Medicine®

SYSTEMATIC REVIEW AND META-ANALYSIS

OPEN

Functional Status Assessment of Patients With COPD

A Systematic Review of Performance-Based Measures and Patient-Reported Measures

Yang Liu, PhD, Honghe Li, PhD, Ning Ding, PhD, Ningning Wang, PhD, and Deliang Wen, PhD

Abstract: Presently, there is no recommendation on how to assess functional status of chronic obstructive pulmonary disease (COPD) patients. This study aimed to summarize and systematically evaluate these measures.

Studies on measures of COPD patients' functional status published before the end of January 2015 were included using a search filters in PubMed and Web of Science, screening reference lists of all included studies, and cross-checking against some relevant reviews. After title, abstract, and main text screening, the remaining was appraised using the Consensus-based Standards for the Selection of Health Measurement Instruments (COSMIN) 4-point checklist. All measures from these studies were rated according to *best-evidence synthesis* and the best-rated measures were selected.

A total of 6447 records were found and 102 studies were reviewed, suggesting 44 performance-based measures and 14 patient-reported measures. The majority of the studies focused on *internal consistency*, *reliability,* and *hypothesis testing*, but only 21% of them employed *good or excellent* methodology. Their common weaknesses include *lack of checks for unidimensionality*, *inadequate sample sizes*, *no prior hypotheses*, and improper methods. On average, patient-reported measures perform better than performance-based measures. The best-rated patient-reported measures are functional performance inventory (FPI), functional performance inventory short form (FPI-SF), living with COPD questionnaire (LCOPD), COPD activity rating scale (CARS), University of Cincinnati dyspnea questionnaire (UCDQ), shortness of breath with daily activities (SOBDA), and short-form pulmonary functional status scale (PFSS-11), and the best-rated performance-based measures are exercise testing: 6-minute walk test (6MWT), endurance treadmill test, and usual 4-meter gait speed (usual 4MGS).

Further research is needed to evaluate the reliability and validity of performance-based measures since present studies failed to provide convincing evidence. FPI, FPI-SF, LCOPD, CARS, UCDQ, SOBDA, PFSS-11, 6MWT, endurance treadmill test, and usual 4MGS performed well and are preferable to assess functional status of COPD patients.

(*Medicine* 95(20):e3672)

Abbreviations: 10MGS = maximal 10-meter gait speed, 10MGS = usual 10-meter gait speed, 12MD = 12-minute distance walk, 2MWT = 2-minute walk test, 30MWT = 30-meter walk test, 3CRT = 3-minute chair rise test, 4MGS = 4-meter gait speed, 4MGS = 4-meter gait speed, 5STS = five-repetition sit-to-stand test, 6MST = 6-minute step test, 6MWT = 6-minute walk test, ADL-D = activity of daily living dyspnea scale, ADLs = activities of daily living, AH = actiheart, AUC = area under the curve, CARS = COPD activity rating scale, CAT = COPD assessment test, CDLM = capacity of daily living during the morning questionnaire, COPD = chronic obstructive pulmonary disease, COSMIN = consensus-based standards for the selection of health measurement instruments, DAM = DynaPort activity monitor, DASI = Duke activity status index, DIF = differential item functioning, DIRECT = DIsability RElated to COPD Tool, ESWT = endurance shuttle walking test, FPI = functional performance inventory, FPI-SF = functional performance inventory short form, GST = grocery shelving task, HRQOL = health-related quality of life, ICC = intraclass correlation coefficient, ISWT/SWT = incremental shuttle walk test, LCADL = London chest activity of daily living scale, LCOPD = living with COPD questionnaire, LoA = limits of agreement, M6MWT = modified 6-minute walk test, MIC = minimal important change, MRADL = Manchester respiratory activities of daily living questionnaire, MSWT = modified SWT, PBRT = 6-minute pegboard and ring test, PFSDQ-M = pulmonary functional status and dyspnea questionnaire-modified, PFSS = pulmonary functional status scale, PFSS-11 = short-form pulmonary functional status scale, PRO = patient-reported outcomes, PW = power walker 610, SAB = SenseWear armband, SAM = StepWatch activity monitor, SCAM = self-contained activity monitor, SCPT = stair climb power test, SDC = smallest detectable change, SOBDA = shortness of breath with daily activities, SRAT = steep ramp anaerobic test, STST = sit-to-stand test, TChester = Chester step test, UCDQ = University of Cincinnati dyspnea questionnaire, UULEX = unsupported upper limb exercise test.

INTRODUCTION

Chronic obstructive pulmonary disease (COPD), characterized by persistent airflow limitation, is usually progressive and associated with an enhanced chronic inflammatory response in the airways and the lung to noxious particles or gases.[1] When the disease becomes aggravated, patients suffer from deteriorated functional status and limitations to daily life. The impaired functional status is proven to be predictors of exacerbations, hospital admissions, and mortality.[2,3] The worsening functional status presents a tough challenge for patients and their families and causes an increasing burden for the

Editor: Chengwu Yang.
Received: January 7, 2016; revised: April 5, 2016; accepted: April 15, 2016.
From the School of Public Health (YL, HL, DW); The Research Centre for Medical Education (ND), China Medical University, Shenyang; and School of Public Health (NW), Dalian Medical University, Dalian, Liaoning, China.
Correspondence: Deliang Wen, China Medical University, No. 77 Puhe Road, Shenyang North New Area, Shenyang, Liaoning Province, People's Republic of China (e-mail: dlwen@cmu.edu.cn).
This study was supported by Natural Science Foundation of China (NSFC) (fund number: 71573032).
The authors have no conflicts of interest to disclose.
Supplemental Digital Content is available for this article.

ISSN: 0025-7974
DOI: 10.1097/MD.0000000000003672

society.[4] Therefore, assessing functional status accurately and systemically is one of demanding require of COPD treatments, as indicated in COPD guidelines.[1,5,6]

Measuring the type and magnitude of functional damage and evaluating treatment effect on functional improvement is a challenging work in clinical practice. Even in some large pulmonary rehabilitation programs, variables like activities of daily living (ADLs) and exercise tolerance were not adequately assessed.[7,8] Functional status contains multidimensional constructs and is often confused with other relevant constructs.[9–13] According to the Wilson-Cleary framework, functional status was broadly defined as the ability to perform particular defined tasks in multiple domains, including physical function, social function, role function, and psychological function.[14,15] Similarly, there are also many aspects of functional status in terms of intension, including functional capacity, functional performance, functional reserve, and functional capacity utilization. Correspondingly, many functional status instruments were proposed for different purposes, including performance-based measures and patient-reported measures.[12,16–19] Two limitations in present studies, undermining the development and validation of these instruments, are (1) lack of an assessment of the quality of methodology used, resulting in unconvincing conclusions of measures' development and/or validation; (2) lack of a clear-defined, systematical, and quantifiable assessment standard, resulting in partial and ambiguous judgments on measures performance.[20–22]

Consensus-based standards for the selection of health measurement instruments (COSMIN) was proposed in 2006.[20–23] Besides evaluating the quality of studies on measurement property critically, COSMIN also includes the measurement properties systematically. It has been used in many systematic reviews to evaluate studies and instruments of various diseases, such as hip and knee osteoarthritis, geriatrics, non-small cell lung carcinoma, and neuro-rehabilitation patients.[24–27] In terms of COPD, COSMIN has been used to assess COPD assessment test (CAT) questionnaire, health-related quality of life (HRQOL) questionnaires, and arm exercise capacity.[28–30] In this article, we employed COSMIN to review both patient-reported measures and performance-based measures of COPD patients' functional status.

The objectives of this review is threefold: (1) to appraise the quality of methodology in the studies on the measures of COPD patients' functional status and to provide insights for future researches, (2) to summarize all candidate instruments and to make recommendations for instrument selection, and (3) to compare performance-based measures and patients-reported measures.

METHOD

Search Strategy

We searched PubMed, Web of Science using a search filter developed by Terwee to identify studies describing development or evaluation of measurement properties of instruments measuring functional status of COPD patients up to the end of January 2015. (See text, Supplemental Digital Content 1, http://links.lww.com/MD/A969, which describes the detailed search strategy.)[31]

1# Construct search

2# Population search

3# Instrument search

4# #1 AND #2 AND #3 AND filter for measurement properties

5# #4 NOT exclusion filter

For supplement, we searched each instrument in the entire database and looked up the references of each included article. Finally, our review was cross-checked against some relevant reviews.[16,17,19,24]

Eligibility Criteria and Study Selection

By applying the inclusion and exclusion criteria (Table 1), 3 reviewers (YL, HL, and ND) independently screened titles and abstracts of the identified records and independently assessed full texts for eligibility. Discussion was conducted when there were differences concerning exclusion criteria. If consensus could not be reached, the final decision was made by the forth reviewer (NW).

Evaluation of Methodological Quality of the Included Studies

Before the evaluation of methodological quality of the included studies, descriptive variables of these studies including authors/year, country, study sample, study design, sex (female, %), mean age years ± SD (range), mean percentage of the

TABLE 1. Inclusion Criteria and Exclusion Criteria for Eligible Studies

Criteria	Inclusion	Exclusion
Criteria 1: Population	COPD patients	All others
Criteria 2: Content	Studies described the development or evaluation of the measurement properties or studies that reported at least one or more psychometric properties were included	Studies where the objective was the evaluation of an intervention or treatment without reporting any measurement properties were excluded. Studies in which the measurement instruments were used as an end point without studying the measurement properties were not considered eligible
Criteria 3: Instrument	Studies of instruments measuring functional status were included	Studies of instruments measuring HRQOL, general health perception, or only symptoms or satisfaction with care and adherence were excluded
Criteria 4: Characters of literature	English; full text; original article	Other languages; conference papers; editorials; commentaries; supplementary
Criteria 5: Journal	Peer-reviewed	All others

COPD = chronic obstructive pulmonary disease, HRQOL = health-related quality of life.

forced vital capacity (FEV1%) predicted ± SD, and patients status were collected. Then the methodological quality of included studies was evaluated according to the COSMIN 4-point checklist.[32] The COSMIN checklist consists of 9 boxes concerning methodological standards on how each measurement property should be assessed, including 5 to 18 items in each box. The overall score (i.e., poor, fair, good, or excellent) for each item was obtained by taking the lowest score for any question within the item.

