ESC European Society of Cardiology

European Journal of Cardiovascular Nursing (2025) **00**, 1–15
https://doi.org/10.1093/eurjcn/zvaf217

REVIEW ARTICLE

Psychometric properties of health-related quality of life patient reported outcome measures for common cardiovascular conditions: a scoping review and COSMIN analysis

Asad Bhatty [1,2,3], Adam B. Smith[2,3,4]*, Martijn Scherrenberg [5,6], Mohammad Haris [1,2,3], Chris Wilkinson[7,8], Chris P. Gale[1,2,3], and Theresa Munyombwe [1]

[1]Leeds Institute of Cardiovascular and Metabolic Medicine, University of Leeds, Leeds, UK; [2]Leeds Institute for Data Analytics, University of Leeds, Leeds, UK; [3]Department of Cardiology, Leeds Teaching Hospitals NHS Trust, Leeds, UK; [4]Leeds Institute of Rheumatic and Musculoskeletal Medicine, Leeds, UK; [5]Faculty of Medicine and Life Sciences, Diepenbeek, Hasselt, Belgium; [6]Faculty of Medicine, Antwerp University, Wilrijk 2610, Belgium; [7]Academic Cardiovascular Unit, South Tees NHS Foundation Trust, James Cook University Hospital, Middlesbrough, UK; and [8]Hull York Medical School, University of York, York, UK

Received 23 June 2025; revised 9 September 2025; accepted 3 November 2025; published 18 January 2026

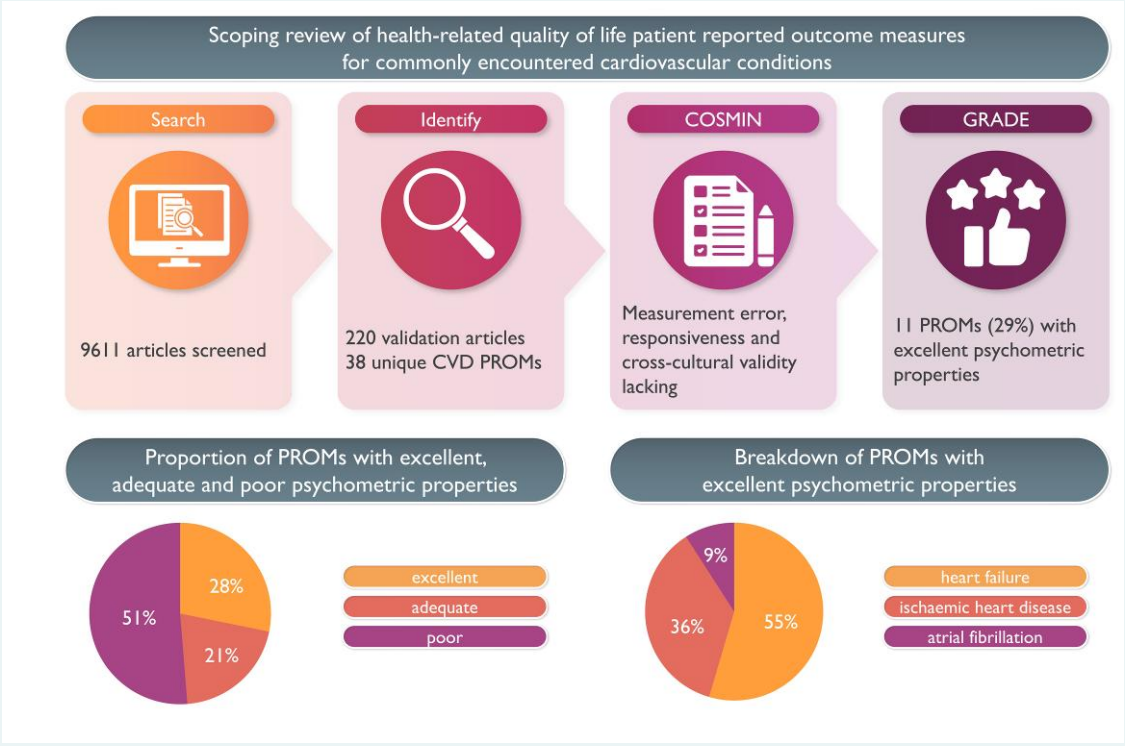
Aims Health-related quality of life (HRQoL) is an important measure of disease status and represents a holistic approach to delivering patient-centered care. We conducted a scoping review of HRQoL patient reported outcome measures (PROMs) for cardiovascular diseases (CVDs) and evaluated their psychometric properties.

