	NR	NR	NR	NR	?	Moderate	B
QOL - AL	–	Moderate	?	Moderate	+	High	NR	NR	+	High	?	Moderate	B
QLASTCM-Lu	?	Moderate	?	High	+	High	+	Moderate	?	Moderate	?	Moderate	B
QLASTCM-Lu (modified)	NR	NR	?	Moderate	+	Moderate	NR	NR	+	Low	?	Low	B
BE-PRO	?	Moderate	+	High	+	High	NR	NR	+	High	?	Moderate	B

Notes: M, Rating the measurement properties; E, Level of evidence; +, sufficient; –, insufficient; ?, indeterminate; B, the level of recommendation is B; C, the level of recommendation is C; NR, Not reported.
Abbreviations: TCMPRO-COPD, Self-reported scale for patients with chronic obstructive pulmonary disease in traditional Chinese medicine; PRO-COPD, Patient-reported outcome instrument for chronic obstructive pulmonary disease with characters of TCM; COPD-PRO, Patient-reported outcome instrument for chronic obstructive pulmonary diseases; mCOPD-PRO, Modified patient-reported outcome scale for chronic obstructive pulmonary disease; sCOPD-PRO, Patient-reported outcome scale for patients with stable chronic obstructive pulmonary disease; CFS-COPD-PRO, Cold and fluid syndrome-COPD-patient report outcome; IPF-TQ32, Health-related quality of life of traditional Chinese medicine scale for patients with idiopathic pulmonary fibrosis; QOL-IPF, Quality of life scale for patients with idiopathic pulmonary fibrosis; CAP-PRO, Patient-reported outcome instrument for community-acquired pneumonia; QOL-AL, Quality of life scale in patients with advanced lung cancer; QLASTCM-Lu, Quality of life assessment instrument for lung cancer patients based on traditional Chinese medicine; QLASTCM-Lu (modified), Modified version of quality of life assessment instrument for lung cancer patients based on traditional Chinese medicine; BE-PRO, the scale combination of disease and syndrome of patient-reported outcomes with bronchiectasis.

and IPF-TQ32[29] have significant differences in length compared with their respective gold standards. Due to the possibility of bias when comparing the long and short versions of a questionnaire, the 2 studies' methodological quality was doubtful. The sCOPD-PRO23,[27] IPF-PRO,[30] QOL-AL,[32] QLASTCM-Lu,[33] and QLASTCM-Lu (modified)[34] showed very good methodological quality in their studies. Except for the QLASTCM-Lu that did not report its correlation coefficient with the gold standard and its criterion validity was evaluated as indeterminate, the correlation coefficients between the other 6 scales and their gold standard were all greater than 0.70, so their criterion validity was sufficient.

Responsiveness

Eleven studies[23–28,31–35] reported responsiveness, among which 3 studies[23,26,28] of them compared the changes of different subgroups, 7 studies[25,27,31–35] compared the changes before and after intervention, and 1 study[24] included both subgroup comparisons and pre- and post-intervention comparisons. Two studies[23,28] did not describe the important characteristics of the subgroups in detail, and 5 studies[25,31,33–35] did not describe the interventions in detail, so these 7 studies' methodological quality was doubtful, while the methodological quality of the remaining 4 studies[24,26,27,32] was very good. The responsiveness of the 11 scales was all evaluated as indeterminate because no hypotheses were defined in the studies for different subgroups or pre- and post-intervention comparisons.

Grading the Quality of the Evidence and Formulating Recommendations

Since each measurement property study of these scales is single, there is no inconsistency between different studies, and the quality of evidence for the measurement property studies was not downgraded for inconsistency. The degradation of the level of evidence is considered in the following three aspects: risk of bias, imprecision, and indirectness. The level of evidence and recommendations are shown in Table 4.

Risk of Bias

Apart from the CFS-COPD-PRO and QLASTCM-Lu (modified) that were not evaluated for content validity, the methodological quality of the remaining 11 scales[23–27,29–33,35] in content validity studies was doubtful and possibly biased, and the level of evidence for content validity of these scales was all reduced by one grade. The level of evidence for structural validity of the PRO-COPD,[24] IPF-PRO,[30] QOL-AL,[32] and QLASTCM-Lu (modified)[34] was also decided to be reduced by one grade because the sample size in their studies did not reach five times the number of items. The methodological quality for the internal consistency study of the IPF-TQ32 was doubtful, and that of the QOL-IPF was inadequate. Therefore, their evidence was at the risk of bias and was downgraded by one level. The methodological quality for 4 reliability studies[23,25,28,33] was doubtful, so the level of evidence for reliability of these 4 scales was reduced by one grade. The 5 criterion validity studies' methodological quality[24,26,29,33,34] was doubtful or inadequate, hence, the level of evidence for criterion validity of these 5 scales was reduced by one grade. The 7 responsiveness studies' methodological quality[23,25,28,31,33–35] was doubtful. As a result, the level of evidence for responsiveness of these 7 scales was also reduced by one grade.