Quality Assessment of Instruments

The quality of the instruments was determined according to the rating system provided by Terwee (Table 2). It contains criteria for content validity, internal consistency, criterion validity, construct validity, reproducibility (agreement and reliability), responsiveness, floor and ceiling effects, and interpretability. Each measurement property was reported by positive (+), intermediate (?), negative (−), or no information available (0). The version provided by Terwee was used in this review.

Data Synthesis and Quality Assessment

To synthesize the evidence, ''best-evidence synthesis'' was performed. As proposed by the Cochrane Back Review Group, the levels of evidence were ''strong,'' ''''moderate,'' ''limited,'' ''conflicting,'' or ''unknown'' (Table 3).[33,34] Methodological quality of the studies (COSMIN score), rating of quality assessment of instruments, consistency between different studies, and the number of studies were taken into consideration using the synthesis. We defined best rated instruments as those which had a ''+++'' (strong positive) in at least one measurement property or a ''+'' or ''++'' in at least three measurement properties according to the results of data synthesis.

Since this study merely reviewed the articles already published without involving any human participants directly, ethical approval is not necessary.

RESULTS

Electronic Literature Search Results

The selection process for all studies is shown in Figure 1. With the search filter, 6447 records were identified. After screening the title and abstract, 6225 records were excluded. The remaining 222 records were screened for full text, among which 145 records were excluded for various reasons shown in Figure 1. Twenty-five additional records were identified through screening of references lists and review articles and searching for each particular instrument in PubMed. A total of 102 articles were analyzed in the review.

Description of the Included Studies and Included Instruments

A total of 95 of the 102 studies were published after 2000. These included cross-sectional studies, longitudinal studies, and randomized double-blind studies. Mean age of subjects include in these studies ranged from 51.0 to 74.7 years. Fifty of the 102 studies declared that subjects include in their studies were stable patients. In total, 58 instruments were identified, including 44 performance-based measures and 14 patient-reported measures. The 44 performance-based measures could be divided into 28 exercise tests and 16 activity monitors. (See table, Supplemental Digital Content 2, http://links.lww.com/MD/A969, which describes the included studies.)

Quality of the Included Studies

The quality of included studies can be found in Tables 4 and 5. The methodological quality of the existing studies ranged from *poor* to *excellent*, with *good* and *excellent* collectively taking 21%.

Studies on performance-based measures

There were 89 studies that analyzed the measurement properties of performance-based measures. *Reliability* and *hypothesis testing* were the most reported measurement properties of this type of instrument (reported in 52 studies and 35 studies, respectively). *Criterion validity* and *responsiveness* were reported in 20 studies and 24 studies, respectively. Unlike the patient-reported measures, performance-based measures had some evidence of *measurement error* from 7 studies.

Of the studies reporting on *reliability*, 2 were *excellent*, 12 were *good*, 12 were *fair*, and 26 were *poor*. Inadequate sample size and *no intraclass correlation coefficient (ICC) or Pearson or Spearman correlations calculated* were the main reasons contributing to poor quality of the study. For *hypothesis testing*, 2 studies were *good*, 16 *fair*, and 17 *poor*. The main weakness lies in inadequate sample sizes and a lack of adequate hypotheses. Evaluating the *criterion validity*, most studies chose the direct observation as the gold criterion. However, 12 studies were considered *poor* in *criterion validity* because of inadequate sample sizes. Among the 20 studies reporting *responsiveness*, 1 study had good quality, 7 studies had fair quality, and 16 studies had poor quality.

Studies on patient-reported measures

Among the 32 studies analyzing the measurement properties of patient-reported measures, *internal consistency*, *reliability*, and *hypothesis testing* were the measurement properties reported most frequently (reported in 21, 20, and 21 studies, respectively), whereas no study reported on the *measurement error* of the patient-reported measures. *Cross-cultural validity* and *criterion validity* were also only reported in 6 studies and 2 studies, respectively. *Content validity* and *cross-cultural validity* were more so evaluated in the development of the scales rather than in the final version.

The qualities of the studies analyzing the *internal consistency* of patients-reported measures were as follows: 1 *excellent*, 3 *good*, 4 *fair*, and 13 *poor*. Studies were deemed *poor* mostly because of the fact that unidimensionality was not properly checked. The quality of the studies analyzing the *reliability* was 1 *excellent*, 5 *good*, 9 *fair*, and 5 *poor*. Inadequate sample size was the decisive factor of lesser quality. The quality of the studies analyzing the *hypothesis testing* was 4 good and 17 fair. The quality of most studies stopped at *fair* because of that they did not formulate any hypotheses in their studies. Studies reporting *responsiveness* did not have high quality because of inadequate sample sizes or to the fact that inappropriate methods were used. *Structural validity* was analyzed in 8 studies and the qualities were mostly determined by the sample size.[111,113–115,119,124–126]

Quality of Psychometric Properties for Outcome Measures

A summary of best-evidence synthesis is provided in Table 6. The summary was driven from the results of study qualities and the quality of psychometric properties for outcome measures (see table, Supplemental Digital Content 3, http://links.lww.com/MD/A969, which describes quality of

TABLE 2. Quality Criteria for Measurement Properties[7]

Property	Rating	Quality Criteria
Reliability		
Internal consistency		
	+	Cronbach's alpha (s) $\geq$0.70
	?	Cronbach's alpha not determined or dimensionality unknown
	−	Cronbach's alpha (s) $<$0.70
Reliability		
	+	ICC/weighted Kappa $\geq$0.70 OR Pearson $r \geq$0.80
	?	Neither ICC/weighted Kappa, nor Pearson r determined
	−	ICC/weighted Kappa $<$0.70 OR Pearson $r <$0.80
Measurement error		
	+	MIC $>$ SDC OR MIC outside the LoA
	?	MIC not defined
	−	MIC $\leq$ SDC OR MIC equals or inside LoA
Validity		
Content validity		
	+	All items are considered to be relevant for the construct to be measured, for the target population, and for the purpose of the measurement AND the questionnaire is considered to be comprehensive
	?	Not enough information available
	−	Not all items are considered to be relevant for the construct to be measured, for the target population, and for the purpose of the measurement OR the questionnaire is considered not to be comprehensive
Construct validity–Structural validity		
	+	Factors should explain at least 50% of the variance
	?	Explained variance not mentioned
	−	Factors explain $<$50% of the variance
- Hypothesis testing		
	+	Correlations with instruments measuring the same construct $\geq$0.50 OR at least 75% of the results are in accordance with the hypotheses AND correlations with related constructs are higher than with unrelated constructs
	?	Solely correlations determined with unrelated constructs
	−	Correlations with instruments measuring the same construct $<$0.50 OR $<$75% of the results are in accordance with the hypotheses OR correlations with related constructs are lower than with unrelated constructs
- Cross-cultural validity		
	+	No differences in factor structure OR no important DIF between language versions
	?	Multiple group factor analysis not applied AND DIF not assessed
	−	Differences in factor structure OR important DIF between language versions
Criterion validity		
	+	Convincing arguments that gold standard is ''gold'' AND correlation with gold standard $\geq$0.70
	?	No convincing arguments that gold standard is ''gold''
	−	Correlation with gold standard $<$0.70
Responsiveness		
Responsiveness		
	+	Correlation with changes on instruments measuring the same construct $\geq$0.50 OR at least 75% of the results are in accordance with the hypotheses OR AUC $\geq$0.70 AND correlations with changes in related constructs are higher than with unrelated constructs
	?	Solely correlations determined with unrelated constructs
	−	Correlations with changes on instruments measuring the same construct $<$0.50 OR $<$75% of the results are in accordance with the hypotheses OR AUC $<$0.70 OR correlations with changes in related constructs are lower than with unrelated constructs

? = indeterminate rating, + = positive rating, − = negative rating, AUC = area under the curve, DIF = differential item functioning, ICC = intraclass correlation coefficient, LoA = limits of agreement, MIC = minimal important change, SDC = smallest detectable change.

TABLE 3. Levels of Evidence for the Quality of the Measurement Property[34]

Level	Rating	Criteria
Strong	+++ or − − −	Consistent findings in multiple studies of good; methodological quality OR in one study of excellent; methodological quality
Moderate	++ or − −	Consistent findings in multiple studies of fair; methodological quality OR in one study of good; methodological quality
Limited	+ or −	One study of fair methodological quality
Conflicting	±	Conflicting findings
Unknown	?	Only studies of poor methodological quality

− = negative rating, OR =, + = positive rating, ? = indeterminate rating.

psychometric properties for outcome measures) using the criteria displayed in Table 3.

Patient-reported measures performed better than performance-based measures. All positive evidence of patient-reported measures were evenly distributed in both reliability and validity. Most of the positive evidence of performance-based measures was confined to exercise testing and reliability measurement property.

Best rated instruments with a ''+++'' in one measurement property or ''+''/''++'' in at least three measurement properties among performance-based measures are 6-minute walk test (6MWT), endurance treadmill test, and usual 4-meter gait speed (usual 4MGS). The best rated patients-reported measures were functional performance inventory (FPI), functional performance inventory short form (FPI-SF), living with COPD questionnaire (LCOPD), COPD activity rating scale (CARS), University of Cincinnati dyspnea questionnaire (UCDQ), shortness of breath with daily activities (SOBDA), and short-form pulmonary functional status scale (PFSS-11).

