



Patient-reported outcomes to assess dyspnoea in interstitial lung disease and pulmonary hypertension: a systematic literature review of measurement properties

Jacqueline M.J. Lemmers , Madelon C. Vonk  and Cornelia H.M. van den Ende

Dept of Rheumatology, Radboud University Medical Center, Nijmegen, The Netherlands.

Corresponding author: Jacqueline M.J. Lemmers (Jacqueline.Lemmers@radboudumc.nl)



Shareable abstract (@ERSpublications)

Currently, four dyspnoea-specific patient-reported outcomes can be recommended for use in ILD patients (including CTD-associated ILD): D12, ERS-IPF-B, KBILD-B and UCSD-SOBQ. The use of D12 is also suitable for patients with PH, including CTD-associated PH. <https://bit.ly/3EsWSb1>

Cite this article as: Lemmers MJM, Vonk MC, van den Ende CHM. Patient-reported outcomes to assess dyspnoea in interstitial lung disease and pulmonary hypertension: a systematic literature review of measurement properties. *Eur Respir Rev* 2022; 31: 220091 [DOI: 10.1183/16000617.0091-2022].

Copyright ©The authors 2022

This version is distributed under the terms of the Creative Commons Attribution Non-Commercial Licence 4.0. For commercial reproduction rights and permissions contact permissions@ersnet.org

Received: 11 May 2022
Accepted: 27 Sept 2022

Abstract

Objective: This Consensus-based Standards for the selection of health measurement INstruments (COSMIN)-based systematic review aims to identify and summarise the quality of measurement properties of dyspnoea-specific patient-reported outcome measures (PROMs) for patients with interstitial lung disease (ILD), pulmonary hypertension (PH) or connective tissue diseases (CTDs).

Methods and results: Relevant articles in PubMed and Embase were screened. Based on COSMIN analysis and the Grading of Recommendations, Assessment, Development and Evaluation approach, overall rating and level of evidence were assessed to formulate recommendations. We identified 26 publications on 10 PROMs. For patients with ILD, including CTD-associated ILD, nine PROMs were evaluated, of which the Dyspnea-12 (D12), EXACT-Respiratory Symptoms Idiopathic Pulmonary Fibrosis Breathlessness subscale (ERS-IPF-B), King's Brief Interstitial Lung Disease Health Status Questionnaire breathlessness and activities subscale (KBILD-B) and the University of California San Diego Shortness of Breath Questionnaire (UCSD-SOBQ) had high-quality evidence for sufficient internal consistency, without high-quality evidence of insufficient measurement properties. We reached this same conclusion regarding the D12 for use in patients with PH, including CTD-associated PH. Most PROMs in this systematic review have moderate- or low-quality evidence on construct validity and responsiveness.

Conclusion: Four dyspnoea-specific PROMs, D12, ERS-IPF-B, KBILD-B and UCSD-SOBQ, can be recommended for use in patients with ILD, including CTD-associated ILD. Of these four, the D12, despite the limited evidence and the lack of evidence on several important domains, is also suitable for use in patients with PH, including CTD-associated PH.

Introduction

Dyspnoea is a cardinal and common symptom in patients with interstitial lung disease (ILD) or pulmonary hypertension (PH), both of which are lethal and frequent complications of connective tissue diseases (CTDs) [1, 2]. In these conditions, dyspnoea has high self-reported disability and causes high disease burden [3, 4].

Patients with ILD, PH or CTDs and their complications are usually monitored by a multidisciplinary team in expert centres that employ guideline-recommended evaluation. Various parameters, including clinical signs and symptoms such as dyspnoea, as well as multiple diagnostic tests are used to evaluate functional capacity, exercise capacity and heart and lung function. This approach is used to evaluate treatment and for the early detection of clinical deterioration, which enables clinicians to tailor the treatment regimen to the individual patient.



Because the course of ILD, PH and CTDs associated with these complications is highly heterogeneous, there is a need to further improve the monitoring of key symptoms such as dyspnoea. A possible option is the use of patient-reported outcome measures (PROMs). PROMs are increasingly recognised as important additional tools for assessing disease severity and response to therapy [5]. The use of a well-validated PROM on dyspnoea could help to evaluate this vital symptom in patients in a structured way, and could possibly be of aid in home-monitoring settings [6, 7].

Furthermore, it has been suggested that dyspnoea be part of the core set for outcome measures in trials for CTD-associated ILD and CTD-associated pulmonary arterial hypertension (PAH), as reported in a consensus statement published by the CTD-ILD working groups of Outcome Measures in Rheumatology (OMERACT) and Outcome Measures in Pulmonary Arterial Hypertension related to Systemic Sclerosis (EPOSS), respectively [8–11]. However, an overview of PROMs and their measurement properties used in literature to assess dyspnoea in patients with ILD, PH or CTDs is lacking to date. Such an overview would support choices for the use of outcome measures both in clinical practice and research.

This review aims to assess measurement properties of self-completed PROMs assessing dyspnoea applied in patients with a diagnosis of ILD or PH, according to the recently updated CONsensus-based Standards for the selection of health Measurement in INstruments (COSMIN) guidelines [12]. With the outcome of this systematic review, we aim to provide clear and evidence-based recommendations for the most suitable available self-reported PROM instrument on dyspnoea for these patients, and to identify gaps in knowledge on the measurement properties.

Methods

Prior to identification of articles and data extraction, the protocol of this systematic review was registered with PROSPERO, the international prospective register of systematic reviews (registration no. CRD42020150519). Furthermore this review follows the COSMIN guidelines for systematic reviews of patient-reported outcomes [12], and adheres to the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) 2020 statement [13].

We first performed an exploratory literature search without date limits in both PubMed and Embase in July 2019. This literature search was assisted by our medical librarian and was designed to identify relevant PROMs, with an English version available, on dyspnoea in our target population. A PROM was included as possibly relevant if no help or specific instruction in a supervised setting was necessary and if it was applied in adult patients with CTD, ILD or PH. In total, our review team agreed upon 20 possibly relevant PROMs regarding dyspnoea in our population of interest for use in the systematic literature search. The exact search terms, search results and a list of these 20 PROMs is included in supplementary file S1.

Literature search strategy; identification of studies on measurement properties of selected PROMs

A systematic literature search without date limits was performed in both PubMed and Embase, undertaken in July 2019 and last updated in March 2022 assisted by our medical librarian. This search was constructed to identify all studies that assessed measurement properties of the 20 selected PROMs. This search included three overall concepts: 1) all identified possibly relevant PROMs and their synonyms; 2) dyspnoea and 3) measurement properties. We used a previously designed, highly sensitive filter for finding studies on measurement properties of measurement instruments for our search in PubMed [14]. For Embase the same highly sensitive filter was adapted to the applicable format with the help of our librarian, and was comparable to the adapted highly sensitive PubMed filter for Embase described by TERWEE *et al.* [14]. The exact terms and complete search strategy are included in supplementary file S2.

