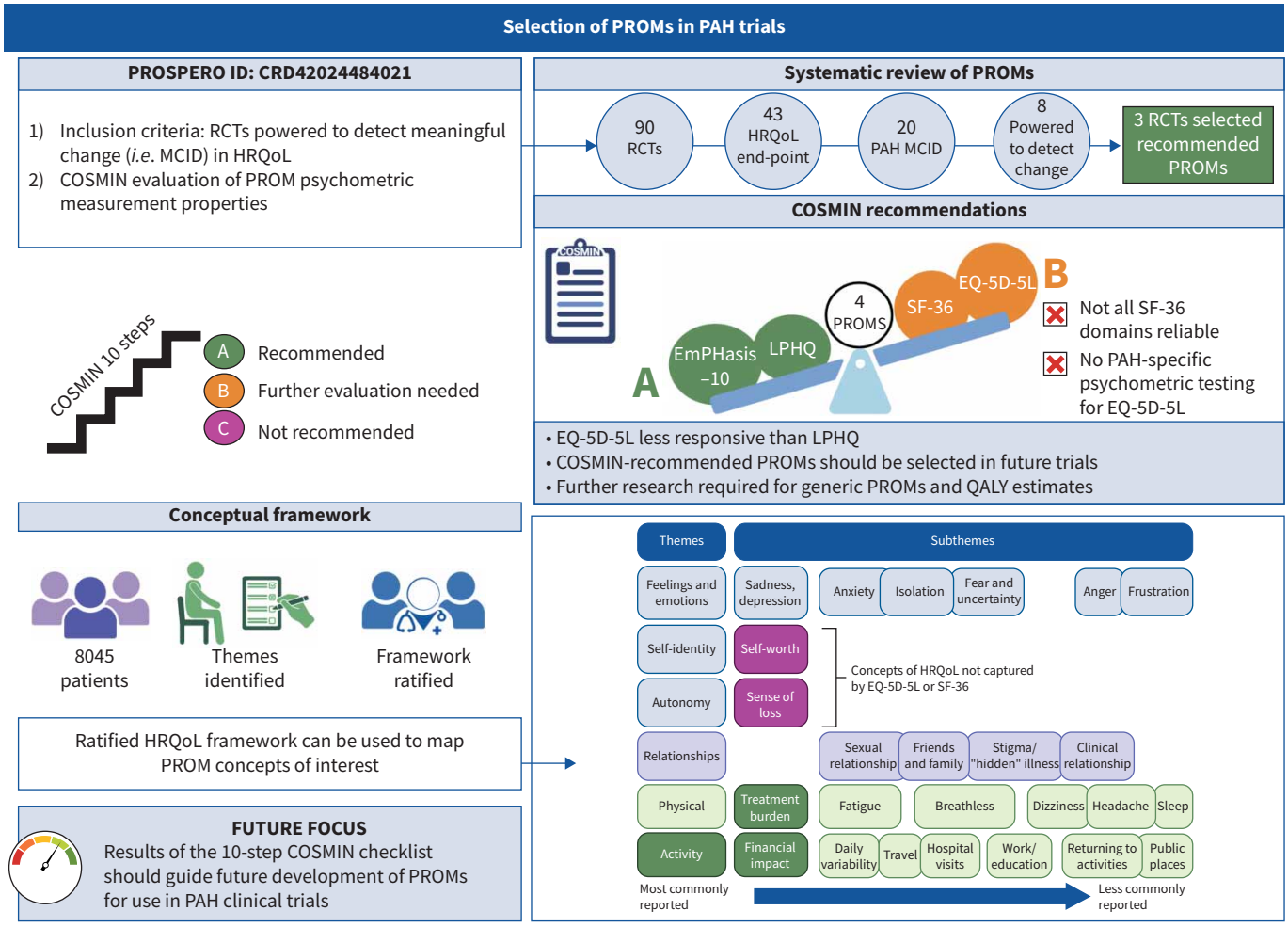
Selection of patient-reported outcome measures in pulmonary arterial hypertension clinical trials: a systematic review, meta-analysis and health-related quality of life framework

Frances Varian, Rebecca Burney, Charlotte Pearson, Ze Ming Goh, Joseph Newman, Gregg Rawlings, Hamza Zafar, David G. Kiely, A.A. Roger Thompson, Robin Condliffe, Mark Toshner, Ciara McCormack, Iain Armstrong, Tessa Peasgood, Jill Carlton and Alexander M.K. Rothman



GRAPHICAL ABSTRACT Selection of patient-reported outcome measures (PROMs) in pulmonary arterial hypertension (PAH) clinical trials. COSMIN: COnsensus-based standards for the Selection of health-Measurement INstruments; EQ-5D-5L: EuroQol-five dimensions-five levels; HRQoL: health-related quality of life; LPHQ: Living with Pulmonary Hypertension Questionnaire; MCID: minimal clinically important difference; QALY: quality-adjusted life-years, RCT: randomised controlled trial; SF-36: 36-Item Short Form Survey. Image created using BioRender.

https://doi.org/10.1183/16000617.0006-2025 Eur Respir Rev 2025; 34: 250006

Selection of patient-reported outcome measures in pulmonary arterial hypertension clinical trials: a systematic review, meta-analysis and health-related quality of life framework

Frances Varian [1,2], Rebecca Burney [1,2], Charlotte Pearson [1,2], Ze Ming Goh [1,2], Joseph Newman [3,4], Gregg Rawlings[5], Hamza Zafar[1,2], David G. Kiely [2], A.A. Roger Thompson [1,2], Robin Condliffe [2], Mark Toshner [3,4], Ciara McCormack [6], Iain Armstrong[2,7], Tessa Peasgood [8], Jill Carlton [8] and Alexander M.K. Rothman [1,2]

[1]Division of Clinical Medicine, School of Medicine and Population Health (formerly Faculty of Medicine, Dentistry and Health), University of Sheffield, Sheffield, UK. [2]Sheffield Pulmonary Vascular Disease Unit, Royal Hallamshire Hospital, Sheffield Teaching Hospitals NHS Foundation Trust, Sheffield, UK. [3]Royal Papworth Hospital, Cambridge, UK. [4]Victor Phillip Dahdaleh Heart and Lung Research Institute, University of Cambridge, Cambridge, UK. [5]Clinical and Applied Psychology Unit, University of Sheffield, Sheffield, UK. [6]Department of Sport Science and Nutrition, Maynooth University, Maynooth, Ireland. [7]Pulmonary Hypertension Association United Kingdom, Sheffield, UK. [8]Sheffield Centre for Health and Related Research (ScHARR), School of Medicine and Population Health, University of Sheffield, Sheffield, UK.

Corresponding author: Alexander M.K. Rothman (a.rothman@sheffield.ac.uk)

Shareable abstract (@ERSpublications)
Only 3 of 90 RCTs have selected COSMIN-recommended PROMs for HRQoL evaluation. EmPHasis-10 and LPHQ can be recommended for use. The ratified conceptual framework can support PROM selection by identifying the HRQoL concepts they are likely to capture. https://bit.ly/3Xlcp58

Cite this article as: Varian F, Burney R, Pearson C, *et al*. Selection of patient-reported outcome measures in pulmonary arterial hypertension clinical trials: a systematic review, meta-analysis and health-related quality of life framework. *Eur Respir Rev* 2025; 34: 250006 [DOI: 10.1183/16000617.0006-2025].

Received: 11 Jan 2025
Accepted: 23 Feb 2025

Abstract

Introduction Health-related quality of life (HRQoL) in pulmonary arterial hypertension (PAH) is valued as an outcome measure by patients, clinicians and regulators. The selection of patient-reported outcome measures (PROMs) for measurement of HRQoL in PAH clinical trials lacks systematic evaluation of their suitability, accuracy and reliability.

Methods We report a systematic review (PROSPERO ID: CRD42024484021) following Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines of PROMs selected in PAH clinical trials. PROM measurement properties were then evaluated according to the 10-step COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) checklist and graded by recommendation for use. Finally, HRQoL was modelled into a conceptual framework using patient interviews and surveys.

Results Screening of 896 records identified 90 randomised controlled trials. 43 trials selected PROMs, of which 20 were sufficiently validated to detect meaningful change. Of these, eight trials were adequately powered, using either EuroQol-five dimensions-five levels (EQ-5D-5L), Short-Form-36 (SF-36) or the Living with Pulmonary Hypertension Questionnaire (LPHQ). The COSMIN evaluation recommended EmPHasis-10 and the LPHQ for use (grade A); whereas, SF-36 and EQ-5D-5L require further study (grade B). A conceptual framework of HRQoL was developed from literature comprising 8045 patients. This framework can be used to visualise the different HRQoL concepts measured by different PROMs.

Conclusion To improve patient-centred research, greater consistency in PROM selection is required. Three of 90 randomised controlled trials have selected COSMIN-recommended PROMs. Whilst the PROMs evaluated require development across the 10 areas of psychometric property measurement, EmPHasis-10 and the LPHQ can be recommended for use. The ratified conceptual framework can further support PROM selection by identifying the HRQoL concepts they are likely to capture.

Background

End-points in randomised controlled trials (RCTs) have traditionally focused on physiological measures, including functional markers such as 6-min walk distance (6MWD) [1, 2]. However, approaches

prioritising clinician-derived end-points [3–5] can undervalue the patient voice. Patient-reported outcome measures (PROMs) are instruments developed to capture and quantify the experience of living with a health condition. Improvement in health-related quality of life (HRQoL) is an important treatment goal for clinicians, regulators and patients, yet it is often not examined in clinical trials [6–11]. Furthermore, significant advancements in the diagnosis and treatment of pulmonary arterial hypertension (PAH) mean people are living longer, with a focus not only on length of life, but also quality. There are many challenges in validating PROMs for accurate measurement of HRQoL. A 10-step checklist for PROM risk of bias was developed by the COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) steering committee [12]. This guidance outlines the systematic evaluation of PROMs and facilitates recommendations about PROM suitability [13, 14]. This has yet to be undertaken for PROMs in pulmonary hypertension (PH), limiting knowledge of appropriate PROM selection in clinical trials.

PROMs can also be used to evaluate the cost-effectiveness of interventions, typically calculated as quality-adjusted life-years (QALYs). To allow such a calculation, PROMs require a value set. Such PROMs are termed a preference-weighted measure (PWM). Value sets are based on the views or preferences of the public and/or patients and vary by country to reflect sociocultural differences [15, 16]. A PWM scores each health state described by the PROM as a single value, or "utility index", on a scale, such that 1 represents full health and 0 represents death. A score below zero indicates a health state considered worse than being dead. The index score of a health state can be combined with time spent in that state to estimate QALYs. QALYs are an important outcome for regulatory and clinical decision-making and therefore dependent upon robust PROMs [17].

PROM selection in clinical trials should follow guidance developed using international Delphi approaches [18] and be evidence-based [14]. Condition-specific PROMs may offer greater sensitivity to changes in HRQoL than generic PROMs, though evidence is limited [11, 19]. The condition-specific PROMs for World Health Organization PH groups (I–V) include the Cambridge Pulmonary Hypertension Outcome Review (CAMPHOR), EmPHasis-10, the Living with Pulmonary Hypertension Questionnaire (LPHQ) and PAH-SYMptoms and imPACT (PAH-SYMPACT) [20, 21]. Guidelines and regulators, such as the US Food and Drug Administration (FDA), recommend that sensitivity to therapeutic change must be interpreted as clinically meaningful (defined as the minimal clinically important difference (MCID)) [11, 13, 18, 22–26]. A generic PROM (*e.g.* Short-Form-36 (SF-36)) may be used, providing the instrument has an MCID validated in the population of interest [14, 18]. MCID interpretation comprises one of 10 steps evaluating measurement properties within the COSMIN risk of bias checklist.