DISCUSSION

The present review provides the first evidence on comparing all candidate instruments measuring functional status in COPD patients according to the COSMIN criteria. It highlighted some areas worthy of future researched, including the lack of adequate positive evidence on measurement properties of performance-based measures compared with patient-reported measures, the weakness limiting the quality of the existing studies, and the important measurement properties neglected by existing studies. Although none of the instruments was tested for all measurement properties, the existing evidence still confirms that some instruments performed better in terms of some measurement properties or some survey types. For clinical

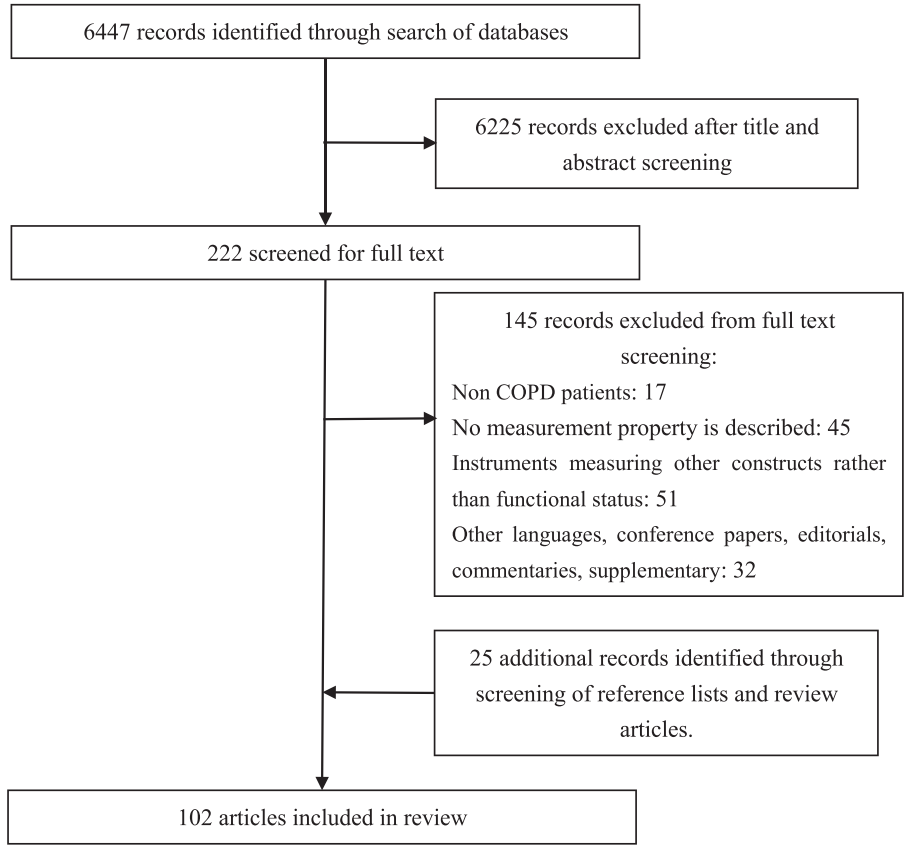

FIGURE 1. Flow diagram of search results.

TABLE 4. Results of Quality of Design, Methods, and Reporting for Included Studies on Performance-Based Measures

	Psychometric Properties								
Authors/Year	**Box A Internal Consistency**	**Box B Reliability**	**Box C Measurement Error**	**Box D Content Validity**	**Box E Structural Validity**	**Box F Hypothesis Testing**	**Box G Cross-Cultural Validity**	**Box H Criterion Validity**	**Box I Responsiveness**
2MWT									
Butland et al/1982[35]		Poor (3,11)				Poor (7,8)			
Eiser et al/2003[36]		Poor (11)							Poor (14)
Leung et al/2006[37]		Fair (3,6,8)				Fair (3,4,8)			Poor (3)
6MWT									
Eiser et al/2003[36]		Poor (11)							Poor (14)
Casanova et al/2007[38]									Poor (8,11,12,13,14)
Brown et al/2008[39]						Fair (40)			
Spencer et al/2008[40]									Fair (3,8,14)
Jenkins and Cecins/ 2010[41]		Poor (11)							
Chatterjee et al/2010[42]		Good (3)							
Hernandes et al/2011[44]		Excellent							
Andersson et al/2011[43]		Fair (3)							
Puhan et al/2011[45]									Fair (8)
Chandra et al/2012[46]									Fair (11,12)
M6MWT									
Burioka et al/2014[47]						Poor (3)			
12MD									
Mungall and Hainsworth/1979[48]		Poor (3,11)				Poor (3,7,8)			
Larson et al/1996[49]		Fair (3,8,11)				Fair (4)			
6MST									
Borel et al/2010[50]		Poor (3,8,11)				Poor (3)			
Marrara et al/2012[51]									Poor
da Costa et al/2014[52]		Fair (3,8)							Poor (11,12,13)
Pessoa et al/2014[53]								Fair (3,4)	
Coquart et al/2015[54]		Fair (3)	Fair (3)						
30MWT									
Andersson et al/2011[43]		Fair (3)	Fair (3)			Fair (3,4)			
ISWT/SWT									
Singh et al/1992[55]		Poor (3)				Poor (3)			
Singh et al/1994[56]						Poor (3,4)			
Eiser et al/2003[36]		Poor (11)							Poor (14)
Rosa et al/2006[57]						Poor (3)			
Perrault et al/2009[58]		Fair (3)							
McKeough et al/2011[59]		Poor (11)							Poor (13)
ESWT									
Revill et al/1999[60]		Poor (3,11)							Poor (3,13)

	Psychometric Properties								
Authors/Year	Box A Internal Consistency	Box B Reliability	Box C Measurement Error	Box D Content Validity	Box E Structural Validity	Box F Hypothesis Testing	Box G Cross-Cultural Validity	Box H Criterion Validity	Box I Responsiveness
Brouillard et al/2008[61]									Poor (3,13)
Revill et al/2009[62]		Poor (11)							
McKeough et al/2011[59]		Poor (11)							Poor (13)
Borel et al/2014[63]		Good (7)							Fair (12)
MSWT									
Campo et al/2006[64]		Fair (3)						Fair (3)	
Usual and fast walking speeds									
Rozenberg et al/2014[65]		Poor (3)				Poor (3)			
Incremental treadmill test									
Mathur et al/1995[66]		Poor (3,11)							
Endurance treadmill test									
Cooper et al/2010[67]		Excellent							
Incremental cycle ergometer test									
Mathur et al/1995[66]		Poor (3,11)							
Covey et al/1999[68]		Good (3)							
Cox et al/1989[69]		Poor (3,11)							
Brown et al/2008[39]						Fair (4)			
Puhan et al/2011[45]									Fair (8)
Endurance cycle ergometer test									
van't Hul et al/2010[70]		Good (3)				Fair (3,4)			
UULEX									
Takahashi et al/2003[71]		Poor (3)				Poor (3)			
Janaudis-Ferreira et al/2013[72]						Fair (3)			Poor (6)
PBRT									
Zhan et al/2006[73]		Poor (3)				Poor (3)			
Janaudis-Ferreira et al/2013[72]						Fair (3)			Poor (6)
Semipaced 3CRT									
Aguilaniu et al/2014[74]		Fair (3)				Fair (3,9,10)			
STST									
Ozalevli et al/2007[75]						Fair (4)			
5STS									
Jones et al/2013[76]		Good (3)				Good (1,7,8)			Good (12)
Usual 4MGS									
Kon et al/2014[77]		Good (3)	Good (3)			Fair (4,7,8)			
Kon et al/2013[78]									Fair (11,12)
Karpman et al/2014[79]		Good (1,3)	Good (1,3)						
Maximal 4MGS									

	Psychometric Properties								
Authors/Year	Box A Internal Consistency	Box B Reliability	Box C Measurement Error	Box D Content Validity	Box E Structural Validity	Box F Hypothesis Testing	Box G Cross-Cultural Validity	Box H Criterion Validity	Box I Responsiveness
Karpman et al/2014[79]		Good (1,3)	Good (1,3)						
Usual 10MGS									
Karpman et al/2014[79]		Good (1,3)	Good (1,3)						
Maximal 10MGS									
Karpman et al/2014[79]		Good (1,3)	Good (1,3)						
GST									
Hill et al/2008[80]		Poor (3,11)				Poor (3)			Fair (3,8)
TChester									
Karloh et al/2013[81]						Poor (3)			
de Camargo et al/ 2011[82]		Fair (3)				Fair (3,4,8)			
SRAT									
Chura et al/2012[83]		Poor (3)							
SCPT									
Roig et al/2010[84]		Poor (3)				Poor (3)			
Glittre ADL-test									
Skumlien et al/2006[85]		Good (3)				Fair (4,8,9)			Fair (3,12,13,14)
Corrêa et al/2011[86]						Poor (3)			
Karloh et al/2014[87]						Poor (3)			
PW 610									
Sant'Anna et al/2012[88]		Fair (3)						Fair (3,4)	
SAB									
Farooqi et al/2013[89]								Poor (3)	
Patel et al/2007[90]		Poor (3)				Poor (3)			
Waschki et al/2012[91]						Fair (4,7,8,9,10)			
Watz et al/2009[92]		Good (11)				Fair (4,7,8)			
Cavalheri et al/2011[93]								Poor (6)	
Hill et al/2010[94]		Poor (3,11)						Poor (3,6)	
Rabinovich et al/ 2013[95]								Good (3)	
Andersson et al/2014[96]								Poor (3,6)	
DAM									
Kanda et al/2012[97]		Poor (3)							
Pitta et al/2005[98]								Poor (3)	
Rabinovich et al/ 2013[95]								Good (3)	
Andersson et al/2014[96]								Poor (3,6)	
Lifecorder PLUS (Kenz Suzuken Co Ltd, Nagoya, Japan)									
Rabinovich et al/ 2013[95]								Good (3)	
Actiwatch Spectrum (Philips Respironics, Bend, OR)									

	Psychometric Properties								
Authors/Year	Box A Internal Consistency	Box B Reliability	Box C Measurement Error	Box D Content Validity	Box E Structural Validity	Box F Hypothesis Testing	Box G Cross-Cultural Validity	Box H Criterion Validity	Box I Responsiveness
Rabinovich et al/2013[95]								Good (3)	
Actimarker									
Sugino et al/2012[99]		Poor (3)				Poor (3)			
AH									
Farooqi et al/2013[89]								Poor (3)	
SAM									
Cindy et al/2012[100]								Poor (3,6)	Poor (3,4)
Moy et al/2012[101]						Fair (4,10)		Poor (5,6)	
ActivPAL									
Cindy et al/2012[100]								Poor (3,6)	Poor (3,4)
SCAM									
Coronado et al/2003[102]									Poor (3,13,14)
RT3 (StayHealthy, Inc., Monrovia, CA)									
Rabinovich et al/2013[95]								Good (3)	
Actigraph GT3X: Actigraph LLC, Pensacola, FL									
Rabinovich et al/2013[95]								Good (3)	
Pedometer: Fitty 3 (Kasper & Richter Company, Uttenreuth, Germany)									
Schönhofer et al/1997[103]		Poor (3)							
Tritrac R3D accelerometer									
Steele et al/2000[104]		Fair (3)				Fair (3,4,8)			
Three-axis accelerometers: Fitbit Ultra (Fitbit Inc., San Francisco)									
Vooijs et al/2014[105]								Poor (3)	
Three-axis accelerometers: Personal Activity Monitor AM300 (PAM BV Doorwerth, the Netherlands)									
Vooijs et al/2014[105]								Poor (3)	

The numbers in () correspond to the item in each COSMIN box of which the assessment is based on.