Eligibility criteria

Published original full-text publications were considered eligible for inclusion if they aimed to assess measurement properties of one or more of the selected PROMs regarding dyspnoea in adult patients with CTD, ILD or PH. PROMs with identified studies on their measurement properties were evaluated in full length (item by item) by the review team. Content validity (face validity) for each PROM with retrieved publications on at least one measurement property was discussed in a consensus meeting. We deemed a PROM relevant for assessing dyspnoea in our target population when $\geq 50\%$ of the items of the PROM or subscale of PROM reflected the construct “dyspnoea”, and the PROM was self-reported (*i.e.* no help or specific instruction in a supervised setting was necessary). Only studies including $\geq 50\%$ of patients with CTD, ILD or PH, or reporting these patients’ subgroup results separately, were included. We excluded publications that only used the identified PROMs as an isolated outcome measurement instrument, *e.g.* in randomised controlled trials or publications in which the PROM was used in a validation study of another instrument (*e.g.* PROMs not included in our review). Furthermore we excluded publications in any other

languages than English, unpublished material, conference abstracts, poster presentations, case reports, editorials, letters or reviews.

Data extraction

Identified publications were imported in Endnote (Clarivate, Philadelphia, PA, USA) and after removal of duplicates publications were entered into the online reviewing tool Rayyan (<https://www.rayyan.ai/>). Inclusion and exclusion criteria were independently applied by the same two reviewers (JL and CE) to titles and abstracts of the hits retrieved. In cases of discrepancy between the two reviewers, articles were read in full to decide on eligibility for inclusion. In case of disagreement, a third independent reviewer (MV) was asked to decide on possible inclusion. Furthermore, all references of the included publications were checked to search for additional studies. This resulted in the final list of publications for inclusion in the review.

Data on the included publications and the included PROMs were extracted by two authors independently (JL and CE). Study characteristics included patient age, gender, disease, disease duration, language of PROM used, country of study and study setting. PROM characteristics (*e.g.* items, scoring, feasibility and availability) were extracted from the questionnaires, background literature or user manuals. Data concerning instrument description and validation were collected using forms based on those suggested in the COSMIN user manual [15].

Measurement properties

According to the COSMIN methodology, eight measurement properties need to be assessed in order to determine the quality of a PROM, summarised in supplementary file S3, table A. However, because no gold standard exists for dyspnoea, criterion validity was not evaluated in our systematic review. For hypotheses testing for construct validity and responsiveness, we formulated a set of hypotheses about the expected direction and magnitude of the correlations between the PROM of interest and other PROMs, other constructs and other outcome measures, and of mean differences in scores between groups (supplementary file S3, table B).

Evaluation of measurement properties and grading of quality of evidence

The following four steps were completed. Of note: one publication can contain multiple studies on different measurement properties; therefore, each evaluation of a measurement property is indicated as “single study”.

- 1) To evaluate the methodological quality, each single study on a measurement property was assessed and rated using the COSMIN risk of bias checklist individually by two raters independently [16]. Studies were graded as “very good”, “adequate”, “doubtful”, “inadequate” or “not assessed” for the methodology used to assess each measurement property. In this checklist for each measurement property, a different set of standards is evaluated and a “worst score counts” principal is applied.
- 2) Based on the results of each single study, the performance of the PROM on a measurement property was rated against the updated criteria for good measurement properties, rated as “sufficient”, “insufficient” or “indeterminate” [15]. These criteria are summarised in supplementary file S4.
- 3) Individual ratings from step 2 for each of the included studies were pooled to provide an overall quality assessment per measurement property of each PROM. When pooling results from multiple studies, a measurement property of a PROM was deemed to have sufficient or insufficient overall quality if $\geq 75\%$ of individual studies were graded as “sufficient” or “insufficient”, respectively; inconsistent quality if $< 75\%$ of the results were graded as either “sufficient” or “insufficient”; and indeterminate quality if all individual studies were rated “indeterminate”.
- 4) The evidence was summarised per measurement property, per PROM in a summary of findings table. The overall quality of evidence per PROM per measurement property (taking all data of included studies in account) was graded based on the Grading of Recommendations Assessment and Evaluation (GRADE) approach for systematic reviews and clinical trials [17]. In this approach the quality of evidence was determined as “high”, “moderate”, “low” or “very low”, and subsequently downgraded for 1) the risk of bias, 2) inconsistency, 3) imprecision or 4) indirect results (supplementary file S5) based on the consensus of two reviewers (CE and JL).

To come to a final recommendation, PROMs were categorised as level A, B or C, as follows:

- A. PROMs with at least low-quality evidence for sufficient internal consistency. Level A instruments are recommended for use and results obtained with these PROMs can be trusted.
- B. PROMs not categorised as A or C. These PROMs have potential to be recommended for use, but require further research to assess their quality.
- C. PROMs with high-quality evidence for an insufficient measurement property. PROMs categorised as C should not be recommended for further use.

Results

We identified 20 possibly relevant PROMs on dyspnoea in our target population, summarised in supplementary file S1. After careful consideration we decided to exclude publications evaluating measurement properties of the Breathlessness Beliefs Questionnaire, Cambridge Pulmonary Hypertension Outcome Review, Chronic Respiratory Questionnaire and St George’s Respiratory Questionnaire because they are not dyspnoea specific; <50% of the included questions addressed dyspnoea or no separate dyspnoea subscale was present/evaluated. The search strategy did not yield publications on measurement properties of six of the 20 initially identified PROMs in our patient population: the Numerical Rating Scale, London Chest Activity of Daily Living scale, Scleroderma Burden Index, Pulmonary Functional Status and Dyspnea Questionnaire-modified, Idiopathic Pulmonary Fibrosis PROM-Dyspnoea and the Multidimensional Dyspnea Profile. As a result we assessed 26 publications evaluating 10 PROMs assessing dyspnoea in patients with CTD, ILD or PH [18–43]. A summary of the literature search results is presented in figure 1.

Of the 10 evaluated PROMS, the Dyspnoea-12 (D12) was initially developed in patients with ILD [44], EXACT-Respiratory Symptoms Idiopathic Pulmonary Fibrosis (ERS-IPF) was developed for patients with IPF, and the King’s Brief Interstitial Lung Disease Health Status Questionnaire (KBILD) was developed for patients with ILD, including IPF and CTD-associated ILD [18, 22]. None of the included PROMs were specifically developed for patients with PH.

PROM summary and feasibility

A summary of the PROM characteristics is presented in table 1. In total we evaluated eight PROMs and two breathlessness subscales (the KBILD breathlessness and activities subscale (KBILD-B) and the ERS-IPF breathlessness subscale (ERS-IPF-B)). Included PROMs evaluated dyspnoea ranging from the day of evaluation up to the last 2 weeks. Most PROMs are easy to use and complete, as documented in the retrieved publications. Translated versions other than English language are available for the following