We follow the COSMIN systematic review process to facilitate recommendations of which PROMs should be selected in PAH clinical trials [13, 14]. This process can be further enhanced by identifying which HRQoL concepts are captured by the PROM [27]. Developing a conceptual framework aids visualisation of the important aspects of HRQoL for people living with PH [28, 29]. The content of PROMs can then be compared to this framework to illustrate which aspects of HRQoL are likely to be measured.

Aims and objectives

This is the first systematic review of PROMs selected for use in clinical trials for adults with PAH [20, 30, 31] to 1) evaluate appropriate selection with a valid MCID and 2) compare measurement properties in accordance with the COSMIN checklist, including grading recommendations for use [12, 14, 23–25, 32]. Additionally, to further inform selection of PROMs, we undertake a literature review to develop a conceptual framework which summarises the impact of relevant HRQoL concepts from the perspective of people living with PH.

Methods

Systematic searches

The protocol for the systematic review of PAH RCTs was registered on PROSPERO (CRD42024484021). The COSMIN systematic evaluation is not independently registered. Methodology adhered to the Cochrane Handbook of Systematic Reviews of Interventions [33] and COSMIN guidance [12, 14]. The reporting structure followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement (see supplementary figure E1) and PRISMA-COSMIN outcome measurement instruments (supplementary figure E3 and table E9) [14]. Medline (1980–December 2023) and the Cochrane Library (2002–December 2023) were searched for RCTs evaluating the effectiveness of any clinical intervention in PAH. Inclusion and exclusion criteria are registered on PROSPERO. After removal of duplicates, one author (F. Varian) screened for relevant titles and abstracts before reviewing the full text for eligibility. Where there was uncertainty about the relevance of an article, a second author (J. Newman) reviewed the title and

abstract/main text. A third author was available to adjudicate discrepancies. This process was repeated for the PRISMA-COSMIN search strategy and reporting structure (supplementary figures E2 and E3). PRISMA-COSMIN studies included PH comprising group 1 and group 4 patients to maximise inclusion of PROM psychometric studies. Forward and backward searches were performed on eligible articles for both searches, and citation searching was performed on the identified systematic reviews.

Data extraction

Five authors (F. Varian, R. Burney, C. Pearson, Z.M. Goh and J. Newman) extracted information independently from all RCTs using a pre-determined template. This included sample and trial characteristics, primary and secondary outcome measures and results, and details of the PROMs used [34].

Risk of bias and strength of evidence

Two authors (Z.M. Goh and R. Burney) assessed the systematic review risk of bias using the Cochrane RoB2 Toolkit and the strength of evidence according to the Grading of Recommendations, Assessment, Development, and Evaluation (GRADE) criteria. The 10-step COSMIN risk of bias checklist was completed by two authors (F. Varian and C. Pearson) [12, 32]. Any disagreements were discussed until consensus was reached. A COSMIN summary of the overall strength of each PROM recommendation was made by grading them into one of the three following categories: A) the PROM is supported by sufficient evidence of psychometric properties and can be trusted for use; B) the PROM has potential for recommendation but lacks sufficient evidence to meet criteria for categories A or C; or C) the PROM has high-quality evidence indicating that a measurement property is insufficient and therefore should not be recommended for use [12, 13, 32]. A description of psychometric terms is available in supplementary table E5.

Data analysis

It is recommended that the MCID is considered when calculating sample sizes for PROM selection [23–25]. To determine whether trials were sufficiently powered for the chosen PROM, the MCID for each instrument was obtained (supplementary table E1). If data were unavailable specifically for PH, an estimated MCID was searched for 1) respiratory conditions and 2) heart failure to maximise PROM inclusion [21, 35–40]. GPower v3.1 was used to estimate sample sizes using the MCID from a two independent means model for 80% power, 5% significance and one-tailed test. Trials insufficiently powered to detect a meaningful change were excluded [41, 42]. Meta-analysis was undertaken per therapy and by PROM in SPSS v28.1 [43].

Conceptual framework and patient and public involvement and engagement (PPIE)

A scoping literature review was conducted independently by two authors (F. Varian and R. Burney) to identify HRQoL concepts in PH [25, 26]. Published studies using primary and secondary analytical methods and grey literature, such as surveys conducted by PH associations, were included (supplementary table E6) [6–9, 45–53]. Seven major themes from an *a priori* model [28, 54] were used to inform the conceptual framework. Subthemes were then extracted and grouped [28, 54]. Final subthemes were weighted from most to least commonly reported. Key professional stakeholders from centres in the UK and Ireland then ratified the framework followed by PPIE obtained from representatives from the Pulmonary Hypertension Association UK (PHA UK) and patient volunteers registered within Sheffield's local PPIE PAH network. Participation was entirely voluntary without reimbursement. The conceptual framework was then used to evaluate the PROMs identified from the COSMIN review in a process called "mapping", allowing visualisation of which HRQoL concepts are captured by each PROM.

Results

Systematic review of PROMs selected in PAH clinical trials

The systematic search identified 896 unique records. After screening, 178 remained for full-text review. Overall, 90 RCTs were identified and 43 used PROMs as a secondary end-point (supplementary table E2). All studies showed some risk of bias (supplementary table E3). The strength of all studies using PROMs was "moderate" (supplementary table E4). There was no reported patient involvement in the selection of PROMs in any RCT (supplementary table E2) [18].

MCIDs are available for three of four PH-specific PROMs, namely EmPHasis-10, the LPHQ and CAMPHOR [21, 35, 55, 56]. The MCID for PAH-SYMPACT was excluded as this was in abstract form only with no reported standard deviation [40, 57, 58]. All available MCID values and methods of derivation are included in supplementary table E1. Figure 1 shows that 20 of the 43 RCTs selected a PROM with an MCID for PAH. Of these, only eight trials met the full inclusion criteria with adequate power to detect a meaningful change in HRQoL (table 1) [59–102]. The PROMs meeting final inclusion criteria (table 1) were SF-36, EQ-5D-5L, the LPHQ and Minnesota Living with Heart Failure-PH (MLWHF-PH). A utility index is available for the use of EQ-5D-5L as a PWM for the PAH population [3],

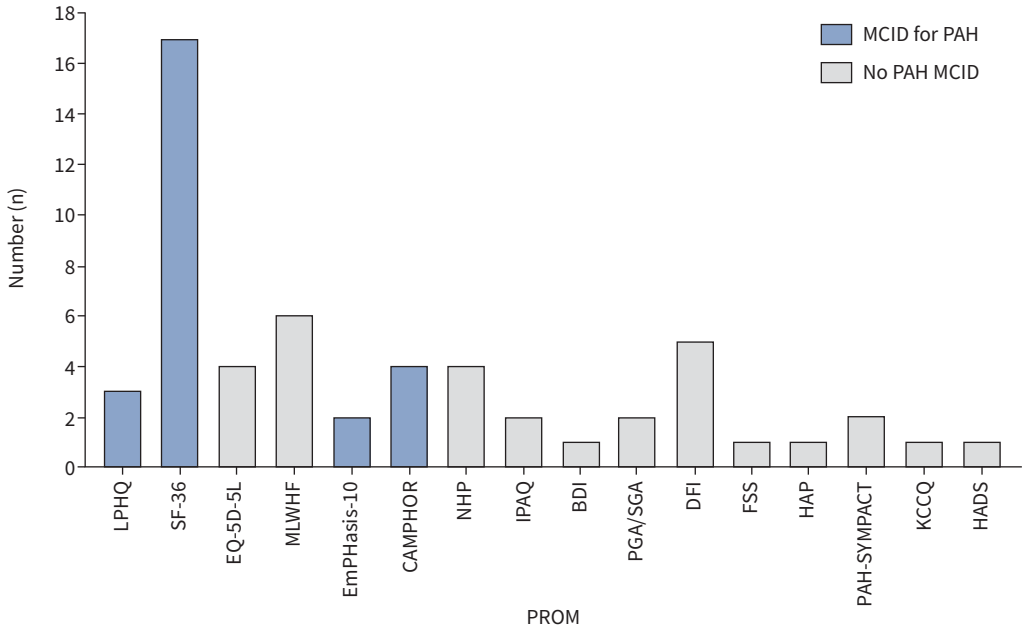


FIGURE 1 Health-related quality of life (HRQoL) instruments in pulmonary arterial hypertension (PAH) randomised controlled trials (RCTs) from the systematic review categorised by ability to distinguish meaningful change in the PAH population. 20 of 43 trials selected an instrument with a minimal clinically important difference (MCID) for HRQoL. 56 total instruments are included as 13 trials included more than one instrument, see supplementary table E2 for all RCTs with an HRQoL end-point. No trials reported results in the context of MCID. The Pulmonary Arterial Hypertension Symptoms and Impact Questionnaire (PAH-SYMPACT) MCID evaluation is underway [58]. BDI: Beck's Depression Inventory; CAMPHOR: Cambridge Pulmonary Hypertension Outcome Review; EQ-5D-5L: EuroQol-five dimensions-five levels; DFI: Dyspnoea Fatigue Index; FSS: Fatigue Severity Score; HADS: Hospital Anxiety and Depression Scale; HAP: Human Activity Profile; IPAQ: International Physical Activity Questionnaire; KCCQ: Kansas City Cardiomyopathy Questionnaire; LPHQ: Living with Pulmonary Hypertension Questionnaire; MLWHF: Minnesota Living with Heart Failure; NHP: Nottingham Health Profile; PGA: patient global assessment; PROM: patient-reported outcome measure; SF-36: 36-item Short-Form Survey; SGA: subjective global assessment.

but a specific MCID has not been established. The MCID from an interstitial lung disease population was therefore used as a surrogate to estimate the sample size, thereby maximising inclusivity [103].

Studies of Bosentan (EARLY) [89], intravenous epoprostenol (PACES) [91] and inhaled treprostinil (TRIUMPH-I) [88] did not meet their primary end-point (6MWD), which was consistent with findings of no improvement in the SF-36 physical functioning domain (table 1). Significant improvements in 6MWD for ambrisentan (ARIES2) [90] and exercise (EU-TRAIN-01) [63] were reported; however, only the MCID was met for the role-physical domain of SF-36 in EU-TRAIN-01 [63].

PATENT-1 [77] and PATENT-2 [78] (riociguat *versus* placebo) were the only RCTs available for meta-analysis (figure 2). Two PROMs (EQ-5D-5L and the LPHQ) were completed by the same patients. Overall, EQ-5D-5L appeared less responsive to changes in HRQoL (Cohen's d effect size (ES)=0.24, SE=0.08, $p<0.001$) compared to the LPHQ (ES= −0.48, SE=0.11, $p<0.001$) (figure 2). Though an exploratory end-point, it is unclear which country-specific value sets were used for EQ-5D-5L, limiting its meaningful interpretation as a PWM [16, 77, 79, 104, 105]. All trials reported statistical significance ($p<0.05$) between arms, rather than interpreting results in the context of the MCID as per regulatory recommendations. A sustained improvement in HRQoL was shown with all dose regimens of riociguat compared to placebo, as measured by the LPHQ [78]. However, EQ-5D-5L was only responsive to changes in the 1.5 mg subgroup and not the 2.5 mg (figure 2) [78]. On further examination, the 1.5 mg subgroup had a statistically higher proportion of patients in World Health Organization functional class (WHO FC) III compared to II (Fisher's exact test, $p<0.05$), supporting the need for evaluation of psychometric property performance.