10MGS = 10-meter gait speed, 10MGS = 10-meter gait speed, 12MD = 12-minute distance walk, 2MWT = 2-minute walk test, 30MWT = 30-meter walk test, 3CRT = 3-minute chair rise test, 4MGS = 4-meter gait speed, 5STS = five-repetition sit-to-stand test, 6MST = 6-minute step test, 6MWT = 6-minute walk test, ADL = activity of daily living, AH = ActiHeart, DAM = DynaPort activity monitor, ESWT = endurance shuttle walking test, GST = grocery shelving task, ISWT/SWT = incremental shuttle walk test, M6MWT = modified 6-minute walk test, MSWT = modified SWT, PBRT = 6-minute pegboard and ring test, PW = power walker, SAB = SenseWear armband, SAM = StepWatch™ activity monitor, SCAM = self-contained activity monitor, SCPT = stair climb power test, SRAT = steep ramp anaerobic test, STST = sit-to-stand test, TChester = Chester step test, UULEX = unsupported upper limb exercise test.

TABLE 5. Results of Quality of Design, Methods, and Reporting for Included Studies on Patient-Reported Measures

	Psychometric Properties								
Authors/Year	**Box A Internal Consistency**	**Box B Reliability**	**Box C Measurement Error**	**Box D Content Validity**	**Box E Structural Validity**	**Box F Hypothesis Testing**	**Box G Cross-Cultural Validity**	**Box H Criterion Validity**	**Box I Responsiveness**
LCADL									
Garrod/2000[106]	Poor (6,7)					Fair (2,4)			
Garrod/2002[107]									Fair (11,12)
		Poor (3)							
Carpes/2008[108]						Fair (2,3,8)			
		Fair (2,3)							
Kovelis et al/2011[109]									Poor (3)
FPI									
Larson et al/1998[110]	Poor (5,6)					Fair (4)			
Leidy/1999[111]	Poor (6)			Good (3)		Good (2)			
		Good (3,7,9)							
Weldam et al/2015[112]	Poor (5,6)					Good (1,2,3)			
		Good (1,2,3)							
Ozkan et al/2009[113]									
FPI-SF									
Leidy and Knebel/2010[114]	Poor (6)			Good (3)		Good (2)			
		Good (3,7,9)							
Guo et al/2011[115]	Poor (5,6)					Fair (4,8)	Poor (3,14)		
		Fair (3)							
Leidy et al/2012[116]				Good (3)					
Wall/2007[117]	Poor (5)					Good (1,2,4,6)			
LCOPD									
Stephen/2011[118]	Good (1,6)			Good (3)	Good (2)	Fair (4,7,8)			
		Fair (11)for UK; Fair (3,11) for USA							
Stephen/2012[119]	Fair (5)					Fair (1,3,4,7,8)	Poor (14)		
		Fair (11)							
CARS									
Morimoto et al/2003[120]	Fair (3)			Poor (5)	Fair (3)	Fair (2,4,8)			
UCDQ									
Lee et al/1998[121]	Fair (8)				Fair (5)				
Hodgev et al/2003[122]	Poor (6)				Poor (4)	Fair (2,4,7,8)			
		Poor (3)							
Binazzi et al/2010[123]									Poor (3)
ADL-D scale									
Yoza et al/2009[124]	Poor (5,)					Fair (2,4,7,8)			
DIRECT									

	Psychometric Properties								
Authors/Year	**Box A Internal Consistency**	**Box B Reliability**	**Box C Measurement Error**	**Box D Content Validity**	**Box E Structural Validity**	**Box F Hypothesis Testing**	**Box G Cross-Cultural Validity**	**Box H Criterion Validity**	**Box I Responsiveness**
Aguilaniu et al/2011[125] SOBDA	Good (3)					Fair (8)			
Wilcox et al/2013[126]	Excellent	Excellent			Excellent	Fair (4,7,8)			
Watkins et al/2013[127]	Poor (5)	Fair (8)				Fair (4,7,8)			Fair (8,11,12,13,14)
CDLM									
Partridge et al/2010[128]	Poor (5)	Fair (5,8)				Fair (4,7,8)			Fair (14)
PFSDQ-M									
Wingårdh et al/2007[129]		Fair (3)							
Kovelis et al/2008[130]	Poor (5)	Fair (3,8,10)						Poor (4)	
Lareau et al/1998[131]	Poor (5)	Fair (2,3,10)			Poor (4)				Poor (4)
Kovelis et al/2011[109]									Poor (3)
PFSS									
Weaver et al/1998[132]	Fair (8)	Poor (3)		Poor (2)	fair (5)	Fair (4,7,8)			
PFSS-11									
Chen et al/2010[133]	Good (1)	Poor			Good (1)	Fair (2,9)			Fair (13,14)
DASI									
Tavares et al/2012[134]		Good (3)				Fair (4)			
Carter et al/2002[135]						Fair (4)		Good (1,4)	
MRADL									
Yohannes et al/2002[136]		Good (3) (mail) Fair (3) (face to face)							

The numbers in () correspond to the item in each COSMIN box of which the assessment is based on.

ADL-D = activity of daily living dyspnea scale, CARS = COPD activity rating scale, CDLM = capacity of daily living during the morning questionnaire, DASI = Duke activity status index, DIRECT = disability related to COPD tool, FPI = functional performance inventory, FPI-SF = functional performance inventory short form, LCADL = London chest activity of daily living scale, LCOPD = living with COPD questionnaire, MRADL = Manchester respiratory activities of daily living questionnaire, PFSDQ-M = pulmonary functional status & dyspnea questionnaire-modified, PFSS = pulmonary functional status scale, PFSS-11 = short-form pulmonary functional status scale, SOBDA = shortness of breath with daily activities, UCDQ = University of Cincinnati dyspnea questionnaire.

TABLE 6. A Summary of Best-Evidence Synthesis

Category	Instrument	Box A Internal Consistency	Box B Reliability	Box C Measurement Error	Box D Content Validity	Box E Structural Validity	Box F Hypothesis Testing	Box G Cross-Cultural Validity	Box H Criterion Validity	Box I Responsiveness
Exercise testing	2MWT		+				+			?
Exercise testing	6MWT		+++							+
Exercise testing	M6MWT						?			
Exercise testing	12MD		++				−			
Exercise testing	6MST		++	?			?		+	?
Exercise testing	30MWT		+	?			+			
Exercise testing	ISWT/SWT		+				?			?
Exercise testing	ESWT		++							+
Exercise testing	MSWT		+							+
Exercise testing	Usual and fast walking speeds		?				?			
Exercise testing	Incremental treadmill test		++							
Exercise testing	Endurance treadmill test		+++							
Exercise testing	Incremental cycle ergometer test		?							?
Exercise testing	Endurance cycle ergometer test		++							
Exercise testing	UULEX		?				+			?
Exercise testing	PBRT		?				+			?
Exercise testing	Semipaced 3CRT		+				?			
Exercise testing	STST						+			
Exercise testing	5STS		++				++			− −
Exercise testing	Usual 4MGS		+++	?			+			−
Exercise testing	Maximal 4MGS		++	?						
Exercise testing	Usual 10MGS		++	?						
Exercise testing	Maximal 10MGS		++	?						
Exercise testing	GST		?				?			−
Exercise testing	TChester		+				+			
Exercise testing	SRAT		?							
Exercise testing	SCPT		?				?			
Exercise testing	Glittre ADL-test		++				−			?
Activity monitor	DAM		?						++	
Activity monitor	PW 610		+						+	
Activity monitor	SAB		+				+		− −	
Activity monitor	Lifecorder PLUS								− −	
Activity monitor	Actiwatch Spectrum								− −	
Activity monitor	AH								?	
Activity monitor	Actimarker		?				?			
Activity monitor	SAM						?		?	?

Category	Instrument	Box A Internal Consistency	Box B Reliability	Box C Measurement Error	Box D Content Validity	Box E Structural Validity	Box F Hypothesis Testing	Box G Cross-Cultural Validity	Box H Criterion Validity	Box I Responsiveness
Activity monitor	ActivPAL								?	?
Activity monitor	SCAM									?
Activity monitor	RT3								− −	
Activity monitor	Actigraph GT3X								− −	
Activity monitor	Fitty 3		?							
Activity monitor	Tritrac R3D accelerometer		+				+			
Activity monitor	FB								?	
Activity monitor	PAM								?	
PRO	LCADL	?	+				++			−
PRO	FPI	?	+++		++		+++			
PRO	FPI-SF	?	++		+++		++	?		
PRO	LCOPD	++	++		++	++	++	?		
PRO	CARS	+			?	++	+			
PRO	UCDQ	+				+	+			?
PRO	ADL-D	?					+			
PRO	DIRECT	++					+			
PRO	SOBDA	+++	+++			− − −	++			?
PRO	CDLM	?	+				+			?
PRO	PFSDQ-M	?	±			?			?	?
PRO	PFSS	+	?		?	−	+			
PRO	PFSS-11	++	?			++	+			?
PRO	DASI		++				±		− −	
PRO	MRADL		++ for mail; + for face to face							

10MGS = 10-meter gait speed, 10MGS = 10-meter gait speed, 12MD = 12-minute distance walk, 2MWT = 2-minute walk test, 30MWT = 30-meter walk test, 3CRT = 3-minute chair rise test, 4MGS = 4-meter gait speed, 4MGS = 4-meter gait speed, 5STS = Five-repetition sit-to-stand test, 6MST = 6-minute step test, 6MWT = 6-minute walk test, ADL-D = activity of daily living dyspnea scale, AH = ActiHeart, CARS = COPD activity rating scale, CDLM = capacity of daily living during the morning questionnaire, DAM = DynaPort activity monitor, DASI = Duke activity status index, DIRECT = disability related to COPD tool, ESWT = endurance shuttle walking test, FPI = functional performance inventory, FPI-SF = functional performance inventory short form, GST = grocery shelving task, ISWT/SWT = incremental shuttle walk test, LCADL = London chest activity of daily living scale, LCOPD = living with COPD questionnaire, M6MWT = modified 6-minute walk test, MRADL = Manchester respiratory activities of daily living questionnaire, MSWT = modified SWT, PBRT = 6-minute pegboard and ring test, PFSDQ-M = pulmonary functional status and dyspnea questionnaire-modified, PFSS = pulmonary functional status scale, PFSS-11 = short-form pulmonary functional status scale, PRO = patient-reported outcomes, PW = power walker, SAB = SenseWear armband, SAM = StepWatch™ activity monitor, SCAM = self-contained activity monitor, SCPT = stair climb power test, SOBDA = shortness of breath with daily activities, SRAT = steep ramp anaerobic test, STST = sit-to-stand test, TChester = Chester step test, UCDQ = University of Cincinnati dyspnea questionnaire, UULEX = unsupported upper limb exercise test.