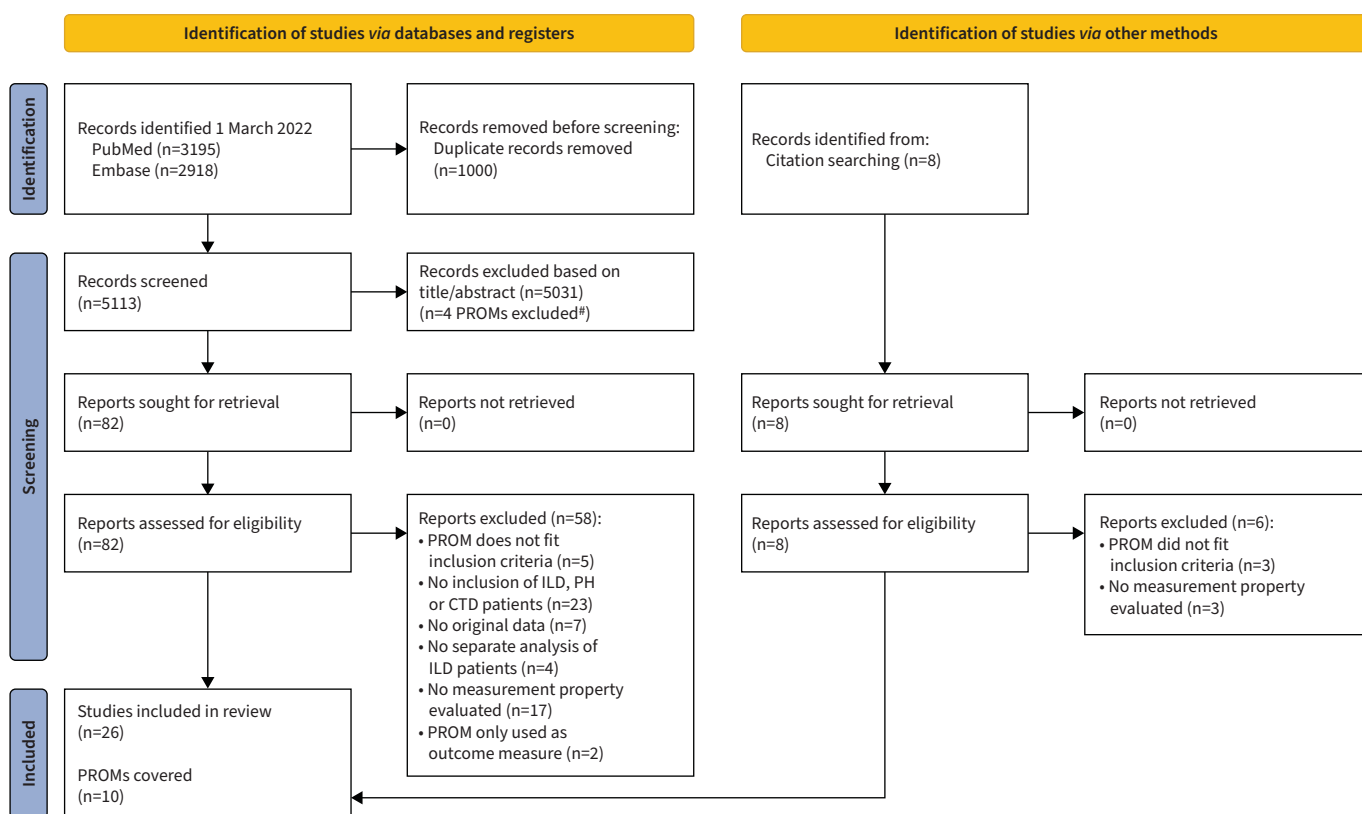


FIGURE 1 Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) 2020 adapted flowchart. #: n=4 patient-reported outcome measures (PROMs) excluded in screening phase, after content evaluation of PROM. ILD: interstitial lung disease; PH: pulmonary hypertension; CTD: connective tissue disease.

TABLE 1 Characteristics of included PROMs

PROM	PROM description	Purpose	Developed for	Scales, items, scoring and recall time	Feasibility, availability and links to more info	Reference
D12	12-item PROM; provides a unidimensional measure that reflects both “physical” (7 items) and “affective” (5 items) aspects of dyspnoea	To measure current level of dyspnoea severity	COPD, ILD and HF	Rating: 12 terms, 4-point Likert scale, anchors “none (0)–severe (3)” Scoring: 0–36, higher scores correspond to greater severity of breathlessness Recall time: “these days”	Completion time: 5–10 min More info: https://eprovide.mapi-trust.org/instruments/dyspnea-12	[44, 45, 53]
ERS-IPF-B	5-item subscale on breathlessness from the IPF-specific version of the ERS (an 11-item daily diary for assessing respiratory symptoms)	To assess the effect of treatment on the severity of respiratory symptoms	COPD	Rating: 5- and 6-point Likert scale Scoring: 0–17 for breathlessness, higher scores represent worse breathlessness Recall time: Today	Completion time: Not reported More info: www.exactproinitiative.com	[22]
FACIT-D	10-item short form; list of common tasks	Describe severity of dyspnoea when completing tasks and the level of difficulty in completing these tasks due to dyspnoea	COPD	Rating: 4-point Likert scale Scoring: Creates two scores: dyspnoea and dyspnoea-related functional limitation, higher scores represent worse dyspnoea or functional limitation Recall time: past 7 days	Completion time: 5–10 min [#] More info: www.facit.org/measures/FACIT-dyspnea	[46, 54]
KBILD-B	Four items on breathlessness and activities; subscale from the 15-item KBILD, which is a self-completed health status questionnaire	Assess the health status of ILD patients	ILD	Rating: 7-point Likert scale Scoring: Domain and total score ranges 0–100; 100 represents best health status Recall time: last 2 weeks	Completion time: 5–7 min [¶] More info: www.kbild.com/	[18, 55]
MRC	Patients classified into one of five MRC grades based on their perceived level of disability	To categorise perceived level of disability	Chronic bronchitis	Rating: 5-point numeric rating scale Scoring: Higher score indicates more impairment Recall time: Current situation	Completion time: Not reported More info: https://eprovide.mapi-trust.org/instruments/modified-medical-research-council-dyspnea-scale	[47]
NS	Dyspnoea scale based on estimates of oxygen consumption for daily activities	Describe degree of dyspnoea associated with the performance of 10 distinct tasks	Not clear	Rating: 10 tasks rated on a 5-point scale Scoring: 0–116, higher score indicating more dyspnoea Recall time: Current situation	Completion time: Not reported More info: Not reported	[31]

Continued

TABLE 1 Continued

PROM	PROM description	Purpose	Developed for	Scales, items, scoring and recall time	Feasibility, availability and links to more info	Reference
OCD	VAS that places a list of activities along a 100 mm line according to their oxygen cost	Identify level of activity at which the patient experiences shortness of breath	Patients with airway obstruction/pulmonary infiltration	Rating: One vertically oriented VAS scale of 100 mm, result expressed as distance from zero to the indicated point Scoring: Higher number represents better function Recall time: Current situation	Completion time: Not reported More info: Not reported	[48]
PROMIS-DS	Evaluation of dyspnoea severity associated with 10 common tasks (part 1 of FACIT-D)	Identify level of dyspnoea severity	COPD	Rating: 4-point Likert scale Scoring: Higher score represents higher levels of dyspnoea severity Recall time: Past 7 days	Completion time: <5 min More info: www.facit.org/measures/PROMIS-SF-v1.0-Dyspnea-Severity-10a	[46]
UCSD-SOBQ	24-item instrument; 21 items rate the severity of dyspnoea while doing various activities, one item on how limiting dyspnoea is and two items on fear associated with overexerting or dyspnoea itself	Assess dyspnoea associated with activities of daily living	COPD, cystic fibrosis and post lung transplantation	Rating: 5-point Likert scale Scoring: Simple summation score yield 0–120, higher score corresponding to greater impairment Recall time: Past week	Completion time: <5 min [#] More info: https://eprovide.mapi-trust.org/instruments/the-university-of-california-san-diego-shortness-of-breath-questionnaire	[49, 57]
VAS-D	VAS for breathlessness	Breathlessness on a day-to day basis	Not reported	Rating: One scale from 0–10 Scoring: Higher number represents worse dyspnoea Recall time: Current situation	Completion time: Not reported More info: Not reported	[41]