TABLE 1 Characteristics of studies powered for health-related quality of life (HRQoL)

Study (year), reference	n	Demographics (mean±SD age (years), female (%))		Intervention	PROM[#]	Additional end-points	Primary outcome met?
		Control	Therapy				
EU-TRAIN-01 (2021) [63]	129	55±12.7 77.6%	52.3±12.4 69%	Exercise training	SF-36 (all)	Primary: 6MWD Secondary: WHO FC, CPET	Yes
PATENT-2 (2015) [78, 79]	396	49±16 80%	49±16 80%	Riociguat	LPHQ EQ-5D-5L	Primary: 6MWD Secondary: CWEs, NTProBNP, WHO FC, haemodynamics	Yes
PATENT-1 (2013) [77]	443	51±17 78%	51±17 79%	Riociguat	LPHQ EQ-5D-5L	Primary: 6MWD Secondary: CWEs, NTProBNP, WHO FC, haemodynamics	Yes
TRIUMPH-I (2010) [88]	235	52 (18–75)[¶] 82%	55 (20–75)[¶] 81%	Inhaled treprostinil	MLWHF-PH	Primary: 6MWD	No
EARLY (2008) [89]	185	44±17 63%	45±18 76%	Bosentan	SF-36 (all)	Primary: 6MWD, PVR Secondary: CWEs, NTProBNP, haemodynamics	No
ARIES2 (2008) [90]	394	51±14 68%	51±15 78%	Ambrisentan	SF-36 (physical)	Primary: 6MWD Secondary: CWE	Yes
PACES (2008) [91]	267	48±13 77%	48±13 82%	Sildenafil+*i.v.* epoprostenol	SF-36 (all)	Primary: 6MWD	No
AIR (2002) [98]	203	53±12 68%	51±13 68%	Inhaled iloprost	EQ-5D-5L and EQ-VAS	Primary: 6MWD Secondary: NYHA, PVR, CWE	Yes

[#]: All secondary end-points. [¶]: Reported as mean (range). 6MWD: 6-min walk distance; CPET: cardiopulmonary exercise test; CWE: clinical worsening event; EQ-5D-5L: EuroQol-five dimensions-five levels; EQ-VAS: EuroQol visual analogue scale; *i.v.*: intravenous; LPHQ: Living with Pulmonary Hypertension Questionnaire; MLWHF-PH: Minnesota Living with Heart Failure – pulmonary hypertension; NTProBNP: N-terminal pro-brain natriuretic peptide; NYHA: New York Heart Failure Association functional score; PROM: patient-reported outcome measure; PVR: pulmonary vascular resistance; SF-36: 36-Item Short-Form Survey; WHO FC: World Health Organization Functional Class.

For COSMIN evaluation, 369 eligible articles were screened with additional citation searching (n=20) from three systematic reviews. EmPHasis-10 was considered relevant for inclusion due to the availability of MCID results and its selection of two recruiting RCTs with adequate sample size [106, 107]. 21 studies demonstrated measurement properties (supplementary figure E3) [20, 35, 42, 55, 108–124]. MLWHF-PH [109] was later renamed the LPHQ and therefore these instruments were pooled for evaluation [20, 35, 108, 109]. SF-36 is available as both a PWM and PROM. However, the MCID for SF-36 is only available for four out of eight domains (physical functioning, role-physical, energy-fatigue and social functioning) [111]. EQ-5D-5L is also a PWM with a value set for PH [3, 125], although no specific MCID is available and there are few psychometric studies. All PROMs selected from the initial systematic review of clinical trials were included in step 10 of the checklist, which evaluates PROM responsiveness [12, 13, 32].

PROM suitability for the PAH population: reporting the 10-step COSMIN checklist

Steps 1 and 2 of the COSMIN risk of bias assessment involves evaluation of PROM development and consideration of how comprehensively a PROM measures HRQoL in PAH. This is described as content validity. Additional descriptions of COSMIN psychometric terms are presented in supplementary table E5. The COSMIN measurement properties of selected PROMs are outlined in table 2 with full details in supplementary tables E7 and E8.

COSMIN recommends that pre- and post-cognitive interviewing with patients and experts must be performed to adequately validate PROM content (steps 1 and 2). The LPHQ was the only instrument to satisfy this criterion. No *post hoc* cognitive interviewing has been performed in the PAH population for SF-36, EmPHasis-10 or EQ-5D-5L and content validation for these PROMs is therefore insufficient (table 2).

COSMIN steps 3, 4 and 5 outline the appropriate sample sizes and tests in which to evaluate the internal structure of the PROM (*i.e.* how well the questionnaire items perform in measuring HRQoL). The structural validity (step 3) should be appropriately analysed and reported as an overall "model fit" [126]. The type of model helps determine PROM scoring. The LPHQ is a multidimensional model with three scoring methods, namely physical, emotional and total scores [35]. SF-36 is also multidimensional with eight domains and physical and mental component scores, whereas EmPHasis-10 was derived as a

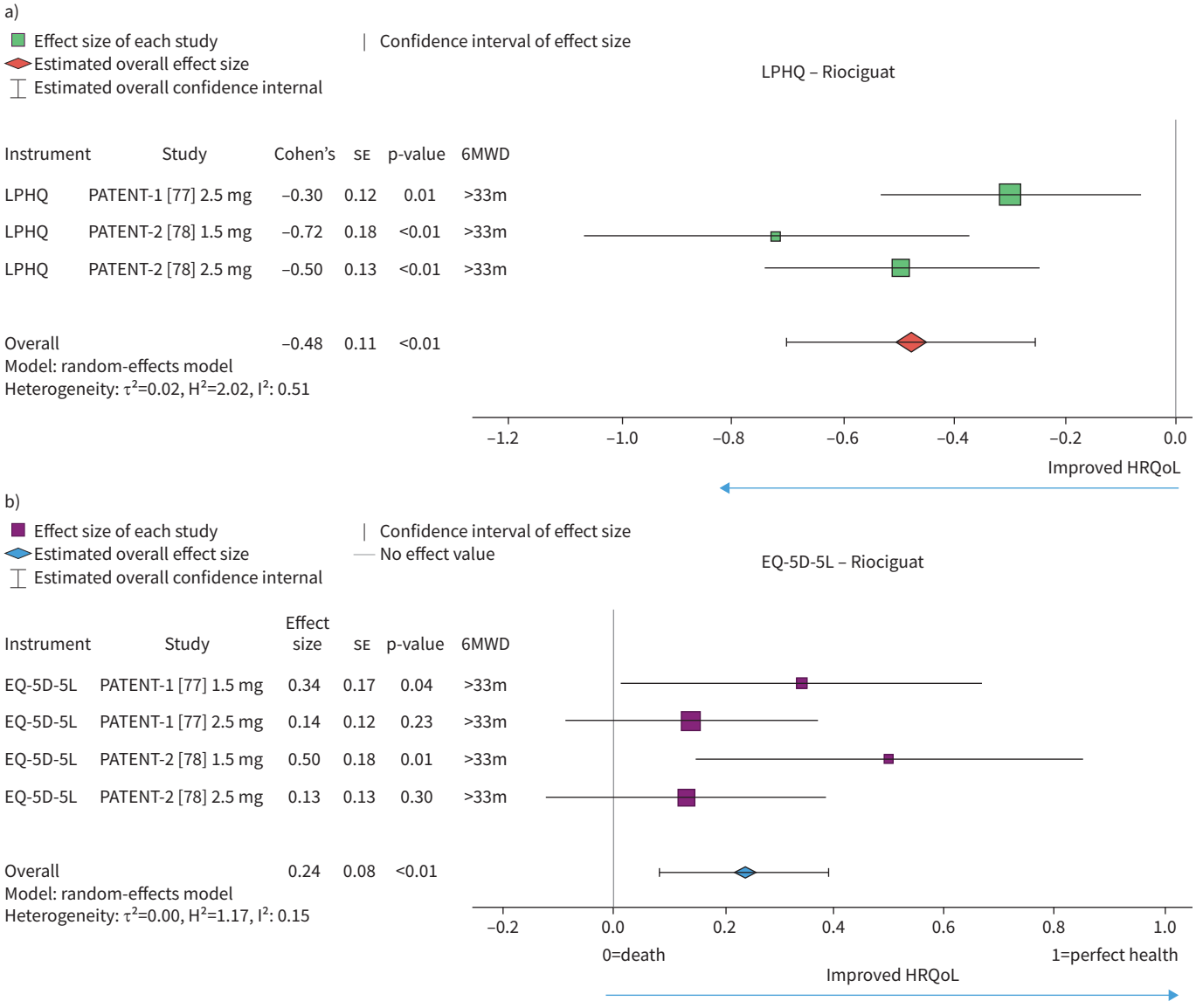


FIGURE 2 Meta-analysis of health-related quality of life (HRQoL) outcomes for riociguat. Patient-reported outcome measure (PROM) instruments: a) Living with Pulmonary Hypertension Questionnaire (LPHQ) and b) EuroQol-five dimensions-five levels (EQ-5D-5L). The 1.5 mg dose in PATENT-1 for LPHQ was excluded as the subgroup was insufficiently powered. Utility index score was not reported with EQ-5D-5L analysis. PROMs delivered at start and week 12 for PATENT-1 and every 2 weeks up to week 8 for the PATENT-2 follow-on study. 12-month follow-up data for EQ-5D-5L from PATENT-2 is not included. No imputation was reported of missing data. PATENT-1 [77] 2.5 mg riociguat, n=254 (World Health Organization Functional Class (WHO FC) III, n=140 (55%) *versus* WHO FC II, n=108 (43%), p>0.05). PATENT-1 [77] 1.5 mg riociguat, n=63 (WHO FC III, n=39 (62%) *versus* WHO FC II, n=19 (30%), p<0.0001). PATENT-2 [78] 2.5 mg riociguat, n=231 (WHO FC III, n=127 (55%) *versus* WHO FC II, n=97 (42%), p>0.05). PATENT-2 [78] 1.5 mg riociguat, n=56 (WHO FC III n= 35 (63%) *versus* WHO FC II, n=17 (30%), p<0.005). All Fisher's exact test. 6MWD: 6-min walk distance.

unidimensional structure (*i.e.* a single, total score) [114]. However, a recent model analysis restructured EmPHasis-10 into three scoring components, namely breathlessness, fatigue and independence [120]. Further clinical evaluation is required. Overall, structural validity is not reported for SF-36 or EQ-5D-5L and is inadequate for the LPHQ and EmPHasis-10 (table 2).