? = indeterminate rating, + = positive rating, − = negative rating.

practice, this review recommends 10 out or 57 instruments assessing functional status of COPD patients. More importantly, it demonstrates how to choose suitable measures according to both the studies on elevating these measures and the requirements of clinical practice.

Comparing Performance-Based Measures with Patient-Reported Measures

According to the summary of best-evidence synthesis, performance-based measures did not have as much positive evidence on measurement property in comparison to patient-reported measures. The lack of adequate positive evidence contradicts their present importance in measuring functional status in COPD patients. Performance-based measures objectively measure what patients actually do by assessing indicators like timing, counting, and distance.[137] It was believed to be more likely to fully characterize a change in functional status than patient-reported measures alone.[138] Some of these performance-based measures have been widely approved and used for many years to evaluate treatment effect, to assess health status, and to explore etiology. For example, 6MWT is a widely used walking test in clinical practice, and it was often used as a standard for other instruments.[139–142] However, its positive evidence confined to *reliability*, which is also a common situation in all performance-based measures. Activity monitors capture the patients' activities of daily living. They are an essential supplement to laboratory tests. Although there are numerous studies (27 studies), the qualities of these studies were poor (16 poor studies), leading to a weaker positive rating in evidence synthesis. More good quality studies need to be conducted in order to assess the measurement properties of these performance-based measures.

Weakness Limiting the Quality and Neglected Measurement Properties in the Existing Studies on Performance-Based Measures and Patient-Reported Measures

The methodological qualities of the studies included in this review ranged from *poor* to *excellent*. *Good* and *excellent* quality studies only took up ~20% of all studies. In terms of performance-based measures, inadequate sample size was one major drawback, probably because performance-based measurements are more difficult to conduct. Some studies on performance-based measures had a sample size <10.[37,47,66,71,90,105] The sample size should be enlarged in future similar studies according to COSMIN criterion, which is ≥ 100 for excellent, $50–99$ for good, $30–49$ for fair, and <30 for poor. However, one thing to note is that the COSMIN checklist was originally developed to assess studies focusing on patient-reported measures. Considering the differences in instrument characteristics and study designs between studies on performance-based measures and studies on patient-reported measures, the sample size criteria may need some adjustment. Methodology on performance-based measures should be discussed in the future. Another obvious drawback affecting studies on performance-based measures was that methods did not meet the COSMIN criteria. For example, some studies measuring *reliability* tend to not calculate the ICC or Pearson or Spearman correlations, no correlation was calculated with other comparative instruments, whereas only *P* values were used when testing *responsiveness*, and no adequate hypotheses were formulated a priori. Qualities of studies on patient-reported measures were better than studies on performance-based measures (Table 5). However, checking for unidimensionality, enlarging sample sizes, and formulating hypotheses a priori may further improve all study quality.

According to the results, the included studies and positive evidence were confined to several measurement properties. Some important measurement properties, including content validity and responsiveness, were neglected or poorly reported. Content validity examines the extent to which the concepts of interest are comprehensively represented by the items of the questionnaire,[24,143] so it is especially important for studies on patient-reported measures. To measure content validity, a clear concept model is to be developed.[144] However, present PROs that aim to measure physical activity in chronic respiratory disease patients or similar populations (chronic heart disease patients or the elderly) are rarely based on a conceptual framework.[145] Additionally, a standard method to assess content validity should be applied. According to COSMIN, an appropriate method is to have experts and the target population to assess the relevance and comprehensiveness of the instrument (s) based on criteria set by COSMIN. The two studies on content validity measurement were determined to be *poor* because they did not meet the above-mentioned criterion. *Responsiveness* is another key issue for future studies on both performance-based measures and patient-reported measures. An important role of functional status measurement is the evaluation of the effect of rehabilitation or treatment. Therefore, it is important for measurement instruments to respond to change. In the present studies measuring *responsiveness*, the rating of *poor* was given because of inadequate sample size. Also, most *fair* studies used *P* values instead of showing correlation with comparative instruments or with AUC values. Further studies exploring the responsiveness of functional status instruments should be conducted by applying appropriate methods.

Choosing Measures According to the Present Evidence

Valid and systematical measures of COPD patients' health status are the base of the accurate quantification of the therapy effects. Facing an impressive and increasing number of measures assessing functional status of COPD patients, clinicians might be confused and feel difficult to find one measure satisfying all of their demands. Another source of confusion is the inconsistent conclusions of reports which employed various measures to evaluate the effect of therapy. It is difficult for clinicians to choose best care for patients by comparing and combining results of these clinical trials.

According to the results, none of measures has been tested for all measurement properties. However, the existing evidence demonstrates that some instruments perform better: 6MWT, endurance treadmill test, and usual 4MGS; and FPI, FPI-SF, LCOPD, CARS, UCDQ, SOBDA, and PFSS-11. These instruments should be preferred in future studies and clinical practice (Table 7). 6MWT was proven to predict the survival in COPD patients well.[146–148] Usual 4MGS needs much shorter course than 6MWT, making it useful for frail patients and applicable in most healthcare settings (including home). MCID was reported to be 0.11 m/s.[77] It is worth to be considered as an instrument for health management of COPD patients. Endurance treadmill test can clearer reflect the physiological limitations.[67] FPI, FPI-SF, LCOPD, CARS, UCDQ, SOBDA, and PFSS-11 are different in terms of their measurement focus and the length of scales. In other words, each measure has its own advantage and most suitable domain. Thus, researchers and clinicians should employ

TABLE 7. Characteristics of Recommended Instruments

Instrument	Category	Measurement Content and Protocol	Output*	Time to Administer	Environment to Administer
6MWT	Exercise testing	Quickly walk on a flat, hard surface in a period of 6 min	Distance	–	A 100-ft hallway which is indoors, along a long, flat, straight, enclosed corridor with a hard surface that is seldom traveled
Usual 4MGS	Exercise testing	Walk at usual speed	Time taken to complete the 4-m course	<2 min	A 4-m flat, unobstructed course marked out with tape
Endurance treadmill test	Exercise testing	Walking on a treadmill with a fixed percentage of the maximum work rate applied as a constant work rate	Time to exhaustion	–	Treadmill
FPI	Patient-reported	Body care, household maintenance, physical exercise, recreation, spiritual activities, and social activities	1–4 (+the activity is not performed for reasons other than health)	65 items	–
FPI-SF	Patient-reported	Body care, household, maintenance, physical exercise, recreation, spiritual activities, and social activities	Three-point scale	32 items	–
LCOPD	Patient-reported	Self-actualization needs, safety and security needs, Independence needs, self-esteem needs, control needs, social and relationship needs	''True'' (scored 1) and ''not true'' (scored 0) response options	22 items	–
CARS	Patient-reported	Self-care activity, domestic activity, outdoor activity, social interaction activity	2 (completely independent), 1 (partially dependent), 0 (dependent)	12 items	–
UCDQ	Patient-reported	Breathlessness during physical activity, breathlessness during speaking activities, when speaking during physical activity	1–5 (+not interested)	30 items	–
SOBDA	Patient-reported	Different levels of exertion and body positions which impact patient's experience of SOB	A scale from ''not at all'' to ''so short of breath that I did not do the activity''	13 items	–
PFSS-11	Patient-reported	Physical functioning, emotional functioning	A 5-point Likert-type response	11 items	–

*Primary outcome measures are shown here.4MGS = 4-meter gait speed, 6MWT = 6-minute walk test, CARS = COPD activity rating scale, SOB, FPI = functional performance inventory, FPI-SF = functional performance inventory short form, LCOPD = living with COPD questionnaire, PFSS = pulmonary functional status scale, SOB = shortness of breath, SOBDA = shortness of breath with daily activities, UCDQ = University of Cincinnati dyspnea questionnaire.

those measures whose measurement properties alignment to their purposes. For example, if the instruments were to be used to measure the therapeutic effect of pulmonary rehabilitation or a respiratory medicine, then the responsiveness of the measurement instruments should be preferred. Finally, functional status measurement of COPD patients is complex, as it contains multidimensional constructs. Different types of instruments have their own strength. It was suggested that both types of measures—performance-based measures and patient-reported measures—are complementary rather than competing when assessing functional status of COPD patients.[149] Finding an optimal combination of measures from both types is worth for further research.

LIMITATIONS

This study has several limitations. First, some comprehensive HRQOL instruments, including dimensions measuring functional status were excluded in our study. The reason is that their reliability and validity were calculated for the whole instrument rather than the dimension of interest, functional status. Second, some studies focusing on evaluating the therapy effect were excluded because they failed to provide enough information on measurement properties. Admittedly, reviewing measurement properties for a certain type of therapy is interesting and valuable for clinical practice, which should be implemented in the future.