PROM: patient-reported outcome measure; D12: Dyspnea-12 Questionnaire; ILD: interstitial lung disease; HF: heart failure; ERS-IPF-B: EXACT-Respiratory Symptoms Idiopathic Pulmonary Fibrosis breathlessness subscale; IPF: idiopathic pulmonary fibrosis; FACIT-DS: Functional Assessment of Chronic Illness Therapy-Dyspnea short form; KBILD-B: King's Brief Interstitial Lung Disease Health Status Questionnaire breathlessness and activities subscale; MRC: Medical Research Council dyspnoea scale; NS: New Dyspnea Scale; OCD: oxygen cost diagram; VAS: visual analogue scale; PROMIS-DS: Patient-Reported Outcomes Measurement Information System-Dyspnea Severity short form; UCSD-SOBQ: University of California San Diego Shortness of Breath Questionnaire; VAS-D: visual analogue scale-dyspnoea. [#]: not assessed in ILD, pulmonary hypertension or connective tissue disease population, results retrieved from populations with other cardio-respiratory diseases; [¶]: not evaluated for specific subscale, time to complete reported for total PROM.

PROMs: D12 (six translations [50–53]), Functional Assessment of Chronic Illness Therapy-Dyspnea (FACIT-D) (>10 translations [54]), KBILD (>10 translations [55]), Medical Research Council dyspnoea scale (MRC) (12 translations [47]), oxygen cost diagram (OCD) (one translation in Spanish [56]), Patient-Reported Outcomes Measurement Information System-Dyspnea Severity short form (PROMIS-DS) (>10 available translations [54]) and University of California San Diego Shortness of Breath Questionnaire (UCSD-SOBQ) (>10 translations [57]).

The MRC is free for research and commercial use with appropriate credit; use of the ERS-IPF, KBILD, UCSD-SOBQ, PROMIS-DS and FACIT-DS is dependent upon the setting (free for use in academic or clinical setting, fees applicable in case of commercial use). No information on user fees is reported for the D12; the questionnaire is provided in the appendix of the original article. For OCD, the New Dyspnea Scale (NS) and the visual analogue scale-dyspnoea (VAS-D), no information is available.

Study summary

An overview of the included publications evaluating measurement properties is provided in table 2. We identified 26 publications, describing 27 separate patient samples. Studies were performed in 26 different countries. Language of PROMs used was not explicitly stated in 15 of the included publications, but most frequently involved English. Patient numbers included in individual studies varied between 11 and 1933.

Most of the studies included patients with ILD (including CTD-associated ILD) or IPF (22 publications on a total of nine PROMs, n=4088 patients). Two publications included analyses of PAH patients; one publication evaluated the D12 in two PAH patient samples (n=201, of which n=41 had CTD-associated PH) [21], while the other evaluated the UCSD-SOBQ and VAS for use in systemic sclerosis (SSc) including associated PAH (53 of 179 patients had associated PH) [38]. SSc was the only CTD without associated ILD or PH in which measurement properties of a dyspnoea-specific PROM were reported. Two publications described performance of the FACIT-D PROM in patients with SSc without associated PH or ILD (n=173) [30, 39].

Risk of bias assessment; methodological quality

The results of the individual risk of bias in each single study per measurement property per PROM are summarised in supplementary files S6, S7 and S8. Studies were mostly graded to be very good or adequate for the methodological quality used to assess measurement properties, except for seven of 69 single studies. Four studies had doubtful/inadequate methodological quality on reliability because of a high risk of recall bias [21, 22, 28, 29]. Two studies had doubtful methodological quality for hypothesis testing on construct validity due to the inclusion of a low number of patients [33] or removal of outliers without specification [37]. Lastly, one study had doubtful methodological quality on responsiveness due to use of an inadequate statistical method [26].

Measurement property assessment and level of evidence

The individual ratings for reported results of measurement properties in each study and overall GRADE levels of evidence per measurement property after pooling of results per PROM are provided in supplementary files S6, S7 and S8. No studies evaluated structural validity or cross-cultural validity/measurement variance. Only one study assessed measurement error of the UCSD-SOBQ [28]. For construct validity and responsiveness, after pooling of individual study results, the overall ratings were either inconclusive or insufficient for all PROMs evaluated.

Summary of findings

A summary of findings per measurement property per PROM is provided in tables 3–5. According to the COSMIN methodology, PROMs with any level of content validity and at least low-quality evidence on internal consistency can be recommended for use when no insufficient measurement properties with high-quality evidence are present.

Therefore D12, ERS-IPF-B, KBILD-B and UCSD-SOBQ can be categorised as a level A recommendations, indicating that these PROMs, or subscales respectively, can be recommended for evaluating dyspnoea in patients with ILD including CTD-associated ILD, and that results obtained with these PROMs can be trusted. We reached this same conclusion regarding the D12 for use in patients with PH, including CTD-associated PH.

All other PROMs included in this review, for which information on internal consistency is lacking, receive a level B recommendation, indicating that they have potential to be recommended for use, but require further research to assess their quality.