Additional checks for internal structure include step 4 – internal consistency: to evaluate whether similar concepts agree (typically Cronbach's alpha ⩾0.7) and step 5 – measurement invariance: stability of PROM responses across different groups (*i.e.* reducing potential confounders *e.g.* age, gender). Internal consistency is sufficient for the LPHQ, EmPHasis-10 and SF-36. This property is not relevant for EQ-5D-5L as items are not inter-related (*i.e.* they all measure different concepts), with only one item per

TABLE 2 Summary of patient-reported outcome measure (PROM) characteristics, measurement properties and evidence quality

	LPHQ/MLWHF [35, 77, 78, 108, 109, 123]	**EmPHasis-10** [55, 114–122]	**SF-36** [42, 63, 89, 90, 91, 98, 108–113]	**EQ-5D-5L** [35, 77, 78, 148]
Characteristics				
Setting	Clinical and trial	Clinical	Clinical	NA
Completion	5–10 min	3–4 min	5–10 min	3–4 min
Number of items/scales	21 Likert	10 Likert	36 Likert/discrete	5 Likert
Recall period	1 week	Recent experience	Varies with item	Today
Derivation PAH	USA, Germany, France	UK and Ireland	NA	NA
Translations	English	>20 [129]	193	208
Cost[#]	Free [149]	Free [150]	Free [151]	Free [152]
Construct(s)	Total, physical and emotional summary scores	Total score	Eight domains[¶], physical and mental scores	Score indexed for population health state
Scoring (best HRQoL)	0–105 (0)	0–50 (0)	Varies	0–1 (1)
Content validity (steps 1–2)				
1) Construct	+	+	–	NA
1) Concept	?	–	–	NA
1) Cognitive interview	?	–	–	NA
2) Patients	+	?	–	NA
2) Experts	–	?	–	NA
Internal structure (steps 3–5)				
3) Structural validity	?	?	NA	NA
4) Internal consistency	+	+	+	Not relevant
5) Measurement invariance/ cross-cultural validity	NA/none	–/USA, Japan, China, Italy, Turkey	NA/none for PAH[+]	NA/none for PAH[+]
Reliability (steps 6–7)				
6) Test–retest	?	+	?	NA
7) Measurement error	–	+	+	NA
Hypothesis testing (steps 8–10)				
8) Criterion validity	NA	NA	NA	NA
9) Convergent/construct validity	+	+	+	?
10) Response to intervention	+	NA	+	?
10) Subgroup responsivity	+	+	?[§]	NA
Quality of evidence summary	Low	Low	Very low	NA
Recommendation (A–C)	A	A	B	B

[#]: Noncommercial use. [¶]: Eight items, as follows: physical, role physical, energy fatigue, social functioning, mental health, role emotional, general health and vitality. Factor coefficients for mental and physical summary scores are held under copyright, reporting a total overall score is not recommended [110, 153]. [+]: Cross-cultural validation may be available for other disease areas/healthy populations however this is outside the scope of this search. If further evaluation is undertaken for measurement invariance in the pulmonary arterial hypertension (PAH) population, it is recommended to re-evaluate availability cross-culturally. [§]: Inconsistencies with item functioning. A full summary of findings is available in supplemental tables E7 and E8. Summary of evidence quality based on a modified Grading of Recommendations, Assessment, Development, and Evaluation (GRADE) approach. Properties with moderate to high evidence quality are shaded grey. +: Sufficient; –: insufficient; ?: indeterminate; NA: not available. Recommendations are made by three categories, as follows: A) PROM can be trusted for use with evidence for sufficient content validity and internal consistency, B) potential to be recommended for use but not categorised as A or C, or C) PROMs with high-quality evidence that a measurement property is insufficient and therefore should not be recommended for use. EQ-5D-5L: EuroQol-five dimensions-five levels; HRQoL: health-related quality of life; LPHQ: Living with Pulmonary Hypertension Questionnaire; MLWHF: Minnesota Living with Heart Failure; SF-36: 36-item Short Form Survey.

domain (table 2) [127]. No studies adequately considered measurement invariance (step 5). Step 5 additionally involves evaluation of cross-cultural validity. To use PROMs in multiple different languages and cultures, appropriate statistical testing must be performed [12, 14]. While multiple translations are available for EQ-5D-5L and SF-36, there is insufficient validation in the PH population (table 2, supplementary table E8) [128]. However, EmPHasis-10 demonstrates strong linguistic testing [117, 118, 122, 129]. Developed in the UK and Ireland, EmPHasis-10 is the only PROM validated in the USA, China, Japan, Italy and Turkey [114, 115, 117, 118, 121, 122]. The LPHQ is available in English only, though was developed in the USA, France and Germany.

Step 6 evaluates PROM performance under stable conditions, termed test–retest reliability, *e.g.* the intraclass correlation coefficient. This ensures reproducible PROM scores under similar conditions. Reliability should be interpreted within the context of step 7 – measurement error, *e.g.* standard error of measurement (SEM). Measurement error defines the natural score variation of responders during a period of

stability. If the mean variation exceeds the MCID (clinical interpretation of the score), then the PROM becomes invalid, as any detectable change is indistinguishable from normal scoring variation. According to COSMIN, this would raise a major bias concern, rendering the PROM as grade C – not recommended for use [12, 14]. Test–retest reliability is sufficient for EmPHasis-10 and indeterminate for the LPHQ due to its limited evaluation of measurement error (supplementary table E8). Only two SF-36 domains (physical functioning and general health) meet adequate test–retest reliability. Other domains have a high risk of bias due to wide confidence intervals and SEM [110, 111].

Criterion validity (step 8) can be used to assess sensitivity and specificity of the instrument; however, this requires a "gold standard" measure/PROM and is therefore excluded from this analysis. Step 9 – hypothesis testing should occur in a stable population and consider 1) how well the PROM correlates with other PROMs developed for PH (convergent validity) and 2) how well the PROM discriminates known subgroups, *e.g.* WHO FC (construct validity). Correlation with other PH-specific PROMs is satisfactory for the LPHQ, EmPHasis-10 and SF-36, and indeterminate for EQ-5D-5L (supplementary table 8). The LPHQ and EmPHasis-10 both show good correlation with WHO FC [35, 108, 123], though SF-36 is inadequate [42, 108, 111, 124]. In-hospital invasive haemodynamic assessments have yet to show strong correlation with any PROM [42, 78, 110–112, 116].

Step 10 evaluates responsiveness. Again, described as a construct approach, this considers 1) the responsiveness of the PROM with the responsiveness of other PROMs (as exemplified in figure 2), 2) the ability to detect changes in subgroups, and 3) the response to an intervention (table 2, figure 2). EmPHasis-10 and the LPHQ show satisfactory subgroup responsivity (*e.g.* improvement from WHO FC III to II) (supplementary table E8). EmPHasis-10 is historically absent from RCT interventions, with much-anticipated trials underway [106, 107].

Summary COSMIN instrument recommendations are grade A for the LPHQ and EmPHasis-10 and grade B for SF-36 and EQ-5D-5L. No PROMs received a grade C recommendation. However, the overall quality of evidence for the LPHQ and Emphasis-10 is low, and for SF-36, very low (table 2).

Developing an HRQoL conceptual framework for PROM content validation

Improving HRQoL matters to people living with PH. Surveys report this as the most important treatment focus (52–83%) over other outcomes such as life expectancy (33–75%, n=1196, UK, Canada) [10, 45]. HRQoL concepts of interest may vary between clinical and trial applications [25, 32, 44, 130]; however, recognising their relationship to PROMs is key for appropriate selection. A conceptual framework developed from the WILSON and CLEARY [28] – and subsequent – models [29, 54] was inductively modified to reflect concepts of HRQoL. These subthemes (*e.g.* "stigma") were identified from combined questionnaires and surveys of 8045 people living with PH globally [19, 115, 131]. Demographics (where available) were reflective of the disease prevalence with a female predominance (79%, n=4700). The average age of patients was 55 years (range 24–80 years) and 88% self-reported to be Caucasian (supplementary table E6).

Figure 3 summarises the conceptual framework, with six themes and 25 subthemes identified. The framework was ratified by six PH consultants, two clinical fellows, one clinical nurse specialist, one physiotherapist and one clinical psychologist. The PPIE was obtained from two PHA UK representatives and five patients with relevant demographic representation. No additional themes or subthemes were identified.

The most frequently reported concepts in figure 3 (left panel) are presented in bold, arranged from most to least common from left to right and those overlapping having similar weighting. The most reported impacts were activity, sadness/depression, self-worth, sense of loss, treatment burden and financial burden. Cultural variation was evident for this latter subtheme and more commonly discussed in surveys and interviews of those living in Canada, the USA and China compared with the UK and Europe.

PROM content was then "mapped" onto the conceptual framework (figure 3 (right panel)) to visualise themes and subthemes likely to be measured. No single PROM covers all subthemes directly. EmPHasis-10 and the LPHQ cover all main themes. Two commonly reported themes, self-identity and autonomy, are not specifically captured by EQ-5D-5L or SF-36. In addition, EQ-5D-5L does not capture the impact of PH on relationships. The LPHQ is the only PROM to directly ask about treatment burden by including items on side-effects, but may also include items that are less impactful in this patient group (*e.g.* diet).

Discussion

This is the first systematic review to evaluate meaningful changes in HRQoL in RCTs in patients with PAH. Based on rigorous methodology following COSMIN guidance, both EmPHasis-10 and the LPHQ

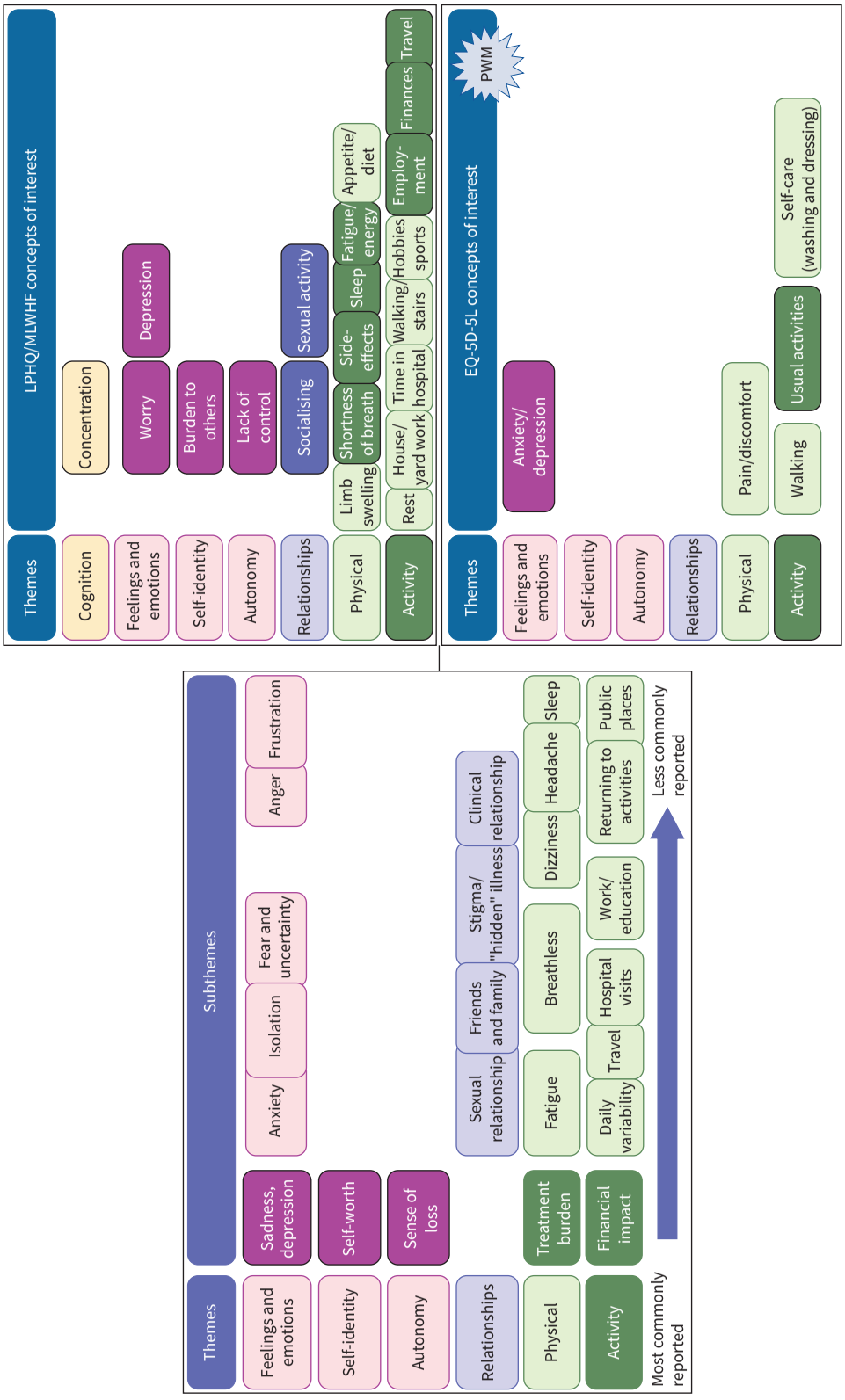


FIGURE 3 (Left) Conceptual framework for pulmonary arterial hypertension (PAH) of health-related quality of life (HRQoL) and (right) example of "mapping" of patient-reported outcome measures (PROMs) used in PAH randomised controlled trials onto the conceptual framework. (Left) Framework of patient-reported themes (n=6) and subthemes (n=25) identified by two independent reviewers on the impact of PAH (majority group 1 PAH) on HRQoL. Directly reported concepts are in bold (n=8,045 from supplementary table E6). Concepts may indirectly cross subthemes (cross-loading). For example, treatment burden may impact on the EuroQol-five dimensions-five levels (EQ-5D-5L) item "pain/discomfort" as a reflection of treatment side-effects; however, this is not included within the scope of this analysis. Further PROMs mapped are available in online supplementary figure E4. The Living with Pulmonary Hypertension Questionnaire (LPHQ) is combined with the Minnesota Living with Heart Failure (MLWHF) questionnaire as these instruments are identical. (Right) Conceptual mapping of PROMs to the HRQoL framework (undertaken by six PAH consultants, two PAH clinical fellows, one nurse specialist, one clinical psychologist and one physiotherapist). The LPHQ covers all main themes compared to EQ-5D-5L. PWM: preference-weighted measure.