In summary, further research is needed to evaluate the measurement properties of performance-based measures because there is a lack of available information and present studies lack in quality. Content validity and responsiveness should be fully assessed in all instruments, and sample size needs to be enlarged. As for choosing measurement tools for functional status in COPD patients, we recommend FPI, FPI-SF, LCOPD, CARS, UCDQ, SOBDA, PFSS-11, 6MWT, endurance treadmill test, and usual 4MGS. These instruments are different in their measurement content or administer requirement, which may tailor to different usage in clinical practice. We also recommend selecting instruments that perform well in certain measurement properties required for certain assessment purposes and combining instruments from both measurement types.

ACKNOWLEDGMENTS

The authors thank Terwee CB for providing us permission to use the COSMIN checklist and instruction to do the systematic review using COSMIN. We also thank Weldam, S.W.M., for responding to their consultation about quality criteria for measurement properties in his article. We thank Nan Jiang for her kind help with correcting English of the manuscript.

REFERENCES

1. From the Global Strategy for the Diagnosis, Management and Prevention of COPD, Global Initiative for Chronic Obstructive Lung Disease (GOLD) 2016. Available from: http://goldcopd.org/. Accessed March 10, 2016.
2. Fan VS, Ramsey SD, Make BJ, et al. Physiologic variables and functional status independently predict COPD hospitalizations and emergency department visits in patients with severe COPD. *COPD.* 2007;4:29–39.
3. Pitta F, Troosters T, Probst VS, et al. Physical activity and hospitalization for exacerbation of COPD. *Chest.* 2006;129:536–544.
4. Wang KY, Sung PY, Yang ST, et al. Influence of family caregiver caring behavior on COPD patients' self-care behavior in Taiwan. *Respir Care.* 2012;57:263–272.
5. Celli BR, MacNee W. ATS/ERS Task Force. Standards for the diagnosis and treatment of patients with COPD: a summary of the ATS/ERS position paper. *Eur Respir J.* 2004;23:932–946.
6. National Clinical Guideline Centre (UK). Chronic Obstructive Pulmonary Disease: Management of Chronic Obstructive Pulmonary Disease in Adults in Primary and Secondary Care [Internet]. London: Royal College of Physicians (UK); 2010 Jun. (NICE Clinical Guidelines, No. 101.) Available from: http://www.ncbi.nlm.nih.gov/books/NBK65039/. Accessed August 25, 2015.
7. Yohannes AM, Connolly MJ. Pulmonary rehabilitation programmes in the UK: a national representative survey. *Clin Rehabil.* 2004;18:444–449.
8. Brooks D, Sottana R, Bell B, et al. Characterization of pulmonary rehabilitation programs in Canada in 2005. *Can Respir J.* 2007;14:87–92.
9. Leidy NK. Functional performance in people with chronic obstructive pulmonary disease. *Image J Nurs Sch.* 1995;27:23–34.
10. Larson JL. Functional performance and physical activity in chronic obstructive pulmonary disease: theoretical perspectives. *COPD.* 2007;4:237–242.
11. Reardon JZ, Lareau SC, ZuWallack R. Functional status and quality of life in chronic obstructive pulmonary disease. *Am J Med.* 2006;119(10 Suppl 1):32–37.
12. Leidy NK. Functional status and the forward progress of merry-go-rounds: toward a coherent analytical framework. *Nurs Res.* 1994;43:196–202.
13. Leidy NK. Subjective measurement of activity in chronic obstructive pulmonary disease. *COPD.* 2007;4:243–249.
14. Wilson IB, Cleary PD. Linking clinical variables with health-related quality of life. A conceptual model of patient outcomes. *JAMA.* 1995;273:59–65.
15. Ferrans CE, Zerwic JJ, Wilbur JE, et al. Conceptual model of health-related quality of life. *J Nurs Scholarsh.* 2005;37:336–342.
16. Kocks JW, Asijee GM, Tsiligianni IG, et al. Functional status measurement in COPD: a review of available methods and their feasibility in primary care. *Prim Care Respir J.* 2011;20:269–275.
17. Fotheringham I, Meakin G, Punekar YS, et al. Comparison of laboratory- and field-based exercise tests for COPD: a systematic review. *Int J Chron Obstruct Pulmon Dis.* 2015;10:625–643.
18. Janaudis-Ferreira T, Beauchamp MK, Robles PG, et al. Measurement of activities of daily living in patients with COPD: a systematic review. *Chest.* 2014;145:253–271.
19. Stull DE, Leidy NK, Jones PW, et al. Measuring functional performance in patients with COPD: a discussion of patient-reported outcome measures. *Curr Med Res Opin.* 2007;23:2655–2665.
20. 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;10:22.
21. 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;63:737–745.
22. Mokkink LB, Terwee CB, Knol DL, et al. Protocol of the COSMIN study: consensus-based standards for the selection of health measurement Instruments. *BMC Medical Res Methodol.* 2006;6:2.

23. Terwee CB, Bot SD, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol.* 2007;60:34–42.
24. Kroman SL, Roos EM, Bennell KL, et al. Measurement properties of performance-based outcome measures to assess physical function in young and middle-aged people known to be at high risk of hip and/or knee osteoarthritis: a systematic review. *Osteoarthritis Cartilage.* 2014;22:26–39.
25. Mijnarends DM, Meijers JM, Halfens RJ, et al. Validity and reliability of tools to measure muscle mass, strength, and physical performance in community-dwelling older people: a systematic review. *J Am Med Dir Assoc.* 2013;14:170–178.
26. Granger CL, McDonald CF, Parry SM, et al. Functional capacity, physical activity and muscle strength assessment of individuals with non-small cell lung cancer: a systematic review of instruments and their measurement properties. *BMC Cancer.* 2013;13:135.
27. Pike S, Lannin NA, Cusick A, et al. A systematic review protocol to evaluate the psychometric properties of measures of function within adult neuro-rehabilitation. *Syst Rev.* 2015;4:86.
28. Gupta N, Pinto LM, Morogan A, et al. The COPD assessment test: a systematic review. *Eur Respir J.* 2014;44:873–884.
29. Weldam SW, Schuurmans MJ, Liu R, et al. Evaluation of quality of life instruments for use in COPD care and research: a systematic review. *Int J Nurs Stud.* 2013;50:688–707.
30. Janaudis-Ferreira T, Beauchamp MK, Goldstein RS, et al. How should we measure arm exercise capacity in patients with COPD? *CHEST.* 2012;141:111–120.
31. Terwee CB, Jansma EP, Riphagen II, et al. Development of methodological PubMed search filter for finding studies on measurement properties of measurement instruments. *Qual Life Res.* 2009;18:1115–1123.
32. Terwee CB, Mokkink LB, Knol DL, et al. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. *Qual Life Res.* 2012;21:651–657.
33. Furlan AD, Pennick V, Bombardier C, et al. Editorial Board, Cochrane Back Review Group. 2009 updated method guidelines for systematic reviews in the Cochrane Back Review Group. *Spine.* 2009;34:1929–1941.
34. van Tulder M, Furlan A, Bombardier C, et al. Editorial Board of the Cochrane Collaboration Back Review Group. Updated method guidelines for systematic reviews in the cochrane collaboration back review group. *Spine.* 2003;28:1290–1299.
35. Butland RJ, Pang J, Gross ER, et al. Two-, six-, and 12-minute walking tests in respiratory disease. *Brit Med J (Clin Res Ed).* 1982;284:1607–1608.
36. Eiser N, Willsher D, Doré CJ. Reliability, repeatability and sensitivity to change of externally and self-paced walking tests in COPD patients. *Respir Med.* 2003;97:407–414.
37. Leung AS, Chan KK, Sykes K, et al. Reliability, validity, and responsiveness of a 2-min walk test to assess exercise capacity of COPD patients. *Chest.* 2006;130:119–125.
38. Casanova C, Cote CG, Marin JM, et al. The 6-min walking distance: long-term follow up in patients with COPD. *Eur Respir J.* 2007;29:535–540.
39. Brown CD, Benditt JO, Sciurba FC, et al. National emphysema Treatment Trial Research Group. Exercise testing in severe emphysema: association with quality of life and lung function. *COPD.* 2008;5:117–124.
40. Spencer LM, Alison JA, McKeough ZJ. Six-minute walk test as an outcome measure: are two six-minute walk tests necessary immediately after pulmonary rehabilitation and at three-month follow-up? *Am J Phys Med Rehabil.* 2008;87:224–228.
41. Jenkins S, Cecins NM. Six-minute walk test in pulmonary rehabilitation: do all patients need a practice test? *Respirology.* 2010;15:1192–1196.
42. Chatterjee AB, Rissmiller RW, Meade K, et al. Reproducibility of the 6-minute walk test for ambulatory oxygen prescription. *Respiration.* 2010;79:121–127.
43. Andersson M, Moberg L, Svantesson U, et al. Measuring walking speed in COPD: test–retest reliability of the 30-metre walk test and comparison with the 6-minute walk test. *Prim Care Respir.* 2011;20:434–440.
44. Hernandes NA, Wouters EF, Meijer K, et al. Reproducibility of 6-minute walking test in patients with COPD. *Eur Respir J.* 2011;38:261–267.
45. Puhan MA, Chandra D, Mosenifar Z, et al. National Emphysema Treatment Trial (NETT) Research Group. The minimal important difference of exercise tests in severe COPD. *Eur Respir J.* 2011;37:784–790.
46. Chandra D, Wise RA, Kulkarni HS, et al. Sciurba FC; NETT Research Group. Optimizing the 6-min walk test as a measure of exercise capacity in COPD. *Chest.* 2012;142:1545–1552.
47. Burioka N, Imada A, Kiyohiro A, et al. Modified six-minute walk test: number of steps per second. *Yonago Acta Med.* 2014;57:61–63.
48. Mungall IP, Hainsworth R. Assessment of respiratory function in patients with chronic obstructive airways disease. *Thorax.* 1979;34:254–258.
49. Larson JL, Covey MK, Vitalo CA, et al. Reliability and validity of the 12-minute distance walk in patients with chronic obstructive pulmonary disease. *Nurs Res.* 1996;45:203–210.
50. Borel B, Fabre C, Saison S, et al. An original field evaluation test for chronic obstructive pulmonary disease population: the six-minute stepper test. *Clin Rehabil.* 2010;24:82–93.
51. Marrara KT, Marino DM, Jamami M, et al. Responsiveness of the six-minute step test to a physical training program in patients with COPD. *J Bras Pneumol.* 2012;38:579–587.
52. da Costa JN, Arcuri JF, Goncalves IL, et al. Reproducibility of cadence-free 6-minute step test in subjects with COPD. *Respir Care.* 2014;59:538–542.
53. Pessoa BV, Arcuri JF, Labadessa IG, et al. Validity of the six-minute step test of free cadence in patients with chronic obstructive pulmonary disease. *Braz J Phys Ther.* 2014;18:228–236.
54. Coquart JB, Lemaître F, Castres I, et al. Reproducibility and sensitivity of the 6-minute stepper test in patients with COPD. *COPD.* 2015;12:533–538.
55. Singh SJ, Morgan MD, Scott S, et al. Development of a shuttle walking test of disability in patients with chronic airways obstruction. *Thorax.* 1992;47:1019–1024.
56. Singh SJ, Morgan MD, Hardman AE, et al. Comparison of oxygen uptake during a conventional treadmill test and the shuttle walking test in chronic airflow limitation. *Eur Respir J.* 1994;7:2016–2020.
57. Rosa FW, Camelier A, Mayer A, et al. Evaluating physical capacity in patients with chronic obstructive pulmonary disease: comparing the shuttle walk test with the encouraged 6-minute walk test. *J Bras Pneumol.* 2006;32:106–113.
58. Perrault H, Baril J, Henophy S, et al. Paced-walk and step tests to assess exertional dyspnea in COPD. *COPD.* 2009;6:330–339.
59. McKeough ZJ, Leung RW, Alison JA. Shuttle walk tests as outcome measures: are two incremental shuttle walk tests and two endurance shuttle walk tests necessary? *Am J Phys Med Rehabil.* 2011;90:35–39.