Inaccuracy

The level of evidence for internal consistency, criterion validity, and responsiveness of the PRO-COPD[24] and QLASTCM-Lu (modified)[34] was downgraded by one level because the sample size was less than 100 when their corresponding measurement properties were evaluated. The level of evidence for criterion validity of the IPF-TQ32[29] and the level of evidence for responsiveness of the QOL-AL[32] were both reduced by one grade for the same reason. However, the level of evidence for internal consistency and criterion validity of the IPF-PRO[30] was reduced by two grades because the sample size was less than 50 in the study on evaluating the measurement properties.

Indirectness

The TCMPRO-COPD[23] can be used for patients at all stages of COPD, but the study population only included hospitalized patients with acute exacerbations of COPD. The measurement concept of the CFS-COPD-PRO[28] is the "cold and fluid" syndrome of COPD, but the study population lacked a specific diagnosis of TCM syndrome. The 2 scales had indirectness, so the level of evidence for each measurement property of them was reduced by one grade.

Recommendation

Through integrating the measurement properties of the included scales and the quality of evidence from their studies, 13 scales[23–35] were all categorized as "B" because only moderate or lower evidence proved that the content validity of them was insufficient or indeterminate.

Discussion

A total of 13 studies involving 13 scales covering COPD, IPF, CAP, lung cancer, and bronchiectasis were included in this study. The scales developed based on TCM theory in the Chinese population have unique features compared with Western PRO measures that contain some health concepts that are widely concerned by Chinese people and some characteristics of TCM, which not only reflect the efficacy of TCM but also fit the Chinese cultural background. However, the TCM PRO scales for respiratory disease have not been fully validated in terms of measurement properties at present, and the overall methodological quality should be further strengthened.

The most significant difference between these TCM scales and Western scales is that they contain some items about symptoms that reflect the changes of TCM syndrome. In fact, most symptoms are common to both TCM and Western medicine, but some are uniquely concerned by TCM, such as the item "dry mouth, bitter mouth" in the PRO-COPD and the item "fear of cold or heat" in the IPF-TQ32. Moreover, items such as "appetite", "sleep", and "stool and urine" were frequently mentioned in TCM scales, but they rarely appear in those Western scales, such as SGRQ.[10] Appetite is emphasized, probably because the Chinese dining culture arose from the fact that China is a traditional agricultural country with a large population, where food production and consumption are regarded as the foundation of life by ordinary Chinese people.[36,37] Similarly, stool and urine are the dross of the human body. If the body's transmission is malfunctioning and leads to abnormal excretion of stool and urine, it will have a serious impact on the body's physical function and psychological status.[38] Therefore, the normality of stool and urine has become an important aspect of examining human QOL, which is not expressed in the concept of Western medical measures. Sleep is emphasized because it is believed that sleep can help maintain a person's vital Qi and restore one's energy in TCM theory.[39,40] Moreover, there is an ancient saying in China that goes, "Diet cures more than the doctors, sleep more than the diet cures". Healthy concepts like "weather adaptation" and "social adaptation" were also highlighted specifically in TCM scales, whereas they were addressed less in Western scales. It may be guided by the idea of "harmony between man and nature" in TCM, which is also an old Chinese saying, indicating a holistic view of life. The meaning conveyed by this phrase is that a person can achieve a dynamic and harmonious state of health by adapting to changes in natural or social environments.[41] Besides, respiratory diseases are also particularly related to the natural environment. These "Chinese-characteristic" concepts revealed potential East–West cultural differences in the definition of health.