receive a grade A recommendation for use, whereas SF-36 and EQ-5D-5L receive a grade B recommendation. Of these PROMs, EmPHasis-10 provides the broadest scope internationally and is validated on three continents. PROMs meeting COSMIN guidance meet historical FDA recommendations [11]; however, this regulatory guidance pre-dates international Delphi consensus [13, 14, 18]. Whilst SF-36 is the most frequently used PROM in PAH RCTs to date, six of the eight SF-36 domains are insufficiently evidenced according to the checklist. No PROMs selected in PAH RCTs to date are adequate for PAH QALY calculations. All 10 areas of the checklist require further study, as the overall quality of evidence summary is low or very low.

To aid future PROM selection and HRQoL evaluation, we developed a conceptual framework that allows visualisation of aspects of HRQoL important to people living with PH. Six themes and 25 subthemes were identified by researchers and ratified in the conceptual framework. Whereas both EmPHasis-10 and the LPHQ likely capture all major themes, two major themes (self-identity and autonomy) are unlikely to be captured by SF-36 and EQ-5D-5L. Further cognitive interviewing discussing how PROMs relate to this conceptual framework from the patients' perspective is required.

Selection of PROMs should be clinically meaningful and patient-centred

PROMs should be resilient to the day-to-day variability in HRQoL. There will be a natural change in score, without a significant change in HRQoL, and this may vary depending on disease severity. Meaningful change in HRQoL, therefore, may not equate to statistical difference [132, 133]. It is common practice in PH to calculate meaningful change thresholds on clinician-derived concepts or by asking patients to report change using the SF-36 physical functioning domain and compare this to functional changes using 6MWD [134]. Caution is advised when interpreting these MCID values, as they capture a narrow view of HRQoL, as illustrated by multiple concepts of the HRQoL framework [134, 135]. PAH-SYMPACT outlined responder thresholds in a recent abstract, calculating the MCID using clinician-derived measures (6MWD, haemodynamics and WHO FC) [58]. On review of the COSMIN guidance, this would be considered hypothesis testing/responsiveness (steps 9 and 10), rather than a patient-centred change in HRQoL. CAMPHOR remains the only PROM in PH to include patient opinion in derivation of a MCID; however, the feasibility of using this 65-item questionnaire in clinical trials has potentially limited historical selection [55, 56, 134]. There is further inaccuracy caused by over-simplifying the MCID at the group level [132, 136]. Multiple MCIDs should ideally be anchored over many individual time-points to improve sensitivity [132, 134, 135]. Other factors influencing MCID include direction of change (improvement or deterioration) and individual baseline value, none of which are available for the PH PROMs evaluated [136]. While highly valuable for clinical trial end-points, MCIDs should be interpreted with caution and within the context of measurement error to determine true change over natural responder variability [22, 132, 134, 136].

HRQoL is a complex construct to accurately measure

HRQoL is a multifactorial construct with diurnal, daily and lifelong variability. Perception varies across the patient's lifespan. Changes in values and priorities (response shift) depends upon whether patients are "pre-diagnosis", "transitioning through diagnosis" or "duration living with PH". The latter group reportedly face challenges with recognising disease progression and monitoring the condition [50, 137]. Registry data shows consistent performance of EmPHasis-10 in patients with recent diagnoses (<6 months), but other time-points are lacking, with further research required [115].

Further complexity is introduced by variations in HRQoL perceptions with age, gender and disease severity [53]. Age and gender have been shown to influence PROMs [130, 138]. These factors require further assessment in the PH population [114, 131, 138]. Perceptions and response to limitations in activity also vary with individual coping strategies and personality types [137]. Responses may therefore differ depending on the choice of PROM. No PROMs used in PAH trials have specifically addressed variations in activity perceptions in longitudinal subgroup analyses. Consideration of stability and changes across subgroups (steps 9 and 10) is also required (*e.g.* WHO FC) [132]. As shown by the meta-analysis (figure 2), EQ-5D-5L may be less responsive to changes in WHO FC II compared to WHO FC III, potentially underestimating the HRQoL treatment benefit in this subgroup. Similar responses were shown when comparing EQ-5D-5L with CAMPHOR [139]. Combining PROMs in a trial setting provides useful comparators of responsiveness and is recommended for psychometric evaluation [12, 14].

Development of the conceptual framework helps visualise important HRQoL concepts captured by PROMs. All PROMs capture limitations in activities; however, two major themes identified (self-identity and autonomy) are unlikely to be captured by SF-36 and EQ-5D-5L (figure 3). While the LPHQ has received criticism for poor symptom saturation [21, 35], concepts such as "time in hospital" and

"side-effects" are uniquely captured. Some symptoms may be less relevant. For example, "palpitations", "problems with limbs" and "diet" were not reported as commonly impactful in the HRQoL framework [6, 9, 35, 45–51]. However, the LPHQ is also the only instrument to consider financial impact, which may have cultural relevance [7, 9, 46, 48]. It is unclear whether PROMs adequately capture treatment burden (a key subtheme) in PH or whether this cross-loads (*i.e.* is captured elsewhere) with other concepts. Future cognitive interviewing is required to validate all the reviewed PROMs. This should consider utilising the conceptual framework to elicit patient interpretation of PROMs and identifying perceptions of themes and subthemes across the disease course. "Mapping" the content of PROMs to the concepts in this HRQoL framework will also help to solidify PROM content validity.

Future selection of PROMs in PAH clinical trials

PROMs offer a descriptor for the patient voice and allow for patient-centred research. This requires appropriate PROM selection with a patient-centred MCID and prioritisation of PPIE preferences that are reported in line with recommendations [18, 140–145]. Greater consistency in PROM selection will improve knowledge of therapeutic outcomes according to lived patient experience. As a minimum, PAH clinical trials should select PROMs with a grade A recommendation for use. PROMs may also offer further value in health economic evaluation, though neither generic PWM (EQ-5D-5L and SF-36) can be recommended at this time [139, 146]. CAMPHOR [147] is currently the only condition-specific PWM with a value set; however, this is underutilised in RCTs and yet to undergo COSMIN evaluation. Future development of PWMs in PH should focus on either improving PROMs with a B grade recommendation and/or developing a value set for those with a grade A recommendation. This will support robust evaluation for QALY outcomes.

Strengths and limitations

Our systematic review of recent publications was designed with rigour, using multiple reviewers, a minimum of dual coders and triangulation to enhance quality. Nevertheless, data informing the conceptual framework was not analysed at source and therefore may be subject to bias. However, following UK PPIE opinion, there were no additional concepts added to the framework and, based on the authors experience in international studies in PAH, we consider the framework to be relevant for other countries. As with adaptation of PROMs cross-culturally, future research is required to ensure individual concepts are applicable to the chosen area. This process could offer further understanding of cultural differences in people living with PAH.

Analysis of instrument power was based on MCID; although, as discussed, this may be inadequate, potentially over- or under-estimating the RCTs included [22]. However, this is currently the only available measurement criteria for estimating sufficient responsiveness and is useful for calculating study size [12, 13, 32]. CAMPHOR has an MCID but did not meet inclusion due to insufficiently powered historical or forthcoming RCTs. As this is currently the only PAH-specific PWM [139, 147], independent COSMIN analysis may be warranted. Finally, it is recognised that all PROMs considered in this analysis were developed prior to COSMIN guideline recommendations and therefore some of the methodological concerns may be overstated due to missing publication details rather than instrument flaws. Despite these challenges and low quality of evidence, two instruments still achieved a grade A recommendation and should be prioritised for selection in future PAH clinical trials.

Conclusion

Eight PAH clinical trials were adequately powered to detect a meaningful change in HRQoL. Only three of these trials selected PROMs recommended for use. Despite their low grade of evidence, both the LPHQ and EmPHasis-10 can be recommended for use in clinical trials. SF-36 and EQ-5D-5L should be used with caution without further examination. Combining these with a grade A recommended PROM may be useful, in addition to offering potential for health economic analyses. The gaps in evidence highlighted using the COSMIN 10-step checklist should be consulted for future psychometric development. These include cognitive interviewing to strengthen content validity, evaluation of natural score variability and further MCID validation, taking into consideration the patient voice, directionality and disease severity. Selection of PROMs internationally also needs to consider cultural validity, which is not necessarily concordant with language availability. PROM selection can be further supported using the ratified conceptual framework to identify the HRQoL concepts they are likely to capture.

Acknowledgements: The authors would like to thank the patients from PHA UK and PHA UK representatives who supported the ratification of the conceptual framework. Thank you to Dr Charlie Elliot (Sheffield Pulmonary Vascular Disease Unit, Royal Hallamshire Hospital, Sheffield, UK) for ratifying the conceptual framework in addition to the authors.

Provenance: Submitted article, peer reviewed.

Author contributions: All authors contributed to the development of the manuscript and approved the final version.

Conflict of interest: D.G. Kiely and I. Armstrong were involved in the derivation of EmPHasis-10 but remained independent in the risk of bias analysis for the COSMIN review. J. Carlton is a co-investigator for the UK EQ-5D-5L study team. T. Peasgood is a member of EuroQol and involved in research development of the EuroQol Health and Wellbeing instrument. The other authors are not affiliated with PROMs evaluated in this review. The authors (A.M.K. Rothman/F. Varian/R. Burney/C. Pearson/Z.M. Goh/J. Newman) performing the data collection, selection and evaluation of review articles have no conflicts of interest. A.A.R. Thompson reports research funding from Heart Research UK, Janssen-Cilang Ltd, British Heart Foundation, and honoraria from Janssen-Cilad Ltd for lectures and education. A.M.K. Rothman reports research funding from Wellcome Trust Clinical Research Career Development Fellowship (206632/Z/17/Z), Medical Research Council (UK) Experimental Medicine Award (MR/W026279/1) and NIHR Biomedical Research Centre Sheffield; contributions in kind from Medtronic Inc., Abbott Laboratories, Endotronix Inc., Novartis, Janssen and Merk; and research support and consulting from NXT Biomedical, Endotronix Inc., SoniVie, Neptune and Gradient. D.G. Kiely has received personal funding from the NIHR Biomedical Research Centre Sheffield, research funding from Ferrer, GSK and Janssen, and consulting and educational funding from Acceleron, Altivant, Ferrer, Gossamer, Janssen, MSD and United Therapeutics. F. Varian has received educational funding from Janssen and is a Medical Research Council (UK) clinical fellow. J. Newman reports research funding from British Heart Foundation, and education and travel funding from Aparito Ltd and United Therapeutics. M. Toshner reports research funding from NIHR Biomedical Research Centre Cambridge and NIHR HTA; and personal support from GSK and Jansen. R. Condliffe has received honoraria for speakers' fees and conference travel from Janssen. The remaining authors have no conflicts to declare.