60. Revill SM, Morgan MD, Singh SJ, et al. The endurance shuttle walk: a new field test for the assessment of endurance capacity in chronic obstructive pulmonary disease. *Thorax.* 1999;54:213–222.
61. Brouillard C, Pepin V, Milot J, et al. Endurance shuttle walking test: responsiveness to salmeterol in COPD. *Eur Respir J.* 2008;31:579–584.
62. Revill SM, Williams J, Sewell L, et al. Within-day repeatability of the endurance shuttle walk test. *Physiotherapy.* 2009;95:140–143.
63. Borel B, Pepin V, Mahler DA, et al. Prospective validation of the endurance shuttle walking test in the context of bronchodilation in COPD. *Eur Respir J.* 2014;44:1166–1176.
64. Campo LA, Chilingaryan G, Berg K, et al. Validity and reliability of the modified shuttle walk test in patients with chronic obstructive pulmonary disease. *Arch Phys Med Rehabil.* 2006;87:918–922.
65. Rozenberg D, Dolmage TE, Evans RA, et al. Repeatability of usual and fast walking speeds in patients with chronic obstructive pulmonary disease. *J Cardiopulm Rehabil Prev.* 2014;34:348–354.
66. Mathur RS, Revill SM, Vara DD, et al. Comparison of peak oxygen consumption during cycle and treadmill exercise in severe chronic obstructive pulmonary disease. *Thorax.* 1995;50:829–833.
67. Cooper CB, Abrazado M, Legg D, et al. Development and implementation of treadmill exercise testing protocols in COPD. *Int J Chron Obstruct Pulmon Dis.* 2010;5:375–385.
68. Covey MK, Larson JL, Alex CG, et al. Test–retest reliability of symptom-limited cycle ergometer tests in patients with chronic obstructive pulmonary disease. *Nurs Res.* 1999;48:9–19.
69. Cox NJ, Hendriks JC, Binkhorst RA, et al. Reproducibility of incremental maximal cycle ergometer tests in patients with mild to moderate obstructive lung diseases. *Lung.* 1989;167:129–133.
70. van't Hul A, Gosselink R, Kwakkel G. Constant-load cycle endurance performance test-retest reliability and validity in patients with COPD. *J Cardiopulm Rehabil.* 2010;23:143–150.
71. Takahashi T, Jenkins SC, Strauss GR, et al. A new unsupported upper limb exercise test for patients with chronic obstructive pulmonary disease. *J Cardiopulm Rehabil.* 2003;23:430–437.
72. Janaudis-Ferreira T, Hill K, Goldstein RS, et al. Relationship and responsiveness of three upper-limb tests in patients with chronic obstructive pulmonary disease. *Physiother Can.* 2013;65:40–43.
73. Zhan S, Cerny FJ, Gibbons WJ, et al. Development of an unsupported arm exercise test in patients with chronic obstructive pulmonary disease. *J Cardiopulm Rehabil.* 2006;26:180–187discussion 188–190.
74. Aguilaniu B, Roth H, Gonzalez-Bermejo J, et al. A simple semipaced 3-minute chair rise test for routine exercise tolerance testing in COPD. *Int J Chron Obstruct Pulmon Dis.* 2014;9:1009–1019.
75. Ozalevli S, Ozden A, Itil O, et al. Comparison of the sit-to-stand test with 6 min walk test in patients with chronic obstructive pulmonary disease. *Respir Med.* 2007;101:286–293.
76. Jones SE, Kon SS, Canavan JL, et al. The five-repetition sit-to-stand test as a functional outcome measure in COPD. *Thorax.* 2013;68:1015–1020.
77. Kon SS, Canavan JL, Nolan CM, et al. The 4-metre gait speed in COPD: responsiveness and minimal clinically important difference. *Eur Respir J.* 2014;43:1298–1305.
78. Kon SS, Patel MS, Canavan JL, et al. Reliability and validity of 4-metre gait speed in COPD. *Eur Respir J.* 2013;42:333–340.
79. Karpman C, Lebrasseur NK, Depew ZS, et al. Measuring gait speed in the out-patient clinic: methodology and feasibility. *Respir Care.* 2014;59:531–537.
80. Hill CJ, Denehy L, Holland AE, et al. Measurement of functional activity in chronic obstructive pulmonary disease: the grocery shelving task. *J Cardiopulm Rehabil Prev.* 2008;28:402–409.
81. Karloh M, Corrêa KS, Martins LQ, et al. Chester step test: assessment of functional capacity and magnitude of cardiorespiratory response in patients with COPD and healthy subjects. *Braz J Phys Ther.* 2013;17:227–235.
82. de Camargo AA, Justino T, de Andrade CH, et al. Chester step test in patients with COPD: reliability and correlation with pulmonary function test results. *Respir Care.* 2011;56:995–1001.
83. Chura RL, Marciniuk DD, Clemens R, et al. Test-retest reliability and physiological responses associated with the steep ramp anaerobic test in patients with COPD. *Pulm Med.* 2012;2012:653831.
84. Roig M, Eng JJ, MacIntyre DL, et al. Associations of the stair climb power test with muscle strength and functional performance in people with chronic obstructive pulmonary disease: a cross-sectional study. *Phys Ther.* 2010;90:1774–1782.
85. Skumlien S, Hagelund T, Bjørtuft O, et al. A field test of functional status as performance of activities of daily living in COPD patients. *Resp Med.* 2006;100:316–323.
86. Corrêa KS, Karloh M, Martins LQ, et al. Can the Glittre ADL test differentiate the functional capacity of COPD patients from that of healthy subjects? *Rev Bras Fisioter.* 2011;15:467–473.
87. Karloh M, Karsten M, Pissaia FV, et al. Physiological responses to the Glittre-ADL test in patients with chronic obstructive pulmonary disease. *J Rehabil Med.* 2014;46:88–94.
88. Sant'Anna T, Escobar VC, Fontana AD, et al. Evaluation of a new motion sensor in patients with chronic obstructive pulmonary disease. *Arch Phys Med Rehabil.* 2012;93:2319–2325.
89. Farooqi N, Slinde F, Håglin L, et al. Validation of SenseWear Armband and ActiHeart monitors for assessments of daily energy expenditure in free-living women with chronic obstructive pulmonary disease. *Physiol Rep.* 2013;1:e00150.
90. Patel SA, Benzo RP, Slivka WA, et al. Activity monitoring and energy expenditure in COPD patients: a validation study. *COPD.* 2007;4:107–112.
91. Waschki B, Spruit MA, Watz H, et al. Physical activity monitoring in COPD: compliance and associations with clinical characteristics in a multicenter study. *Resp Med.* 2012;106:522–530.
92. Watz H, Waschki B, Meyer T, et al. Physical activity in patients with COPD. *Eur Respir J.* 2009;33:62–272.
93. Cavalheri V, Donária L, Ferreira T, et al. Energy expenditure during daily activities as measured by two motion sensors in patients with COPD. *Resp Med.* 2011;105:922–929.
94. Hill K, Dolmage TE, Woon L, et al. Measurement properties of the SenseWear armband in adults with chronic obstructive pulmonary disease. *Thorax.* 2010;65:486–491.
95. Rabinovich RA, Louvaris Z, Raste Y, et al. Validity of physical activity monitors during daily life in patients with COPD. *Eur Respir J.* 2013;42:1205–1215.
96. Andersson M, Janson C, Emtner M. Accuracy of three activity monitors in patients with chronic obstructive pulmonary disease: a comparison with video recordings. *COPD.* 2014;11:560–567.
97. Kanda M, Minakata Y, Matsunaga K, et al. Validation of the triaxial accelerometer for the evaluation of physical activity in Japanese patients with COPD. *Internal Med.* 2012;51:369–375.
98. Pitta F, Troosters T, Spruit MA, et al. Activity monitoring for assessment of physical activities in daily life in patients with chronic obstructive pulmonary disease. *Arch Phys Med Rehab.* 2005;86:1979–1985.