Although these scales have TCM characteristics and are in line with Chinese cultural backgrounds, the methodological quality of their studies has some flaws. Content validity is the most significant measurement property, which refers to the degree to which the content of a PRO scale is an adequate reflection of the construct to be measured, and its evaluation content includes relevance, comprehensiveness, and comprehensibility.[18] However, none of the content validity studies for the included scales involved the comprehensiveness of the items. For the relevance evaluation, most of them were quantitative studies by experts, lacking qualitative interviews with experts. What's more, only three scales (TCMPRO-COPD,[23] PRO-COPD,[24] and QLASTCM-Lu[33]) evaluated the relevance of items in the scale using semi-structured interviews. For the comprehensibility evaluation, although qualitative interviews with patients were conducted in all studies, there are limitations in the data analysis process (recording transcription, analysis methods, and researcher qualifications). To strengthen the consistency between the items and measurement concepts, it is recommended to conduct in-depth interviews with patients or experts in the future to gain a better understanding of their perspectives and comprehension of the scale. At the same time, relevant research design and data analysis should strictly follow the COSMIN guidelines.

Internal consistency depends on the acceptability of structural validity to a certain extent.[17] Before evaluating internal consistency, it is necessary to verify the suitability of the conceptual structure of the scale which ensures that the scores of the scale can fully reflect the content of the domains contained in the scale, while internal consistency ensures the correlation between each item in the scale.[18] The methodological quality for most studies[23–28,31–35] was satisfactory in

terms of internal consistency, but only the studies of the mCOPD-PRO[26] and BE-PRO[35] were rated as sufficient in terms of structural validity, as they used confirmatory factor analysis. The COSMIN guidelines suggested that CFA is superior to EFA.[17] CFA is applicable when there is a preconstructed conceptual structure describing the relationship between the items and the measurement factors in more detail.[42]

Different from content validity, structure validity, and internal consistency, the evaluation of reliability, criterion validity, and responsiveness mainly reflect the overall performance of the scale rather than individual items. The TCMPRO-COPD,[23] COPD-PRO,[25] CFS-COPD-PRO,[28] and QLASTCM-Lu[33] in this study were reported for reliability, but none of them indicated whether the patient composition in the two surveys was stable, whether the survey environment and methods were similar, or whether the survey time interval was appropriate. For example, it set a time interval of 9 to 24 hours in the study of TCMPRO-COPD,[23] but the reason for the time interval has not been clearly explained. Therefore, the methodological quality of its study was rated as doubtful. Generally, the interval time is usually set to 2 weeks in reliability studies, and long- or short-time intervals lead to overestimation or underestimation of reliability.[43,44] During the measurement interval, the similarity between pre- and post-measurement scenarios is also very important. If the measurement scenario changes, the reliability of the instrument may be underestimated. Consequently, more attention should be paid to this issue in future research designs.

The COSMIN guidelines indicate that bias may occur when a long and short versions of a questionnaire are contrasted in the study of criterion validity.[21] For example, the COPD assessment test (CAT) score is chosen as the gold standard of the mCOPD-PRO,[26] and the number of items in the CAT score is much smaller than that in the mCOPD-PRO. In addition, as the gold standard for disease-specific scales, the universal scale is not suitable. For example, the World Health Organization Quality of Life Assessment (WHOQOL-BREF) was used as the gold standard of the PRO-COPD to evaluate its criterion validity.[24] Responsiveness can be evaluated from two angles. The first is whether the scale can distinguish the changes in the measured construct of the same group over time, and the second is whether the scale can distinguish the differences in the measured construct among different groups.[18] Responsiveness should attach importance to the combination of both.

The 13 scales have not been fully validated in terms of measurement properties at present, and they have not been reported in evaluation studies, such as measurement errors and hypothesis testing. Measurement errors include systematic errors and random errors. For quantitative data, the measurement standard error can be calculated through retesting, while for qualitative data, percentage consistency can be evaluated.[45] Hypothesis testing includes two types: One of which is the hypothesis of a relationship with other measurement tools, and another is the assumption of differences between different subgroups.[46] Additionally, the analysis of item functional differences, ceiling-to-floor effects, and the minimum clinical significance difference (MCID) of the scale are also important aspects of scale evaluation. Therefore, it is necessary to further strengthen the attention and interpretation of relevant international standards and norms to improve and supplement the validation of the measurement properties of the scales. On the one hand, on-site investigations with large samples and multiple centers should be actively conducted to promote item stability. On the other hand, the reporting standards of PRO measures should strictly referenced to increase research transparency and clarify the risk of bias in the study.[47]

According to the modified GRADE approach, the recommended levels for these 13 scales were the same, all of which were B-level, mainly because they do not have any level of evidence to indicate that their content validity is sufficient. However, based on existing evidence, a qualitative recommendation can also be made. Among the 6 scales targeting COPD, the measurement property evaluation of the mCOPD-PRO is relatively comprehensive, without insufficient measurement properties, and they can be applied to clinical studies targeting COPD. The measurement property evaluation of the IPF-TQ32 among the 2 scales targeting IPF is relatively good, with sufficient internal consistency and criterion validity. The QLASTCM-Lu may be recommended for clinical studies on lung cancer, as it does not have insufficient measurement properties. In addition, there is 1 scale for CAP and 1 scale for bronchiectasis, and the BE-PRO is relatively good.