TABLE 2 Characteristics of included publications

Reference	PROM(s)	Subjects (n)	Patient population	Age (years)	Women (%)	Disease duration (years)	Language PROM used	Country where study performed	Setting
BIRRING <i>et al.</i> 2021 [23]	KBILD-B	663	Progressive fibrosing ILD (100%)	65.2±9.7	46	Not stated	Not stated	Argentina, Belgium, Canada, Chile, China, France, Germany, Italy, Japan, Republic Korea, Poland, Russian Federation, Spain, UK, USA	OPC
CHEN <i>et al.</i> 2021 [26]	UCSD-SOBQ	1933	ILD (100%): 539 IPF 701 CTD-ILD 344 UC-ILD 151 HP 198 other ILD	62±13	45	Not stated	English/French	Canada	OPC
KIM <i>et al.</i> 2021 [27]	MRC UCSD-SOBQ	238	IPF (100%)	72.6±7.6 (n=231)	28.2	3.19 (2.8) (n=194)	English	UK	OPC/home
PRIOR <i>et al.</i> 2019 [24] [#]	KBILD-B	150	IPF (100%)	72.9±6.2	18.7	0.5 (range 0–9.3)	Danish	Denmark	OPC
PRIOR <i>et al.</i> 2020 [42] [#]	KBILD-B UCSD-SOBQ	150	IPF (100%)	72.9±6.2	18.7	0.5 (IQR 0–21)	Danish	Denmark	OPC
SILVA <i>et al.</i> 2021 [28]	UCSD-SOBQ MRC	30	ILD (100%): 18 IPF 5 CTD-PF 4 particle-induced PF 3 NSIP	59±10	50	Not stated	Brazilian, Portuguese version UCSD	Brazil	OPC
TOPCU <i>et al.</i> 2021 [36]	VAS-D MRC	39	ILD (100%): 14 RA-ILD 25 CTD-ILD	60±7 53.6±10	71 92	10 (IQR 10) 9.7 (IQR 10)	Not stated	Turkey	OPC
WAPENAAR <i>et al.</i> 2017 [25]	KBILD-B	176	ILD 100% 108 IPF 68 other ILD	66.8±9.6	39.2	Not stated	Dutch, French, Italian, Swedish	France, The Netherlands, Italy, Sweden	OPC/home
SWIGRIS <i>et al.</i> 2010 [19]	D12 UCSD-SOBQ	48	100% CTD-ILD	52.8±19.7	74	Not stated	English	USA	OPC
YORKE and ARMSTRONG 2014 [21] [¶]	D12	176	PAH: 86 IPAH 49 CH-PAH 41 CTD-PAH	54.3±14.4	70.5	Not stated	English	UK, Ireland	OPC/home
YORKE and ARMSTRONG 2014 [21] [¶]	D12	25	PAH	Not stated	Not stated	Not stated	Not stated	UK, Ireland	OPC/home
YORKE <i>et al.</i> 2011 [20]	D12 MRC	101	ILD (100%)	67±10.9	70	Not stated	Not stated	England	OPC
BACCI <i>et al.</i> 2018 [22]	ERS-IPF-B	168	IPF (100%)	67.8±6.62	23.8	<5 years	Not stated	Australia, Canada, Israel, Peru, South Korea, USA	OPC
HINCHCLIFF <i>et al.</i> 2011 [30]	FACIT-DS MRC	73	SSc (100%)	51 (range 22–72)	84	7 (range 0–45)	Not stated	USA	OPC

Continued

TABLE 2 Continued

Reference	PROM(s)	Subjects (n)	Patient population	Age (years)	Women (%)	Disease duration (years)	Language PROM used	Country where study performed	Setting
HINCHCLIFF <i>et al.</i> 2015 [39]	FACIT-DS	100	SSc (100%)	52 (range 27–71)	90	5 (range 0–35) (from onset Raynaud)	Not stated	USA (Chicago)	OPC
NOLAN <i>et al.</i> 2019 [40]	KBILD-B	209	ILD (100%)	70±10	39	Not stated	English	UK	OPC rehabilitation programme
PATEL <i>et al.</i> 2012 [18]	KBILD-B	173	ILD (100%)	60±13	60	Not stated	English	UK	OPC
KHADAWARDI <i>et al.</i> 2017 [43]	MRC	115	ILD (100%): 50 IPF 65 non-IPF-ILD	70±10 65±13	56	Maximum 5 years	Not stated	Canada	OPC
BADDINI <i>et al.</i> 2002 [31]	NS OCD MRC	30	IPF (100%)	59±2	40	Not stated	Not stated	Brazil	OPC
TZANAKIS <i>et al.</i> 2005 [32]	OCD MRC	25	IPF (100%)	66±11	16	2.6 (1.7)	Not stated	Greece	OPC
DE JESÚS-BERRIOS <i>et al.</i> 2015 [33]	OCD MRC	66	ILD (17%)	52 (range 25–70)	54	Not stated	Spanish	Puerto Rico	OPC
YOUNT <i>et al.</i> 2016 [29]	PROMIS-DS	220	IPF (100%)	61±5.6	29.6	Not stated	Not stated	USA	Home/online survey
CHUNG <i>et al.</i> 2013 [38]	UCSD-SOBQ VAS-D	179	SSc (100%): 53 SSc-PH 126 SSc	56±11	87	10 (11)	Not stated	USA	Not reported
SWIGRIS <i>et al.</i> 2012 [37]	UCSD-SOBQ	180	IPF (100%)	69±9	17	2 (1.9)	Not stated	USA, Canada	OPC
YATES <i>et al.</i> 2018 [41]	VAS-D KBILD-B	64	ILD (100%)	66±14	41	Not stated	English	UK	Not reported
MANALI <i>et al.</i> 2010 [34]	MRC	25	IPF (100%)	68±8.3	68	Not stated	Not stated	Greece	OPC
MAHLER <i>et al.</i> 1988 [35]	MRC OCD	153	ILD (15%)	57±15	36	Not stated	Not stated	USA	OPC

Age presented as years (mean±SD) unless otherwise stated. PROM: patient-reported outcome measure; KBILD-B; King's Brief Interstitial Lung Disease Health Status Questionnaire breathlessness and activities subscale; ILD: interstitial lung disease; OPC: outpatient clinic; UCSD-SOBQ; University of California San Diego Shortness of Breath Questionnaire; IPF: idiopathic pulmonary disease; CTD: connective tissue disease; UC: unclassifiable; HP: hypersensitivity pneumonitis; MRC: Medical Research Council dyspnoea scale; PF: pulmonary fibrosis; NSIP: nonspecific interstitial pneumonia; VAS-D: visual analogue scale-dyspnoea; RA: rheumatoid arthritis; PAH: pulmonary arterial hypertension; CH: congenital heart disease; D12: Dyspnea 12 Questionnaire; ERS-IPF-B: EXACT-Respiratory Symptoms Idiopathic Pulmonary Hypertension breathlessness subscale; FACIT-DS; Functional Assessment of Chronic Illness Therapy-Dyspnea short form; IPAH: idiopathic pulmonary hypertension; SSc: systemic sclerosis; NS: New Dyspnea Scale; OCD: oxygen cost diagram; PROMIS-DS: Patient-Reported Outcomes Measurement Information System-Dyspnea Severity short form. †: same patient population, different measurement properties evaluated; ‡: test-retest cohort evaluated in 25 different PAH patients.

TABLE 3 Summary of pooled findings on the quality of evidence on measurement properties of included PROMs in ILD, including CTD-associated ILD

PROM	Internal consistency	Reliability	Measurement error	Construct validity	Responsiveness	Recommendation
D12	+/high	+/moderate	x	±/moderate	x	A
ERS-IPF-B	+/high	+/very low	x	±/moderate	?/moderate	A
FACIT-DS	x	x	x	x	x	x
KBILD-B	+/high	+/high	x	–/moderate	±/moderate	A
MRC	x	x	x	±/moderate	±/moderate	B
NS	x	x	x	–/very low	x	B
OCD	x	x	x	±/low	x	B
PROMIS-DS	x	+/low	x	x	x	B
UCSD-SOBQ	+/high	+/very low	+/very low	±/moderate	±/moderate	A
VAS-D	x	x	x	–/very low	±/moderate	B

Quality of evidence was categorised as high, moderate, low or very low. Overall rating of PROM quality categorised as sufficient (+), insufficient (–), inconsistent (±) or indeterminate (?). The symbol x indicates that the measurement property was not evaluated in the study. PROM: patient-reported outcome measure; ILD: interstitial lung disease; CTD: connective tissue disease; D12: Dyspnea 12 Questionnaire; ERS-IPF-B: EXACT-Respiratory Symptoms Idiopathic Pulmonary Hypertension breathlessness subscale; FACIT-DS: Functional Assessment of Chronic Illness Therapy-Dyspnea short form; KBILD-B: King's Brief Interstitial Lung Disease Health Status Questionnaire breathlessness and activities subscale; MRC: Medical Research Council dyspnoea scale; NS: New Dyspnea Scale; OCD: oxygen cost diagram; PROMIS-DS: Patient-Reported Outcomes Measurement Information System-Dyspnea Severity short form; UCSD-SOBQ: University of California San Diego Shortness of Breath Questionnaire; VAS-D: visual analogue scale-dyspnoea.