Support statement: Wellcome Trust Clinical Research Career Development Fellowship (A.M.K. Rothman: 206632/Z/17/Z), BHF Intermediate Fellowship (A.A.R. Thompson: FS/18/13/33281), MRC Experimental Medicine grant (A.M.K. Rothman/M. Toshner/D.G. Kiely: MR/W026279/1), BHF Clinical Research Training Fellowship (H. Zafar/A.M.K. Rothman: FS/CRTF/23/24465, M. Toshner/J. Newman: FS/CRTF/22/24390). The research was carried out at the National Institute for Health and Care Research (NIHR) Sheffield and Cambridge Cardiorespiratory Biomedical Research Centres. A.M.K. Rothman is grateful to Richard Hughes, whose generous philanthropic support has helped to make this work possible. Funding information for this article has been deposited with the Crossref Funder Registry.

References

1 Weatherald J, Boucly A, Peters A, *et al.* The evolving landscape of pulmonary arterial hypertension clinical trials. *Lancet* 2022; 400: 1884–1898.
2 Sitbon O, Nikkho S, Benza R, *et al.* Novel composite clinical endpoints and risk scores used in clinical trials in pulmonary arterial hypertension. *Pulm Circ* 2020; 10: 2045894020962960.
3 Nafees B, de Freitas HM, Beaudet A, *et al.* A health state utility study to elicit societal values associated with pulmonary hypertension. *Patient Prefer Adherence* 2023; 17: 2119–2130.
4 Dong W, Zhang Z, Chu M, *et al.* Cost-effectiveness analysis of selexipag for the combined treatment of pulmonary arterial hypertension. *Front Pharmacol* 2023; 14: 1122866
5 Ekhlasi M, Sheikhi S, Majd ZK, *et al.* Cost-effectiveness analysis of macitentan in comparison with bosentan in the treatment of pulmonary arterial hypertension in Iran. *Value Health Reg Issues* 2023; 34: 78–85.
6 Flattery MP, Pinson JM, Savage L, *et al.* Living with pulmonary artery hypertension: patients' experiences. *Heart Lung* 2005; 34: 99–107.
7 Guillevin L, Armstrong I, Aldrighetti R, *et al.* Understanding the impact of pulmonary arterial hypertension on patients' and carers' lives. *Eur Respir Rev* 2013; 22: 535–542.
8 Pulmonary Hypertension Association UK. What it means to live with PH today. Date last accessed: 5 October 2023. Date last updated: September 2017. www.phauk.org/research-survey-work/living-ph-report/
9 Armstrong I, Billings C, Kiely DG, *et al.* The patient experience of pulmonary hypertension: a large cross-sectional study of UK patients. *BMC Pulm Med* 2019; 19: 67.
10 Tremblay É, Gosselin C, Mai V, *et al.* Assessment of clinical worsening end points as a surrogate for mortality in pulmonary arterial hypertension: a systematic review and meta-analysis of randomized controlled trials. *Circulation* 2022; 146: 597–612.
11 US Food and Drug Administration. Guidance for industry patient-reported outcome measures: use in medical product development to support labeling claims. Date last accessed: 31 March 2025. Date last updated: 17 October 2019. https://www.fda.gov/media/77832/download
12 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.

13 Terwee CB, Prinsen CAC, Chiarotto A, *et al.* COSMIN methodology for evaluating the content validity of patient-reported outcome measures: a Delphi study. *Qual Life Res* 2018; 27: 1159–1170.
14 Elsman EBM, Mokkink LB, Terwee CB, *et al.* Guideline for reporting systematic reviews of outcome measurement instruments (OMIs): PRISMA-COSMIN for OMIs 2024. *Health Qual Life Outcomes* 2024; 22: 48.
15 Devlin N, Shah K, van Hout B, *et al.* Valuing health-related quality of life: an EQ-5D-5L value set for England. Date last accessed: 18 February 2024. Date last updated: January 2016. www.ohe.org/wp-content/uploads/2016/04/25-Jan-OHE-research-paper_value-set-paper-CORRECTED.pdf
16 Gerlinger C, Bamber L, Leverkus F, *et al.* Comparing the EQ-5D-5L utility index based on value sets of different countries: impact on the interpretation of clinical study results. *BMC Res Notes* 2019; 12: 18.
17 Whitehead SJ, Ali S. Health outcomes in economic evaluation: the QALY and utilities. *Br Med Bull* 2010; 96: 5–21.
18 Prinsen CAC, Vohra S, Rose MR, *et al.* Core Outcome Measures in Effectiveness Trials (COMET) initiative: protocol for an international Delphi study to achieve consensus on how to select outcome measurement instruments for outcomes included in a 'core outcome set'. *Trials* 2014; 15: 247.
19 Humbert M, Kovacs G, Hoeper MM, *et al.* 2022 ESC/ERS guidelines for the diagnosis and treatment of pulmonary hypertension. *Eur Heart J* 2022; 43: 3618–3731.
20 Yarlas A, Mathai SC, Nathan SD, *et al.* Considerations when selecting patient-reported outcome measures for assessment of health-related quality of life in patients with pulmonary hypertension. *Chest* 2022; 162: 1163–1175.
21 Rose SW, Highland KB, Kelkar AA. Clinical utility of patient-reported outcome instruments in the management of pulmonary hypertension. *JACC Heart Fail* 2024; 12: 366–376.
22 US Food and Drug Administration. Patient-focused drug development guidance public workshop methods to identify what is important to patients & select, develop or modify fit-for-purpose clinical outcomes assessments. Date last accessed: 15 July 2024. Date last updated: 16 October 2018. www.fda.gov/media/116277/download
23 Lapin BR. Considerations for reporting and reviewing studies including health-related quality of life. *Chest* 2020; 158: S49–S56.
24 Leidy NK, Revicki DA, Genesté B. Recommendations for evaluating the validity of quality of life claims for labeling and promotion. *Value Health* 1999; 2: 113–127.
25 Calvert M, Kyte D, Mercieca-Bebber R, *et al.* Guidelines for inclusion of patient-reported outcomes in clinical trial protocols. *JAMA* 2018; 319: 483–494.
26 Reeve BB, Wyrwich KW, Wu AW, *et al.* ISOQOL recommends minimum standards for patient-reported outcome measures used in patient-centered outcomes and comparative effectiveness research. *Qual Life Res* 2013; 22: 1889–1905.
27 Finch AP, Brazier JE, Mukuria C. What is the evidence for the performance of generic preference-based measures? A systematic overview of reviews. *Eur J Health Econ* 2018; 19: 557–570.
28 Wilson IB, Cleary PD. Linking clinical variables with health-related quality of life. *JAMA* 1995; 273: 59–65.
29 McDool E, Carlton J, Powell PA, *et al.* Measuring health-related quality of life in amyotrophic lateral sclerosis. *Neurology* 2024; 103: e209549.
30 Chen H, Taichman DB, Doyle RL. Health-related quality of life and patient-reported outcomes in pulmonary arterial hypertension. *Proc Am Thorac Soc* 2008; 5: 623–630.
31 Rival G, Lacasse Y, Martin S, *et al.* Effect of pulmonary arterial hypertension-specific therapies on health-related quality of life. *Chest* 2014; 146: 686–708.
32 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.
33 Higgins JPT, Thomas J, Chandler J, *et al*. Cochrane Handbook for Systematic Reviews of Interventions version 6.5 (updated August 2024). Chichester, Cochrane, 2024.
34 US Food and Drug Administration. Surrogate endpoint resources for drug and biologic development. Date last accessed: 24 May 2024. Date last updated: 24 July 2018. www.fda.gov/drugs/development-resources/surrogate-endpoint-resources-drug-and-biologic-development
35 Bonner N, Abetz L, Meunier J, *et al.* Development and validation of the living with pulmonary hypertension questionnaire in pulmonary arterial hypertension patients. *Health Qual Life Outcomes* 2013; 11: 161.
36 Witt S, Krauss E, Barbero MAN, *et al.* Psychometric properties and minimal important differences of SF-36 in idiopathic pulmonary fibrosis. *Respir Res* 2019; 20: 47.
37 Gonzalez-Saenz de Tejada M, Bilbao A, Ansola L, *et al.* Responsiveness and minimal clinically important difference of the Minnesota living with heart failure questionnaire. *Health Qual Life Outcomes* 2019; 17: 36.
38 Butler J, Shahzeb Khan M, Lindenfeld J, *et al.* Minimally clinically important difference in health status scores in patients with HFrEF *vs* HFpEF. *JACC Heart Fail* 2022; 10: 651–661.
39 McClure NS, Al Sayah F, Xie F, *et al.* Instrument-defined estimates of the minimally important difference for EQ-5D-5L index scores. *Value Health* 2017; 20: 644–650.

40 Chin KM, Gomberg-Maitland M, Channick RN, *et al.* Psychometric validation of the pulmonary arterial hypertension-symptoms and impact (PAH-SYMPACT) questionnaire: results of the SYMPHONY trial. *Chest* 2018; 154: 848–861.

41 Ulrich S. A further step toward meaningful trial outcomes for patients with pulmonary arterial hypertension: minimal important difference in 6-minute-walk distance. *Am J Respir Crit Care Med* 2023; 207: 972–974.

42 Taichman DB, Shin J, Hud L, *et al.* Health-related quality of life in patients with pulmonary arterial hypertension. *Respir Res* 2005; 6: 92.

43 Peasgood T, Brazier J. Is meta-analysis for utility values appropriate given the potential impact different elicitation methods have on values? *Pharmacoeconomics* 2015; 33: 1101–1105.

44 Aiyegbusi OL, Nair D, Peipert JD, *et al.* A narrative review of current evidence supporting the implementation of electronic patient-reported outcome measures in the management of chronic diseases. *Ther Adv Chronic Dis* 2021; 12: 204062232110159.

45 Pulmonary Hypertension Association UK. Breathless not voiceless: what it means to live with pulmonary hypertension today. Date last accessed: 2 April 2024. Date last updated: November 2023. www.phauk.org/research/pha-uk-led-research/breathless-not-voiceless-what-it-means-to-live-with-ph-today/

46 Péloquin J, Robichaud-Ekstrand S, Pepin J. Quality of life perception by women suffering from stage III or IV primary pulmonary hypertension and receiving prostacyclin treatment. *Can J Nurs Res* 1998; 30: 113–136.

47 Yorke J, Armstrong I, Bundock S. Impact of living with pulmonary hypertension: a qualitative exploration. *Nurs Health Sci* 2014; 16: 454–460.

48 Zhai Z, Zhou X, Zhang S, *et al.* The impact and financial burden of pulmonary arterial hypertension on patients and caregivers. *Medicine (Baltimore)* 2017; 96: e6783.

49 McDonough A, Matura LA, Carroll DL. Symptom experience of pulmonary arterial hypertension patients. *Clin Nurs Res* 2011; 20: 120–134.