There are some limitations to consider in this work. Like any other review, the results obtained are also influenced by the biases originating from the studies included in the review. However, the methodological quality evaluation carried out provides a valuable means to determine the feasibility and reliability of the scales. Furthermore, other (content-specific)

databases were not searched, such as CINAHL or PsycINFO. However, PubMed and Embase were searched in this study, which are considered the minimum databases that need to be searched. One limitation worth mentioning in this study is that there was only one article found for each PRO scale, which might negatively impact the evaluation of the measurement properties and quality of evidence for the PRO scales included, thus undermining the generalizability of our findings. Finally, although the systematic review of the included articles was conducted strictly in accordance with the COSMIN guidelines, the analysis and evaluation processes are subjective in part, so the authors guaranteed the participation of multiple reviewers to ensure consistency and consensus.

Conclusion

The 13 scales included in this study contained some concepts that may be considered as TCM features, which not only reflect the efficacy of TCM but also fit the Chinese cultural background. However, the existing evidence suggested that the measurement properties of these scales for respiratory diseases were not comprehensive enough, and the overall methodological quality was relatively low. It is necessary to further improve and supplement the validation of the measurement properties of the scales.

Acknowledgments

We would like to express our gratitude to all the authors whose articles were included in this systematic review.

Author Contributions

All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

Funding

Disclosure

The authors declared no conflicts of interest for this work.

References

1. FDA. CDER patient-focused drug development; 2022. Available from: https://www.fda.gov/drugs/development-approval-process-drugs/cder-patient-focused-drug-development. Accessed June 2, 2023.
2. Hareendran A, Gnanasakthy A, Winnette R, Revicki D. Capturing patients' perspectives of treatment in clinical trials/drug development. *Contemp Clin Trials*. 2012;33(1):23–28. doi:10.1016/j.cct.2011.09.015
3. McColl E, Junghard O, Wiklund I, Revicki DA. Assessing symptoms in gastroesophageal reflux disease: how well do clinicians' assessments agree with those of their patients? *Am J Gastroenterol*. 2005;100(1):11–18. doi:10.1111/j.1572-0241.2005.40945.x
4. U.S. Department of Health and Human Services FDA Center for Drug Evaluation and Research, U.S. Department of Health and Human Services FDA Center for Biologics Evaluation and Research, U.S. Department of Health and Human Services FDA Center for Devices and Radiological Health. Guidance for industry: patient-reported outcome measures: use in medical product development to support labeling claims: draft guidance. *Health Qual Life Outcomes*. 2006;4(1):79. doi:10.1186/1477-7525-4-79
5. Calvert M, Kyte D, Price G, Valderas JM, Hjollund NH. Maximising the impact of patient reported outcome assessment for patients and society. *BMJ*. 2019;364:k5267. doi:10.1136/bmj.k5267
6. Guo J, Chen Y, Zhang W, Tong S, Dong J. Moderate and severe exacerbations have a significant impact on health-related quality of life, utility, and lung function in patients with chronic obstructive pulmonary disease: a meta-analysis. *Int J Surg*. 2020;78:28–35. doi:10.1016/j.ijsu.2020.04.010
7. Magge A, Ashraf S, Quittner AL, Metersky ML. Quality of life in patients with bronchiectasis: a 2-year longitudinal study. *Ann Transl Med*. 2019;7 (14):334. doi:10.21037/atm.2019.06.62

8. Cox IA, Borchers Arriagada N, de Graaff B, et al. Health-related quality of life of patients with idiopathic pulmonary fibrosis: a systematic review and meta-analysis. *Eur Respir Rev.* 2020;29(158):200154. doi:10.1183/16000617.0154-2020

9. Molen TVD, 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(1):13. doi:10.1186/1477-7525-1-13

10. 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(6):1321–1327. doi:10.1164/ajrccm/145.6.1321

11. Stewart AL, Napoles-Springer A. Health-related quality-of-life assessments in diverse population groups in the United States. *Med Care.* 2000;38(9 Suppl):II102–II124. doi:10.1097/00005650-200009002-00017

12. Zhang S, Chen Y. Research status of patient reported outcome scale of traditional Chinese medicine. *Zhong Hua Zhong Yi Yao Za Zhi.* 2018;33 (03):1026–1029.