TABLE 4 Summary of pooled findings on the quality of evidence on measurement properties of included PROMs in PH, including CTD-associated PH

PROM	Internal consistency	Reliability	Measurement error	Construct validity	Responsiveness	Recommendation
D12	+/high	+/very low	x	–/moderate	x	A
UCSD-SOBQ	x	x	x	±/low	±/very low	B
VAS	x	x	x	±/very low	–/low	B

Quality of evidence was categorised as high, moderate, low or very low. Overall rating of PROM quality categorised as sufficient (+), insufficient (–) or inconsistent (±). The symbol x indicates that the measurement property was not evaluated in the study. PROM: patient-reported outcome measure; PH: pulmonary hypertension; CTD: connective tissue disease; D12: Dyspnea 12 Questionnaire; UCSD-SOBQ: University of California San Diego Shortness of Breath Questionnaire; VAS: visual analogue scale.

TABLE 5 Summary of pooled findings on the rating and quality of evidence on measurement properties of included PROMS in CTD (SSc patients only)

PROM	Internal consistency	Reliability	Measurement error	Construct validity	Responsiveness	Recommendation
FACIT-DS	x	x	x	±/low	±/low	B
MRC	x	x	x	+/low	x	B
UCSD-SOBQ	x	x	x	±/low	±/low	B
VAS-D	x	x	x	±/low	–/moderate	B

Quality of evidence was categorised as high, moderate, low or very low. Overall rating of PROM quality categorised as sufficient (+), insufficient (–) or inconsistent (±). The symbol x indicates that the measurement property was not evaluated in the study. PROM: patient-reported outcome measure; CTD: connective tissue disease; SSc: systemic sclerosis; D12: Dyspnea 12 Questionnaire; FACIT-DS: Functional Assessment of Chronic Illness Therapy-Dyspnea short form; MRC: Medical Research Council dyspnoea scale; UCSD-SOBQ: University of California San Diego Shortness of Breath Questionnaire; VAS-D: visual analogue scale-dyspnoea.

Lastly it must be noted that none of the included PROMs in this review have sufficient ratings for all evaluated measurement properties.

Recommendation

- Of the PROMs with a level A recommendation in patients with ILD (including CTD-associated ILD), KBILD-B is the only PROM with high-quality evidence for reliability; the level of evidence for reliability of D12 is moderate, and for ERS-IPF-B and UCSD-SOBQ is very low. All four PROMs have moderate-quality evidence on construct validity and responsiveness. Therefore, based on available evidence and despite limitations, given that evidence on other important measurement properties is lacking, the best PROM on dyspnoea in patients with ILD (including CTD-associated ILD) is the KBILD-B.
- The D12 is the only included dyspnoea-specific PROM with a level A recommendation for use in patients with PH (including CTD-associated PH).

Discussion

To our knowledge, this is the first study that systematically reviews measurement properties of dyspnoea-specific PROMs in patients with ILD, PH or CTD using the COSMIN methodology.

We identified 10 relevant PROMs with information on at least one measurement property for this target population. Based on current evidence, four of these receive a level A recommendation using the COSMIN criteria; therefore, ERS-IPF-B, KBILD-B and UCSD-SOBQ can be recommended for use in ILD patients, and D12 in both ILD and PH patients, with or without associated CTD.

This study highlights gaps in current knowledge relating to currently used PROMs and the quality assessment of commonly used dyspnoea PROMs in these populations. The results of our systematic literature search demonstrated that for six of the initially identified possibly relevant PROMs, no studies could be found that examined their measurement properties, indicating that these PROMs cannot be recommended for the evaluation of dyspnoea in this population. Furthermore, among the PROMs evaluated, no studies adequately assessed the structural validity or cross-cultural validity/measurement invariance and only one study, with inadequate methodology and a low level of evidence, evaluated measurement error. Additionally, most of the included PROMs in this systematic review had inconsistent scores and moderate/low-quality evidence for construct validity and responsiveness. These essential properties need further exploration to strengthen future recommendations. Lastly, this systematic review indicates that the majority of the studies identified in our search included patients with ILD, with or without CTDs, and very few studies included patients with PH. For patients with CTD without associated ILD or PH, only the FACIT-D was evaluated in patients with SSc. Currently only low-quality evidence on construct validity and responsiveness is available and information on internal consistency was not evaluated for this PROM in these patients. We conclude that this PROM shows potential for use in patients with SSc; however, further research is necessary to be able to make an evidence-based recommendation.

Our study has two major strengths. First, we performed an extensive literature search, assisted by an experienced medical librarian based on the recommendations developed by TERWEE *et al.* [14] to identify relevant studies. Furthermore we used a well-developed consensus-based methodology reviewing the evidence on measurement properties [15].

Our study also has important limitations. First, we acknowledge that decisions regarding content validity of PROMs included in this review could be questioned. However our research team combined experience from different expertise areas, and we applied a conservative threshold of 50% or more of items on dyspnoea in the PROMs to demonstrate adequate content validity. Therefore, we feel that this methodology allowed us to identify the most appropriate PROMs of dyspnoea in this population. Second, we recognise that the exploratory search on PROMs or subscales reflecting dyspnoea in our target population is not exhaustive. We only considered PROMs for which an English version was available, which could have excluded possibly relevant PROMs or publications on measurement properties available in other languages. However, several of the PROMs we identified in this study were also recently described by the American Thoracic Society research statement as important frequently used domain-specific patient-centred outcomes for dyspnoea in ILD [58]. Third, we used the expertise present in our research team to formulate a hypothesis regarding construct validity and responsiveness regarding correlations with specific outcomes used in clinical practice. We recognise different thresholds or less strict hypotheses could lead to different conclusions on sufficiency of the evidence. Furthermore, it is important to emphasise that the evaluation regarding quality of individual study methodology using COSMIN guidelines, and overall study quality

using the GRADE approach, especially the grading of inconsistency and indirectness of the evidence, is subjective and open to differences in opinion [12]. We ensured consistency in rating across studies, and thereby minimised variability, by grading studies independently using two authors and resolving any conflicts by discussion with a third team member if needed.

Points for clinical practice

- Currently four dyspnoea-specific PROMs can be recommended for use in patients with ILD (including CTD-associated ILD): D12, ERS-IPF-B, KBILD-B and UCSD-SOBQ.
- The use of D12 is also suitable for patients with PH, including CTD-associated PH; however, evidence is based on a limited amount of evidence.
- There is not enough evidence on measurement properties of dyspnoea-specific PROMs to give a recommendation for use in patients with CTD without associated ILD or PH.