50 Pulmonary Hypertension Association UK. Helping people get the most from their PH treatments: Phoenix study proposal findings. Date last accessed: 18 July 2023. Date last updated: March 2021. www.phauk.org/app/uploads/2021/04/Phoenix-study-research-report.pdf

51 Pulmonary Hypertension Association UK. Clinical trials: what matters to you. Date last accessed: 2 April 2024. Date last updated: 2022. www.phauk.org/clinical-trials-what-matters-to-you/.

52 McGoon MD, Ferrari P, Armstrong I, *et al.* The importance of patient perspectives in pulmonary hypertension. *Eur Respir J* 2019; 53: 1801919.

53 Rawlings GH, Beail N, Armstrong I, *et al.* Adults' experiences of living with pulmonary hypertension: a thematic synthesis of qualitative studies. *BMJ Open* 2020; 10: e041428.

54 Brazier J, Peasgood T, Mukuria C, *et al.* The EQ-HWB: overview of the development of a measure of health and wellbeing and key results. *Value Health* 2022; 25: 482–491.

55 Hendriks PM, van Thor MCJ, Wapenaar M, *et al.* The longitudinal use of EmPHasis-10 and CAMPHOR questionnaire health-related quality of life scores in patients with pulmonary arterial hypertension and chronic thromboembolic pulmonary hypertension. *Respir Med* 2021; 186: 106525.

56 Bunclark K, Doughty N, Michael A, *et al.* A minimal clinically important difference measured by the Cambridge Pulmonary Hypertension Outcome Review for patients with idiopathic pulmonary arterial hypertension. *Pulm Circ* 2021; 11: 2045894021995055.

57 McCollister D, Shaffer S, Badesch DB, *et al.* Development of the pulmonary arterial hypertension–symptoms and impact (PAH-SYMPACT®) questionnaire: a new patient-reported outcome instrument for PAH. *Respir Res* 2016; 17: 72.

58 Benjamin K, Pena J, Wang F, *et al.* Evaluation of responder thresholds for PAH-SYMPACT domain scores in patients with sotatercept treatment in the STELLAR study. *Am J Respir Crit Care Med* 2024; 209: A7293.

59 Hoeper MM, Badesch DB, Ghofrani HA, *et al.* Phase 3 trial of sotatercept for treatment of pulmonary arterial hypertension. *N Engl J Med* 2023; 388: 1478–1490.

60 Howard LS, Rosenkranz S, Frantz RP, *et al.* Assessing daily life physical activity by actigraphy in pulmonary arterial hypertension: insights from the randomized controlled study with selexipag (TRACE). *Chest* 2023; 163: 407–418.

61 Ozcan Kahraman B, Tanriverdi A, Savci S, *et al.* Effects of inspiratory muscle training in patients with pulmonary hypertension. *Am J Cardiol* 2023; 203: 406–413.

62 Rosenkranz S, Feldman J, McLaughlin VV, *et al.* Selonsertib in adults with pulmonary arterial hypertension (ARROW): a randomised, double-blind, placebo-controlled, phase 2 trial. *Lancet Respir Med* 2022; 10: 35–46.

63 Grünig E, MacKenzie A, Peacock AJ, *et al.* Standardized exercise training is feasible, safe, and effective in pulmonary arterial and chronic thromboembolic pulmonary hypertension: results from a large European multicentre randomized controlled trial. *Eur Heart J* 2021; 42: 2284–2295.

64 Hemnes AR, Silverman-Lloyd LG, Huang S, *et al.* A mobile health intervention to increase physical activity in pulmonary arterial hypertension. *Chest* 2021; 160: 1042–1052.

65 Humbert M, McLaughlin V, Gibbs JSR, *et al.* Sotatercept for the treatment of pulmonary arterial hypertension. *N Engl J Med* 2021; 384: 1204–1215.

66 Hoeper MM, Al-Hiti H, Benza RL, *et al.* Switching to riociguat *versus* maintenance therapy with phosphodiesterase-5 inhibitors in patients with pulmonary arterial hypertension (REPLACE): a multicentre, open-label, randomised controlled trial. *Lancet Respir Med* 2021; 9: 573–584.

67 Chin KM, Sitbon O, Doelberg M, *et al.* Three- *versus* two-drug therapy for patients with newly diagnosed pulmonary arterial hypertension. *J Am Coll Cardiol* 2021; 78: 1393–1403.

68 Aslan GK, Akıncı B, Yeldan I, *et al.* A randomized controlled trial on inspiratory muscle training in pulmonary hypertension: effects on respiratory functions, functional exercise capacity, physical activity, and quality of life. *Heart Lung* 2020; 49: 381–387.

69 Kahraman BO, Savci S, Ozsoy I, *et al.* Effects of neuromuscular electrical stimulation in patients with pulmonary arterial hypertension: a randomized controlled pilot study. *J Cardiol* 2020; 75: 702–708.

70 Yılmaz BC, Güçlü MB, Keleş MN, *et al.* Effects of upper extremity aerobic exercise training on oxygen consumption, exercise capacity, dyspnea and quality of life in patients with pulmonary arterial hypertension. *Heart Lung* 2020; 49: 564–571.

71 McLaughlin V, Vachiery JL, Oudiz RJ, *et al.* Patients with pulmonary arterial hypertension with and without cardiovascular risk factors: results from the AMBITION trial. *J Heart Lung Transplant* 2019; 38: 1286–1295.

72 Babu AS, Padmakumar R, Nayak K, *et al.* Effects of home-based exercise training on functional outcomes and quality of life in patients with pulmonary hypertension: a randomized clinical trial. *Indian Heart J* 2019; 71: 161–165.

73 Gerhardt F, Dumitrescu D, Gärtner C, *et al.* Oscillatory whole-body vibration improves exercise capacity and physical performance in pulmonary arterial hypertension: a randomised clinical study. *Heart* 2017; 103: 592–598.

74 Tapson VF, Jing ZC, Xu KF, *et al.* Oral treprostinil for the treatment of pulmonary arterial hypertension in patients receiving background endothelin receptor antagonist and phosphodiesterase type 5 inhibitor therapy (the FREEDOM-C2 study). *Chest* 2013; 144: 952–958.

75 van Campen JSJA, de Boer K, van de Veerdonk MC, *et al.* Bisoprolol in idiopathic pulmonary arterial hypertension: an explorative study. *Eur Respir J* 2016; 48: 787–796.

76 Sitbon O, Channick R, Chin KM, *et al.* Selexipag for the treatment of pulmonary arterial hypertension. *N Engl J Med* 2015; 373: 2522–2533.

77 Ghofrani HA, Galiè N, Grimminger F, *et al.* Riociguat for the treatment of pulmonary arterial hypertension. *N Engl J Med* 2013; 369: 330–340.

78 Rubin LJ, Galiè N, Grimminger F, *et al.* Riociguat for the treatment of pulmonary arterial hypertension: a long-term extension study (PATENT-2). *Eur Respir J* 2015; 45: 1303–1313.

79 Galiè N, Müller K, Scalise AV, *et al.* PATENT PLUS: a blinded, randomised and extension study of riociguat plus sildenafil in pulmonary arterial hypertension. *Eur Respir J* 2015; 45: 1314–1322.

80 Saglam M, Arikan H, Vardar-Yagli N, *et al.* Inspiratory muscle training in pulmonary arterial hypertension. *J Cardiopulm Rehabil Prev* 2015; 35: 198–206.

81 Ulrich S, Keusch S, Hildenbrand FF, *et al.* Effect of nocturnal oxygen and acetazolamide on exercise performance in patients with pre-capillary pulmonary hypertension and sleep-disturbed breathing: randomized, double-blind, cross-over trial. *Eur Heart J* 2015; 36: 615–623.

82 Chan L, Chin LMK, Kennedy M, *et al.* Benefits of intensive treadmill exercise training on cardiorespiratory function and quality of life in patients with pulmonary hypertension. *Chest* 2013; 143: 333–343.

83 Weinstein AA, Chin LMK, Keyser RE, *et al.* Effect of aerobic exercise training on fatigue and physical activity in patients with pulmonary arterial hypertension. *Respir Med* 2013; 107: 778–784.

84 Hoeper MM, Barst RJ, Bourge RC, *et al.* Imatinib mesylate as add-on therapy for pulmonary arterial hypertension. *Circulation* 2013; 127: 1128–1138.

85 Tapson VF, Torres F, Kermeen F, *et al.* Oral treprostinil for the treatment of pulmonary arterial hypertension in patients on background endothelin receptor antagonist and/or phosphodiesterase type 5 inhibitor therapy (the FREEDOM-C study). *Chest* 2012; 142: 1383–1390.

86 Kawut SM, Bagiella E, Lederer DJ, *et al.* Randomized clinical trial of aspirin and simvastatin for pulmonary arterial hypertension. *Circulation* 2011; 123: 2985–2993.

87 Ghofrani HA, Morrell NW, Hoeper MM, *et al.* Imatinib in pulmonary arterial hypertension patients with inadequate response to established therapy. *Am J Respir Crit Care Med* 2010; 182: 1171–1177.

88 McLaughlin V, Benza RL, Rubin LJ, *et al.* Addition of inhaled treprostinil to oral therapy for pulmonary arterial hypertension. *J Am Coll Cardiol* 2010; 55: 1915–1922.

89 Galiè N, Rubin L, Hoeper M, *et al.* Treatment of patients with mildly symptomatic pulmonary arterial hypertension with bosentan (EARLY study): a double-blind, randomised controlled trial. *Lancet* 2008; 371: 2093–2100.

90 Galiè N, Olschewski H, Oudiz RJ, *et al.* Ambrisentan for the treatment of pulmonary arterial hypertension: results of the ambrisentan in pulmonary arterial hypertension, randomized, double-blind, placebo-controlled, multicenter, efficacy (ARIES) study 1 and 2. *Circulation* 2008; 117: 3010–3019.