13. Haifeng W, Hailong Z, Jiansheng L, et al. Effectiveness and safety of traditional Chinese medicine on stable chronic obstructive pulmonary disease: a systematic review and meta-analysis. *Complement Ther Med.* 2015;23(4):603–611. doi:10.1016/j.ctim.2015.06.015

14. Chen X, Kang F, Lai J, Deng X, Guo X, Liu S. Comparative effectiveness of phlegm-heat clearing Chinese medicine injections for AECOPD: a systematic review and network meta-analysis. *J Ethnopharmacol.* 2022;292:115043. doi:10.1016/j.jep.2022.115043

15. Huang SY, Cui HS, Lyu MS, Huang GR, Hou D, Yu MX. Efficacy of traditional Chinese medicine injections for treating idiopathic pulmonary fibrosis: a systematic review and network meta-analysis. *PLoS One.* 2022;17(7):e0272047. doi:10.1371/journal.pone.0272047

16. Wang Q, Wang Q, Wang SF, et al. Oral Chinese herbal medicine as maintenance treatment after chemotherapy for advanced non-small-cell lung cancer: a systematic review and meta-analysis. *Curr Oncol.* 2017;24(4):e269–e276. doi:10.3747/co.24.3561

17. Mokkink LB, de Vet HCW, Prinsen CAC, et al. COSMIN risk of bias checklist for systematic reviews of patient-reported outcome measures. *Qual Life Res.* 2018;27(5):1171–1179. doi:10.1007/s11136-017-1765-4

18. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol.* 2010;63(7):737–745. doi:10.1016/j.jclinepi.2010.02.006

19. Mokkink LB, Terwee CB, Knol DL, et al. The COSMIN checklist for evaluating the methodological quality of studies on measurement properties: a clarification of its content. *BMC Med Res Methodol.* 2010;10(1):22. doi:10.1186/1471-2288-10-22

20. Terwee CB, Prinsen CAC, Chiarotto A, et al. COSMIN methodology for evaluating the content validity of patient-reported outcome measures: a delphi study. *Qual Life Res.* 2018;27(5):1159–1170. doi:10.1007/s11136-018-1829-0

21. Prinsen CAC, Mokkink LB, Bouter LM, et al. COSMIN guideline for systematic reviews of patient-reported outcome measures. *Qual Life Res.* 2018;27(5):1147–1157. doi:10.1007/s11136-018-1798-3

22. Prinsen CA, Vohra S, Rose MR, et al. How to select outcome measurement instruments for outcomes included in a "core outcome set" - a practical guideline. *Trials.* 2016;17(1):449. doi:10.1186/s13063-016-1555-2

23. Ren M, Sun ZT, Feng JH, et al. Evaluating of self-reported scale for patients with chronic obstruction pulmonary disease in TCM. *Tianjin J Tradit Chin Med.* 2011;28(02):101–103.

24. Zhou D. *Application and Evaluation of the "Patient Reported Outcome" Traditional Chinese Medicine Evaluation Scale for Chronic Obstructive Pulmonary Disease.* Hunan: Hunan University of Chinese Medicine; 2016.

25. Li JS, Wang MH, Yu XQ, Li SY, Xie Y. Development and validation of a patient reported outcome instrument for chronic obstructive pulmonary diseases. *Chin J Integr Med.* 2015;21(9):667–675. doi:10.1007/s11655-014-1982-4

26. Li J, Wang J, Xie Y, Feng Z. Development and validation of the modified patient-reported outcome scale for Chronic Obstructive Pulmonary Disease (mCOPD-PRO). *Int J Chron Obstruct Pulmon Dis.* 2020;15:661–669. doi:10.2147/COPD.S240842

27. Zhu YB, Li YL, Wang W, Jiang FC. Development of a patient-reported outcome scale for stable chronic obstructive pulmonary disease and its clinical applicability. *J Integr Med.* 2011;9(08):857–865.