Conclusion

Research performed with PROMs of poor or unknown quality constitutes a waste of resources and is unethical [59]. Therefore it is crucial to select PROMs with good-quality evidence on the sufficiency of their measurement properties. Clinicians need to be aware of the strengths and weaknesses of available PROMs, particularly when deciding which measure to use in a specific patient population. To summarise, there are four dyspnoea-specific PROMs that can be recommended for use in patients with ILD, including CTD-associated ILD. These are the D12, ERS-IPF-B, KBILD-B and UCSD-SOBQ, of which KBILD-B (despite limitations) can currently be considered the best PROM to evaluate dyspnoea. Of these PROMs, the D12 can also be recommended for use in patients with PH, with or without associated CTD.

Provenance: Submitted article, peer reviewed.

Acknowledgements: We would like to acknowledge O.Y. Chan, information specialist affiliated to the Radboud University Medical Library, Nijmegen, the Netherlands, for assistance on both constructing and performing the systematic literature search.

Data availability statement: The authors confirm that the data supporting the findings of this study are available within the article or its supplementary materials.

Conflict of interest: The authors have no conflicts of interest to report.

References

- 1 Tyndall AJ, Bannert B, Vonk M, *et al.* Causes and risk factors for death in systemic sclerosis: a study from the EULAR Scleroderma Trials and Research (EUSTAR) database. *Ann Rheum Dis* 2010; 69: 1809–1815.
- 2 Elhai M, Meune C, Boubaya M, *et al.* Mapping and predicting mortality from systemic sclerosis. *Ann Rheum Dis* 2017; 76: 1897–1905.
- 3 Jaeger VK, Distler O, Maurer B, *et al.* Functional disability and its predictors in systemic sclerosis: a study from the DeSSciphir project within the EUSTAR group. *Rheumatology (Oxford)* 2018; 57: 441–450.
- 4 Baron M, Sutton E, Hudson M, *et al.* The relationship of dyspnoea to function and quality of life in systemic sclerosis. *Ann Rheum Dis* 2008; 67: 644–650.
- 5 Black N. Patient reported outcome measures could help transform healthcare. *BMJ* 2013; 346: f167.
- 6 Moor CC, van Leuven SI, Wijnsbeek MS, *et al.* Feasibility of online home spirometry in systemic sclerosis-associated interstitial lung disease: a pilot study. *Rheumatology (Oxford)* 2021; 60: 2467–2471.
- 7 Nakshbandi G, Moor CC, Wijnsbeek MS. Home monitoring for patients with ILD and the COVID-19 pandemic. *Lancet Respir Med* 2020; 8: 1172–1174.
- 8 Saketkoo LA, Mittoo S, Huscher D, *et al.* Connective tissue disease related interstitial lung diseases and idiopathic pulmonary fibrosis: provisional core sets of domains and instruments for use in clinical trials. *Thorax* 2014; 69: 428–436.
- 9 Khanna D, Mittoo S, Aggarwal R, *et al.* Connective tissue disease-associated interstitial lung diseases (CTD-ILD) - report from OMERACT CTD-ILD Working Group. *J Rheumatol* 2015; 42: 2168–2171.
- 10 Khanna D, Distler O, Avouac J, *et al.* Measures of response in clinical trials of systemic sclerosis: the Combined Response Index for Systemic Sclerosis (CRISS) and Outcome Measures in Pulmonary Arterial Hypertension related to Systemic Sclerosis (EPOSS). *J Rheumatol* 2009; 36: 2356–2361.

- 11 Distler O, Behrens F, Pittrow D, *et al.* Defining appropriate outcome measures in pulmonary arterial hypertension related to systemic sclerosis: a Delphi consensus study with cluster analysis. *Arthritis Rheum* 2008; 59: 867–875.
- 12 Prinsen CAC, Mokkink LB, Bouter LM, *et al.* COSMIN guideline for systematic reviews of patient-reported outcome measures. *Qual Life Res* 2018; 27: 1147–1157.
- 13 Page MJ, McKenzie JE, Bossuyt PM, *et al.* The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 2021; 372: n71.
- 14 Terwee CB, Jansma EP, Riphagen, II, *et al.* Development of a methodological PubMed search filter for finding studies on measurement properties of measurement instruments. *Qual Life Res* 2009; 18: 1115–1123.
- 15 Mokkink L, Prinsen CAC, Patrick DL *et al.* COSMIN Methodology for Systematic Reviews of Patient-Reported Outcome Measures (PROMs) User Manual version 1.0 February. 2018. https://www.cosmin.nl/wp-content/uploads/COSMIN-syst-review-for-PROMs-manual_version-1_feb-2018.pdf.
- 16 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: 1171–1179.
- 17 Schünemann H, Brozek J, Guyatt G, *et al.* (eds). GRADE Handbook for grading the quality of evidence and the strength of recommendations using the GRADE approach. 2013. <https://gdt.gradeapro.org/app/handbook/handbook.html>
- 18 Patel AS, Siegert RJ, Brignall K, *et al.* The development and validation of the King’s Brief Interstitial Lung Disease (KBILD) health status questionnaire. *Thorax* 2012; 67: 804–810.
- 19 Swigris JJ, Yorke J, Sprunger DB, *et al.* Assessing dyspnea and its impact on patients with connective tissue disease-related interstitial lung disease. *Respir Med* 2010; 104: 1350–1355.
- 20 Yorke J, Swigris J, Russell AM, *et al.* Dyspnea-12 is a valid and reliable measure of breathlessness in patients with interstitial lung disease. *Chest* 2011; 139: 159–164.
- 21 Yorke J, Armstrong I. The assessment of breathlessness in pulmonary arterial hypertension: reliability and validity of the Dyspnoea-12. *Eur J Cardiovasc Nurs* 2014; 13: 506–514.
- 22 Bacci ED, O’Quinn S, Leidy NK, *et al.* Evaluation of a respiratory symptom diary for clinical studies of idiopathic pulmonary fibrosis. *Respir Med* 2018; 134: 130–138.
- 23 Birring SS, Bushnell DM, Baldwin M, *et al.* The psychometric properties of the King’s Brief Interstitial Lung Disease questionnaire and thresholds for meaningful treatment response in patients with progressive fibrosing interstitial lung diseases. *Eur Respir J* 2021; 59: 2101790.
- 24 Prior TS, Hilberg O, Shaker SB, *et al.* Validation of the King’s Brief Interstitial Lung Disease questionnaire in idiopathic pulmonary fibrosis. *BMC Pulm Med* 2019; 19: 255.
- 25 Wapenaar M, Patel AS, Birring SS, *et al.* Translation and validation of the King’s Brief Interstitial Lung Disease (KBILD) questionnaire in French, Italian, Swedish, and Dutch. *Chron Respir Dis* 2017; 14: 140–150.
- 26 Chen T, Tsai APY, Hur SA, *et al.* Validation and minimum important difference of the UCSD Shortness of Breath Questionnaire in fibrotic interstitial lung disease. *Respir Res* 2021; 22: 202.
- 27 Kim JW, Clark A, Birring SS, *et al.* Psychometric properties of patient reported outcome measures in idiopathic pulmonary fibrosis. *Chron Respir Dis* 2021; 18: 14799731211033925.
- 28 Silva H, Mantoani LC, Zamboti CL, *et al.* Validation of the Brazilian Portuguese version of the University of California San Diego Shortness of Breath Questionnaire in patients with interstitial lung disease. *J Bras Pneumol* 2021; 47: e20210172.
- 29 Yount SE, Beaumont JL, Chen SY, *et al.* Health-related quality of life in patients with idiopathic pulmonary fibrosis. *Lung* 2016; 194: 227–234.
- 30 Hinchcliff M, Beaumont JL, Thavarajah K, *et al.* Validity of two new patient-reported outcome measures in systemic sclerosis: Patient-Reported Outcomes Measurement Information System 29-item Health Profile and Functional Assessment of Chronic Illness Therapy-Dyspnea short form. *Arthritis Care Res (Hoboken)* 2011; 63: 1620–1628.
- 31 Baddini Martinez JA, Martinez TY, Lovetro Galhardo FP, *et al.* Dyspnea scales as a measure of health-related quality of life in patients with idiopathic pulmonary fibrosis. *Med Sci Monit* 2002; 8: CR405-10.
- 32 Tzanakis N, Samiou M, Lambiri I, *et al.* Evaluation of health-related quality-of-life and dyspnea scales in patients with idiopathic pulmonary fibrosis. Correlation with pulmonary function tests. *Eur J Intern Med* 2005; 16: 105–112.
- 33 De Jesús-Berrios Y, Santos-Rodríguez RA, Dexter D, *et al.* Usefulness of the culturally adapted oxygen-cost diagram in the assessment of dyspnea in Puerto Rico. *P R Health Sci J* 2015; 34: 14–19.
- 34 Manali ED, Lyberopoulos P, Triantafillidou C, *et al.* MRC chronic dyspnea scale: relationships with cardiopulmonary exercise testing and 6-minute walk test in idiopathic pulmonary fibrosis patients: a prospective study. *BMC Pulm Med* 2010; 10: 32.
- 35 Mahler DA, Wells CK. Evaluation of clinical methods for rating dyspnea. *Chest* 1988; 93: 580–586.
- 36 Topcu A, Mursaloglu HH, Yalcinkaya Y, *et al.* Evaluation of rheumatoid arthritis and connective tissue disease-related interstitial lung disease with pulmonary physiologic test, HRCT, and patient-based measures of dyspnea and functional disability. *Clin Rheumatol* 2021; 40: 3797–3805.