91 Simonneau G, Rubin LJ, Galiè N, *et al.* Addition of sildenafil to long-term intravenous epoprostenol therapy in patients with pulmonary arterial hypertension: a randomized trial. *Ann Intern Med* 2008; 149: 521–530.
92 Hoeper MM. Combining inhaled iloprost with bosentan in patients with idiopathic pulmonary arterial hypertension. *Eur Respir J* 2006; 28: 691–694.
93 Mereles D, Ehlken N, Kreuscher S, *et al.* Exercise and respiratory training improve exercise capacity and quality of life in patients with severe chronic pulmonary hypertension. *Circulation* 2006; 114: 1482–1489.
94 Galié N, Badesch D, Oudiz R, *et al.* Ambrisentan therapy for pulmonary arterial hypertension. *J Am Coll Cardiol* 2005; 46: 529–535.
95 Wilkins MR, Paul GA, Strange JW, *et al.* Sildenafil *versus* endothelin receptor antagonist for pulmonary hypertension (SERAPH) study. *Am J Respir Crit Care Med* 2005; 171: 1292–1297.
96 Humbert M, Barst RJ, Robbins IM, *et al.* Combination of bosentan with epoprostenol in pulmonary arterial hypertension: BREATHE-2. *Eur Respir J* 2004; 24: 353–359.
97 Barst RJ, McGoon M, McLaughlin V, *et al.* Beraprost therapy for pulmonary arterial hypertension. *J Am Coll Cardiol* 2003; 41: 2119–2125.
98 Olschewski H, Simonneau G, Galiè N, *et al.* Inhaled iloprost for severe pulmonary hypertension. *N Engl J Med* 2002; 347: 322–329.
99 Barst RJ, Rubin LJ, Long WA, *et al.* A comparison of continuous intravenous epoprostenol (prostacyclin) with conventional therapy for primary pulmonary hypertension. *N Engl J Med* 1996; 334: 296–301.
100 Farha S, Saygin D, Park MM, *et al.* Pulmonary arterial hypertension treatment with carvedilol for heart failure: a randomized controlled trial. *JCI Insight* 2017; 2: e95240.
101 Han X, Zhang Y, Dong L, *et al.* Treatment of pulmonary arterial hypertension using initial combination therapy of bosentan and iloprost. *Respir Care* 2017; 62: 489–496.
102 Mehta S, Sastry BKS, Souza R, *et al.* Macitentan improves health-related quality of life for patients with pulmonary arterial hypertension. *Chest* 2017; 151: 106–118.
103 Tsai APY, Hur SA, Wong A, *et al.* Minimum important difference of the EQ-5D-5L and EQ-VAS in fibrotic interstitial lung disease. *Thorax* 2021; 76: 37–43.
104 Ghofrani HA, Grimminger F, Grünig E, *et al.* Predictors of long-term outcomes in patients treated with riociguat for pulmonary arterial hypertension: data from the PATENT-2 open-label, randomised, long-term extension trial. *Lancet Respir Med* 2016; 4: 361–371.
105 Xie F, Pickard AS, Krabbe PFM, *et al.* A checklist for reporting valuation studies of multi-attribute utility-based instruments (CREATE). *Pharmacoeconomics* 2015; 33: 867–877.
106 Varian F, Dick J, Battersby C, *et al.* Pulmonary hypertension: intensification and personalization of combination Rx (PHoenix): a phase IV randomized trial for the evaluation of dose–response and clinical efficacy of riociguat and selexipag using implanted technologies. *Pulm Circ* 2024; 14: e12337.
107 Deliu N, Das R, May A, *et al.* StratosPHere 2: a response-adaptive randomised placebo-controlled phase II trial to evaluate hydroxychloroquine and phenylbutyrate in pulmonary arterial hypertension caused by mutations in BMPR2. *Trials* 2024; 25: 680.
108 Chua R, Keogh AM, Byth K, *et al.* A comparison and validation of three measures of quality of life in patients with pulmonary hypertension. *Intern Med J* 2006; 36: 705–710.
109 Zlupko M, Harhay MO, Gallop R, *et al.* Evaluation of disease-specific health-related quality of life in patients with pulmonary arterial hypertension. *Respir Med* 2008; 102: 1431–1438.
110 Twiss J, McKenna S, Ganderton L, *et al.* Psychometric performance of the CAMPHOR and SF-36 in pulmonary hypertension. *BMC Pulm Med* 2013; 13: 45.
111 Gilbert C, Brown MCJ, Cappelleri JC, *et al.* Estimating a minimally important difference in pulmonary arterial hypertension following treatment with sildenafil. *Chest* 2009; 135: 137–142.
112 Gomberg-Maitland M, Thenappan T, Rizvi K, *et al.* United States validation of the Cambridge pulmonary hypertension outcome review (CAMPHOR). *J Heart Lung Transplant* 2008; 27: 124–130.
113 Li JL, Xiao F, Liu HT, *et al.* Long-term outcomes in health-related quality of life influence chronic disease management in patients with pulmonary hypertension. *Front Cardiovasc Med* 2022; 9: 1008253
114 Yorke J, Corris P, Gaine S, *et al.* emPHasis-10: development of a health-related quality of life measure in pulmonary hypertension. *Eur Respir J* 2014; 43: 1106–1113.
115 Borgese M, Badesch D, Bull T, *et al.* EmPHasis-10 as a measure of health-related quality of life in pulmonary arterial hypertension: data from PHAR. *Eur Respir J* 2021; 57: 2000414.
116 Favoccia C, Kempny A, Yorke J, *et al.* EmPHasis-10 score for the assessment of quality of life in various types of pulmonary hypertension and its relation to outcome. *Eur J Prev Cardiol* 2019; 26: 1338–1340.
117 Takeyasu R, Tamura Y, Abe K, *et al.* Psychometric validation of a Japanese version of the emPHasis-10 questionnaire, a patient-reported outcome measure for pulmonary hypertension – multicenter study in Japan. *Circ Rep* 2020; 2: 255–259.
118 Shi Y, Dong X, Hu X, *et al.* Cross-cultural validation of the Chinese version of the EmPHasis-10 questionnaire in connective tissue disease patients with pulmonary arterial hypertension and its relationship with risk stratification. *BMC Pulm Med* 2022; 22: 264.

119 Lewis RA, Armstrong I, Bergbaum C, *et al.* EmPHasis-10 health-related quality of life score predicts outcomes in patients with idiopathic and connective tissue disease-associated pulmonary arterial hypertension: results from a UK multicentre study. *Eur Respir J* 2021; 57: 2000124.

120 Rawlings GH, Gaskell C, Beail N, *et al.* Exploratory and confirmatory factor analysis of emPHasis-10: the health-related quality-of-life measure in pulmonary hypertension. *Pulm Circ* 2024; 14: e12378.

121 Odevoglu P, Demir R, Okumus G, *et al.* Validity and reliability of the Turkish version of the EmPHasis-10 questionnaire in patients with pulmonary hypertension. *J Eval Clin Pract* 2019; 25: 896–902.

122 Surace A, Torre R, Di Simone E, *et al.* Studio di validazione dell'emPHasis-10 per la popolazione Italiana [Validation study of emPHasis-10 for the Italian population]. *NSC Nursing* 2019; 1: 1–18.

123 Cenedese E. Measurement of quality of life in pulmonary hypertension and its significance. *Eur Respir J* 2006; 28: 808–815.

124 Fernandes CJ, Martins BC, Jardim CV, *et al.* Quality of life as a prognostic marker in pulmonary arterial hypertension. *Health Qual Life Outcomes* 2014; 12: 130.

125 Davies EW, Llewellyn S, Beaudet A, *et al.* Elicitation of health state utilities associated with the mode of administration of drugs acting on the prostacyclin pathway in pulmonary arterial hypertension. *Patient Prefer Adherence* 2018; 12: 1079–1088.

126 Mokkink LB, Prinsen CAC, Patrick DI, *et al.*. COSMIN methodology for systematic reviews of patient-reported outcome measures (PROMs). Date last accessed: 30 May 2024. Date last updated: February 2018. https://cosmin.nl/wp-content/uploads/COSMIN-syst-review-for-PROMs-manual_version-1_feb-2018.pdf

127 Feng YS, Kohlmann T, Janssen MF, *et al.* Psychometric properties of the EQ-5D-5L: a systematic review of the literature. *Qual Life Res* 2021; 30: 647–673.

128 Alagappan T. The cross-cultural adaptation process of a patient-reported outcome measure. *J Sci Soc* 2023; 50: 13.

129 Foster E, Guillen A, Lara K, *et al.* Linguistic validation of the emPHasis-10 questionnaire: a patient-reported outcome instrument for assessing QoL in pulmonary hypertension (PH). *Value Health* 2015; 18: A744.

130 Calvert MJ, Cruz Rivera S, Retzer A, *et al.* Patient reported outcome assessment must be inclusive and equitable. *Nat Med* 2022; 28: 1120–1124.

131 Mair KM, Johansen AKZ, Wright AF, *et al.* Pulmonary arterial hypertension: basis of sex differences in incidence and treatment response. *Br J Pharmacol* 2014; 171: 567–579.

132 Hays RD, Peipert JD. Between-group minimally important change *versus* individual treatment responders. *Qual Life Res* 2021; 30: 2765–2772.

133 King MT. A point of minimal important difference (MID): a critique of terminology and methods. *Expert Rev Pharmacoecon Outcomes Res* 2011; 11: 171–184.

134 Swigris J, Foster B, Johnson N. Determining and reporting minimal important change for patient-reported outcome instruments in pulmonary medicine. *Eur Respir J* 2022; 60: 2200717.

135 Devji T, Carrasco-Labra A, Qasim A, *et al.* Evaluating the credibility of anchor based estimates of minimal important differences for patient reported outcomes: instrument development and reliability study. *BMJ* 2020; 369: m1714.

136 Hays RD, Woolley JM. The concept of clinically meaningful difference in health-related quality of life research. *Pharmacoeconomics* 2000; 18: 419–423.

137 Kingman M, Hinzmann B, Sweet O, *et al.* Living with pulmonary hypertension: unique insights from an international ethnographic study. *BMJ Open* 2014; 4: e004735.

138 Hertler C, Seiler A, Gramatzki D, *et al.* Sex-specific and gender-specific aspects in patient-reported outcomes. *ESMO Open* 2020; 5: e000837.

139 McKenna SP, Ratcliffe J, Meads DM, *et al.* Development and validation of a preference based measure derived from the Cambridge pulmonary hypertension outcome review (CAMPHOR) for use in cost utility analyses. *Health Qual Life Outcomes* 2008; 6: 65.

140 Bours MJL. Moving towards increased implementation of evidence-based measurements through comprehensive and transparent reporting: the PRISMA–COSMIN 2024 reporting guideline. *Qual Life Res* 2024; 33: 2047–2048.

141 Terwee CB, Ahmed S, Alhasani R, *et al.* Comparable real-world patient-reported outcomes data across health conditions, settings, and countries: the PROMIS international collaboration. *NEJM Catal Innov Care Deliv* 2024; 5: 0045.

142 US Food and Drug Administration. Patient-focused drug development: incorporating clinical outcome assessments into endpoints for regulatory decision-making. Date last accessed: 14 October 2024. Date last updated: April 2023. www.fda.gov/media/166830/download

143 Mercieca-Bebber R, King MT, Calvert MJ, *et al.* The importance of patient-reported outcomes in clinical trials and strategies for future optimization. *Patient Relat Outcome Meas* 2018; 9: 353–367.

144 Snyder C, Crossnohere N, King M, *et al.* The PROTEUS-Trials Consortium: optimizing the use of patient-reported outcomes in clinical trials. *Clin Trials* 2022; 19: 277–284.

145 Crossnohere NL, Brundage M, Calvert MJ, *et al.* International guidance on the selection of patient-reported outcome measures in clinical trials: a review. *Qual Life Res* 2021; 30: 21–40.

146 National Institute for Health and Care Excellence. NICE health technology evaluations: the manual. Date last accessed: 15 July 2024. Date last updated: 31 October 2023. www.nice.org.uk/process/pmg36/resources/nice-health-technology-evaluations-the-manual-pdf-72286779244741

147 Meads DM, McKenna SP, Doughty N, *et al.* The responsiveness and validity of the CAMPHOR utility index. *Eur Respir J* 2008; 32: 1513–1519.

148 EuroQol Research Foundation. EQ-5D-5L user guide. Date last accessed: 15 April 2024. Date last updated: September 2019. https://euroqol.org/wp-content/uploads/2023/11/EQ-5D-5LUserguide-23-07.pdf

149 University of Minnesota. Living with pulmonary hypertension questionnaire (LPHQ). Date last accessed: 1 August 2024. Date last updated: 2020. https://license.umn.edu/product/living-with-pulmonary-hypertension-questionnaire-lphq

150 Pulmonary Hypertension Association UK. EmPHasis-10 questionnaire. Date last accessed: 1 August 2024. Date last updated: 2024. www.phauk.org/pha-uk-resources/emphasis-10-questionnaire

151 QualityMetric. The SF-36v2® health survey. Date last accessed: 1 August 2024. Date last updated: 2025. www.qualitymetric.com/health-surveys/the-sf-36v2-health-survey/

152 EuroQoL. How to obtain EQ-5D. Date last accessed: 1 August 2024. Date last updated: 2 December 2024. https://euroqol.org/register/obtain-eq-5d/how-to-obtain-eq-5d

153 Lins L, Carvalho FM. SF-36 total score as a single measure of health-related quality of life: scoping review. *SAGE Open Med* 2016; 4: 205031211667172.