99. Sugino A, Minakata Y, Kanda M, et al. Validation of a compact motion sensor for the measurement of physical activity in patients with chronic obstructive pulmonary disease. *Respiration.* 2012;83:300–307.
100. Cindy Ng LW, Jenkins S, Hill K. Accuracy and responsiveness of the stepwatch activity monitor and ActivPAL in patients with COPD when walking with and without a rollator. *Disabil Rehabil.* 2012;34:1317–1322.
101. Moy ML, Danilack VA, Weston NA, et al. Daily step counts in a US cohort with COPD. *Respir Med.* 2012;106:962–969.
102. Coronado M, Janssens JP, de Muralt B, et al. Walking activity measured by accelerometry during respiratory rehabilitation. *J Cardiopulm Rehabil.* 2003;23:357–364.
103. Schönhofer B, Ardes P, Geibel M, et al. Evaluation of a movement detector to measure daily activity in patients with chronic lung disease. *Eur Respir J.* 1997;10:2814–2819.
104. Steele BG, Holt L, Belza B, et al. Quantitating physical activity in COPD using a triaxial accelerometer. *Chest.* 2000;117:1359–1367.
105. Vooijs M, Alpay LL, Snoeck-Stroband JB, et al. Validity and usability of low-cost accelerometers for internet-based self-monitoring of physical activity in patients with chronic obstructive pulmonary disease. *Interact J Med Res.* 2014;3:e14.
106. Garrod R, Bestall JC, Paul EA, et al. Development and validation of a standardized measure of activity of daily living in patients with severe COPD: the London Chest Activity of Daily Living scale (LCADL). *Resp Med.* 2000;94:589–596.
107. Garrod R, Paul EA, Wedzicha JA. An evaluation of the reliability and sensitivity of the London Chest Activity of Daily Living Scale (LCADL). *Resp Med.* 2002;96:725–730.
108. Carpes MF, Mayer AF, Simon KM, et al. The Brazilian Portuguese version of the London Chest Activity of Daily Living scale for use in patients with chronic obstructive pulmonary disease. *J Bras Pneumol.* 2008;34:143–151.
109. Kovelis D, Zabatiero J, Oldemberg N, et al. Responsiveness of three instruments to assess self-reported functional status in patients with COPD. *COPD.* 2011;8:334–339.
110. Larson JL, Kapella MC, Wirtz S, et al. Reliability and validity of the functional performance inventory in patients with moderate to severe chronic obstructive pulmonary disease. *J Nurs Meas.* 1998;6:55–73.
111. Leidy NK. Psychometric properties of the functional performance inventory in patients with chronic obstructive pulmonary disease. *Nurs Res.* 1999;48:20–28.
112. Weldam SW, Lammers JJ, de Bruin-Veelers MC, et al. The Dutch functional performance inventory: validity and reliability in patients with chronic obstructive lung disease. *Nurs Res.* 2015;64:44–52.
113. Ozkan S, Gemicioglu B, Durna Z, et al. Turkish version of the functional performance inventory used on patients with chronic obstructive pulmonary disease. *Saudi Med J.* 2009;30:1098–1100.
114. Leidy NK, Knebel A. In search of parsimony: reliability and validity of the Functional Performance Inventory-Short Form. *Int J Chron Obstruct Pulmon Dis.* 2010;5:415–423.
115. Guo AM, Han JN, Leidy NK, et al. Validation of the Chinese version of the functional performance inventory short form in patients with chronic obstructive pulmonary disease. *J Clin Nurs.* 2011;20:1613–1622.
116. Leidy NK, Hamilton A, Becker K. Assessing patient report of function: content validity of the functional performance inventory-short form (FPI-SF) in patients with chronic obstructive pulmonary disease (COPD). *Int J Chron Obstruct Pulmon Dis.* 2012;7:543–554.
117. Wall MP. Predictors of functional performance in community-dwelling people with COPD. *J Nurs Scholarsh.* 2007;39:222–228.
118. McKenna SP, Meads DM, Doward LC, et al. Development and validation of the living with chronic obstructive pulmonary disease questionnaire. *Qual Life Res.* 2011;20:1043–1052.
119. McKenna SP, Twiss J, Crawford SR, et al. The living with chronic obstructive pulmonary disease scale was successfully adapted for use in Southern European (Italian and Spanish) and Eastern European (Russian) cultures. *J Clin Epidemiol.* 2012;65:906–914.
120. Morimoto M, Takai K, Nakajima K, et al. Development of the chronic obstructive pulmonary disease activity rating scale: reliability, validity and factorial structure. *Nurs Health Sci.* 2003;5:23–30.
121. Lee L, Friesen M, Lambert IR, et al. Evaluation of dyspnea during physical and speech activities in patients with pulmonary diseases. *Chest.* 1998;113:625–632.
122. Hodgev V, Kostianev S, Marinov B. University of cincinnati dyspnea questionnaire for evaluation of dyspnoea during physical and speech activities in patients with chronic obstructive pulmonary disease: a validation analysis. *Clin Physiol Funct Imaging.* 2003;23:269–274.
123. Binazzi B, Lanini B, Romagnoli I, et al. Dyspnea during speech in chronic obstructive pulmonary disease patients: effects of pulmonary rehabilitation. *Respiration.* 2011;81:379–385.
124. Yoza Y, Ariyoshi K, Honda S, et al. Development of an activity of daily living scale for patients with COPD: the activity of daily living dyspnoea scale. *Respirology.* 2009;14:429–435.
125. Aguilaniu B, Gonzalez-Bermejo J, Regnault A, et al. Disability related to COPD tool (DIRECT): towards an assessment of COPD-related disability in routine practice. *Int J Chron Obstruct Pulmon Dis.* 2011;6:387–398.
126. Wilcox TK, Chen WH, Howard KA, et al. Item selection, reliability and validity of the Shortness of Breath with Daily Activities (SOBDA) questionnaire: a new outcome measure for evaluating dyspnea in chronic obstructive pulmonary disease. *Health Qual Life Outcomes.* 2013;11:196.
127. Watkins ML, Wilcox TK, Tabberer M, et al. Shortness of breath with daily activities questionnaire: validation and responder thresholds in patients with chronic obstructive pulmonary disease. *BMJ Open.* 2013;3:e003048.
128. Partridge MR, Miravitlles M, Ståhl E, et al. Development and validation of the capacity of daily living during the morning questionnaire and the global chest symptoms questionnaire in COPD. *Eur Respir J.* 2010;36:96–104.
129. Wingårdh A, Engström CP, Claesson L. Test-retest of a Swedish version of the pulmonary functional status and dyspnea questionnaire-modified. *Scand J Occup Ther.* 2007;14:183–191.
130. Kovelis D, Segretti NO, Probst VS, et al. Validation of the modified pulmonary functional status and dyspnea questionnaire and the Medical Research Council scale for use in Brazilian patients with chronic obstructive pulmonary disease. *J Bras Pneumol.* 2008;34:1008–1018.
131. Lareau SC, Meek PM, Roos PJ. Development and testing of the modified version of the Pulmonary Functional Status and Dyspnea Questionnaire (PFSDQ-M). *Heart Lung.* 1998;27:159–168.
132. Weaver TE, Narsavage GL, Guilfoyle MJ. The development and psychometric evaluation of the pulmonary functional status scale: an instrument to assess functional status in plumonary disease. *J Cardiopulm Rehabil.* 1998;18:105–111.
133. Chen YJ, Narsavage GL, Culp SL, et al. The development and psychometric analysis of the short-form pulmonary functional status scale (PFSS-11). *Res Nurs Health.* 2010;33:477–485.

134. Tavares Ldos A, Barreto Neto J, Jardim JR, et al. Cross-cultural adaptation and assessment of reproducibility of the Duke Activity Status Index for COPD patients in Brazil. *J Bras Pneumol.* 2012;38:684–691.

135. Carter R, Holiday DB, Grothues C, et al. Criterion validity of the Duke Activity Status Index for assessing functional capacity in patients with chronic obstructive pulmonary disease. *J Cardiopulm Rehabil.* 2002;22:298–308.

136. Yohannes AM, Greenwood YA, Connolly MJ. Reliability of the Manchester Respiratory Activities of Daily Living Questionnaire as a postal questionnaire. *Age Ageing.* 2002;31:355–358.

137. Dobson F, Hinman RS, Hall M, et al. Measurement properties of performance-based measures to assess physical function in hip and knee osteoarthritis: a systematic review. *Osteoarthritis Cartilage.* 2012;20:1548–1562.

138. Boer LM, Asijee GM, van Schayck OC, et al. How do dyspnoea scales compare with measurement of functional capacity in patients with COPD and at risk of COPD. *Prim Care Respir J.* 2012;21:202–207.

139. Marquis N, Larivée P, Dubois MF, et al. Are improvements maintained after in-home pulmonary telerehabilitation for patients with chronic obstructive pulmonary disease? *Int J Telerehabil.* 2015;6:21–30.

140. Coyle ME, Shergis JL, Huang ET, et al. Acupuncture therapies for chronic obstructive pulmonary disease: a systematic review of randomized, controlled trials. *Altern Ther Health Med.* 2014;20:10–23.

141. Divo M, Pinto-Plata V. Role of exercise in testing and in therapy of COPD. *Med Clin North Am.* 2012;96:753–766.

142. Deepak TH, Mohapatra PR, Janmeja AK, et al. Outcome of pulmonary rehabilitation in patients after acute exacerbation of chronic obstructive pulmonary disease. *Indian J Chest Dis Allied Sci.* 2014;56:7–12.

143. Guyatt GH, Feeny DH, Patrick DL. Measuring health related quality of life. *Ann Intern Med.* 1993;118:622–629.

144. Aaronson N, Alonso J, Burnam A, et al. Assessing health status and quality-of-life instruments: attributes and review criteria. *Qual Life Res.* 2002;11:193–205.

145. Gimeno-Santos E, Frei A, Dobbels F, et al. Validity of instruments to measure physical activity may be questionable due to a lack of conceptual frameworks: a systematic review. *Health Qual Life Outcomes.* 2011;9:86.

146. Pinto-Plata VM, Cote C, Cabral H, et al. The 6-min walk distance: change over time and value as a predictor of survival in severe COPD. *Eur Respir J.* 2004;23:28–33.

147. Bowen JB, Votto JJ, Thrall RS, et al. Functional status and survival following pulmonary rehabilitation. *Chest.* 2000;118: 697–703.

148. Meriem M, Cherif J, Toujani S, et al. Sit-to-stand test and 6-min walking test correlation in patients with chronic obstructive pulmonary disease. *Ann Thorac Med.* 2015;10:269–273.

149. Downs CA. Functional assessment of chronic obstructive pulmonary disease. *J Am Acad Nurse Pract.* 2011;23:161–167.