- 37 Swigris JJ, Han M, Vij R, *et al.* The UCSD Shortness of Breath Questionnaire has longitudinal construct validity in idiopathic pulmonary fibrosis. *Respir Med* 2012; 106: 1447–1455.
- 38 Chung L, Chen H, Khanna D, *et al.* Dyspnea assessment and pulmonary hypertension in patients with systemic sclerosis: utility of the University of California, San Diego, Shortness of Breath Questionnaire. *Arthritis Care Res (Hoboken)* 2013; 65: 454–463.
- 39 Hinchcliff ME, Beaumont JL, Carns MA, *et al.* Longitudinal evaluation of PROMIS-29 and FACIT-dyspnea short forms in systemic sclerosis. *J Rheumatol* 2015; 42: 64–72.
- 40 Nolan CM, Birring SS, Maddocks M, *et al.* King's Brief Interstitial Lung Disease questionnaire: responsiveness and minimum clinically important difference. *Eur Respir J* 2019; 54: 1900281.
- 41 Yates H, Adamali HI, Maskell N, *et al.* Visual analogue scales for interstitial lung disease: a prospective validation study. *QJM* 2018; 111: 531–539.
- 42 Prior TS, Hoyer N, Hilberg O, *et al.* Responsiveness and minimal clinically important difference of SGRQ-I and KBILD in idiopathic pulmonary fibrosis. *Respir Res* 2020; 21: 91.
- 43 Khadawardi H, Mura M. A simple dyspnoea scale as part of the assessment to predict outcome across chronic interstitial lung disease. *Respirology* 2017; 22: 501–507.
- 44 Yorke J, Moosavi SH, Shuldham C, *et al.* Quantification of dyspnoea using descriptors: development and initial testing of the Dyspnoea-12. *Thorax* 2010; 65: 21–26.
- 45 Bech TW, Eklund M, Spaak E, *et al.* Feasibility of completing Multidimensional Dyspnea Profile and Dyspnea-12 over the telephone in patients with oxygen-dependent disease. *BMJ Open Respir Res* 2021; 8: e001027.
- 46 Yount SE, Choi SW, Victorson D, *et al.* Brief, valid measures of dyspnea and related functional limitations in chronic obstructive pulmonary disease (COPD). *Value Health* 2011; 14: 307–315.
- 47 Fletcher CM. Modified Medical Research Council Dyspnea Scale (mMRC). 1952. <https://eprovide.mapi-trust.org/instruments/modified-medical-research-council-dyspnea-scale> Date last accessed: 12 April 2022.
- 48 McGavin CR, Artvinli M, Naoe H, *et al.* Dyspnoea, disability, and distance walked: comparison of estimates of exercise performance in respiratory disease. *Br Med J* 1978; 2: 241–243.
- 49 Eakin EG, Resnikoff PM, Prewitt LM, *et al.* Validation of a new dyspnea measure: the UCSD Shortness of Breath Questionnaire. University of California, San Diego. *Chest* 1998; 113: 619–624.
- 50 Sundh J, Ekstrom M. Dyspnoea-12: a translation and linguistic validation study in a Swedish setting. *BMJ Open* 2017; 7: e014490.
- 51 Amado Diago CA, Puente Maestu L, Abascal Bolado B, *et al.* Translation and validation of the multidimensional Dyspnea-12 questionnaire. *Arch Bronconeumol (Engl Ed)* 2018; 54: 74–78.
- 52 Beaumont M, Couturaud F, Jego F, *et al.* Validation of the French version of the London Chest Activity of Daily Living scale and the Dyspnea-12 questionnaire. *Int J Chron Obstruct Pulmon Dis* 2018; 13: 1399–1405.
- 53 Moosavi SH, Shuldham C, Yorke J, *et al.* D-12 described in ePROVIDE. 2010. <https://eprovide.mapi-trust.org/instruments/dyspnea-12> Date last accessed: 12 April 2022.
- 54 FACIT Group. FACIT-Dyspnea Languages. www.facit.org/measure-languages/FACIT-Dyspnea-Languages Date last accessed: 12 April 2022.
- 55 Birring S. King's Brief Interstitial Lung Disease Questionnaire (KBILD). 2012. www.kbild.com Date last accessed: 12 April 2022.
- 56 Serrano J, De Jesús-Berríos Y, Santos RA, *et al.* Adaptation of the Medical Research Council Dyspnea Scale and the Oxygen Cost Diagram for its use in Puerto Rico. *P R Health Sci J* 2007; 26: 135–140.
- 57 Mapi Research Trust. The University of California, San Diego Shortness of Breath Questionnaire (SOBQ) 2022. <https://eprovide.mapi-trust.org/instruments/the-university-of-california-san-diego-shortness-of-breath-questionnaire> Date last accessed: 12 April 2022.
- 58 Aronson KI, Danoff SK, Russell AM, *et al.* Patient-centered outcomes research in interstitial lung disease: an Official American Thoracic Society research statement. *Am J Respir Crit Care Med* 2021; 204: e3–e23.
- 59 Ioannidis JP, Greenland S, Hlatky MA, *et al.* Increasing value and reducing waste in research design, conduct, and analysis. *Lancet* 2014; 383: 166–175.