28. Wei R, You LN, Xia J, Lei L, Wu ZY, Pan X. Evaluation of reliability and validity of CFS-COPD-PRO for COPD patients of "cold and fluid" pattern. *West J Tradit Chin Med.* 2021;34(02):80–84.

29. Zang NZ, Pang LJ, Li P, et al. Establishment and evaluation of IPF-TCM-HRQOL32: a health-related quality of life of TCM scale for patients with idiopathic pulmonary fibrosis. *China J Tradit Chin Med Pharm.* 2016;31(05):1638–1642.

30. Liang LJ, Liang LR, Dai HP. Research on life quality scale for patients with idiopathic pulmonary fibrosis. *Chin J Integr Tradit West Med.* 2016;36 (06):668–673.

31. Li JS, Wang MH, Yu XQ, Xie Y, Li SY. Development and assessment of efficacy measurement instruments for community-acquired pneumonia. *China J Chin Med.* 2016;31(11):1654–1661.

32. Diao J. *Study on Quality of Life Scale in Patient with Advance Lung Cancer.* Nanjing: Nanjing University of Chinese Medicine; 2007.

33. Yang Z, Lu JG, You SF, et al. Development of quality of life assessment system for cancer based on traditional Chinese medicine-lung cancer (QLASTCM-LU). *Mod Prev Med.* 2011;38(18):3636–3639+3645.

34. Wang TT, He LY, Zhang M, et al. Development of improved version of Quality of Life Assessment Instrument for Lung Cancer Patients Based on Traditional Chinese Medicine (QLASTCM-Lu). *Chin J Integr Med.* 2019;25(11):831–836. doi:10.1007/s11655-018-2991-5

35. Guan JR. *Development of the Scale Combination of Disease and Syndrome of Patient-Reported Outcomes with Bronchiectasis Based on Modern Testing Theory.* Henan: Henan University of Chinese Medicine; 2022.

36. Zhao RG. *Chinese Food Culture.* Beijing: Higher Education Press; 2008.

37. Han G, Ji J. A discussion on Chinese "eating" culture from the greeting "Have you eaten". *Youth Literator.* 2009;2009(5):76–77.

38. Hou ZK. On the necessity of developing quality of life instruments in traditional Chinese medicine. *J Chin Integr Med.* 2011;9(5):468–482. doi:10.3736/jcim20110502

39. Lou CX, Pang ZR, Cui J. Practical significance of health preservation thoughts in Huangdi's orthodox classic. *Li Shizhen Medicine Medical Res.* 2009;20(11):2837–2839.

40. Zhou X, Yu CQ, Wang HW, et al. The relationship between Chinese traditional theory of essence-Qi-spirit and health. *J Tianjin Univ Trad Chinese Med.* 2013;32(1):8–11.

41. Zhang H, Zhang FY, Zhao YN, Li Z. Health concepts under the view of Chinese culture. *Chinese J Nurs.* 2015;50(10):1236–1239.

42. Paula JJD, Albuquerque MR, Bicalho MAC, Romano-Silva MA. Confirmatory factor analysis of the general activities of daily living scale: further evidences of internal validity. *Braz J Psychiatry.* 2017;39(4):379–380. doi:10.1590/1516-4446-2017-2256

43. Devellis RB. *Scale Development: Theory and Applications 3ed*. Chongqing: Chongqing University Press; 2016.
44. Moreira RS, Bassolli L, Coutinho E, et al. Reproducibility and reliability of the quality of life questionnaire in patients with atrial fibrillation. *Arq Bras Cardiol*. 2016;106(3):171–181. doi:10.5935/abc.20160026
45. Terwee CB, Mokkink LB, Knol DL, Ostelo RW, Bouter LM, de Vet HC. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. *Qual Life Res*. 2012;21(4):651–657. doi:10.1007/s11136-011-9960-1
46. Elsman EBM, Butcher NJ, Mokkink LB, et al. Study protocol for developing, piloting and disseminating the PRISMA-COSMIN guideline: a new reporting guideline for systematic reviews of outcome measurement instruments. *Syst Rev*. 2022;11(1):121. doi:10.1186/s13643-022-01994-5
47. Gagnier JJ, Lai J, Mokkink LB, Terwee CB. COSMIN reporting guideline for studies on measurement properties of patient-reported outcome measures. *Qual Life Res*. 2021;30(8):2197–2218. doi:10.1007/s11136-021-02822-4