Methods and results Randomized trials and observational studies that developed and validated HRQoL PROMs for adults with ischaemic heart disease (IHD), aortic stenosis (AS), atrial fibrillation (AF), heart failure (HF), or generic CVD were included, published from inception of databases to 8 February 2025 using PubMed, Web of Science, and Embase. Independent reviewers selected and extracted the psychometric properties of each PROM in accordance with the Consensus-based Standards for the selection of health Measurement Instruments (COSMIN) checklist: content validity, reliability, internal consistency, structural validity, criterion/convergent, cross-cultural validity, measurement error, hypothesis testing, and responsiveness. Each PROM was graded using the Grading of Recommendations Assessment, Development, and Evaluation approach. Of 9430 articles, 220 studies for 38 different PROMs were included (HF $n = 17$, 45%; AF $n = 11$, 29%; IHD $n = 7$, 18%; generic $n = 2$, 5%; AS $n = 1$, 3%). Eleven PROMs (29%) satisfied all nine COSMIN criteria; the majority ($n = 19$, 50%) required further validation and eight were deemed inadequate for clinical use (21%).

Conclusion This scoping review of HRQoL PROMs in individuals with common CVDs found evidence that many PROMs do not fulfill all nine COSMIN criteria for methodological quality, and for some CVDs there is a limited choice of suitable PROMs for HRQoL measurement. There is an opportunity to improve HRQoL evaluation for use within routine care and research.

* Corresponding author. Email: a.b.smith@leeds.ac.uk

Graphical Abstract



Novelty

- PROMs that evaluate HRQoL and are validated for common CVDs vary in their psychometric properties; most require further validation studies prior to use, particularly for cross-cultural validity, measurement error, and responsiveness.
- The quality of patient reported outcome data generated from such instruments may have limitations in informing clinical care and the generalisability of trial results.
- Most of the HRQoL CVD PROMs that met all nine COSMIN criteria were validated for HF (six instruments, 55%), a minority for IHD (four instruments, 36%), and one for AF (one instrument, 9%).

Introduction

Health-related quality of life (HRQoL) is an important measure of disease status and represents a holistic approach to delivering patient-centered care.[1–3] HRQoL assesses an individual's perception of the impact a condition has on their physical health, psychological and social functioning, and emotional well-being,[4,5] and is commonly quantified using validated questionnaires or patient reported outcome measures (PROMs).[6]

The adoption of HRQoL within routine care and research has broad appeal to healthcare providers and patients alike, because patients value improvements in their HRQoL similarly to additional life years gained, and the use of PROMs is associated with increased patient satisfaction.[7–9] For clinicians, studies demonstrate that patients directly reporting their symptoms and HRQoL is more accurate than a clinician's interpretation, and that a low baseline HRQoL is associated with poor long-term outcomes.[10–12]

Evaluation of HRQoL features in the European Society of Cardiology (ESC) quality indicators in some commonly treated cardiovascular disease (CVD) conditions including transcatheter aortic valve intervention (TAVI) for aortic stenosis (AS), heart failure (HF), and atrial fibrillation (AF).[13–16] Regulators recommend their use in evaluating pharmaceutical and device labeling claims in trials.[17]

To benefit from measuring HRQoL, a comprehensive evaluation of existing questionnaires is required to ensure generated patient data are accurate, valid in the specific disease and that longitudinal changes in PROM scores reliably reflect a change in disease state.[18,19] The Consensus-based Standards for the selection of health Measurement Instruments (COSMIN) initiative is a widely employed methodological questionnaire assessment, evaluating over 116 items in multiple domains.[20] Previous work has investigated the qualitative assessment of CVD PROMs. However, these studies either did not include contemporary validation studies for some CVD conditions,[21–23] restricted to investigating condition specific PROMs that included HRQoL or other PROMs,[21–24] or assessed each PROM's adherence to regulatory requirements therefore limiting applicability to trials and clinical practice.[25]

To date, there has been no comprehensive review of the psychometric properties across the common cardiovascular conditions of ischaemic heart disease (IHD), AF, HF, AS, and generic CVD HRQoL PROMs for use in routine clinical practice. We aimed to conduct a scoping review of HRQoL PROMs for these conditions and evaluate their psychometric properties, using the COSMIN framework and determine their applicability in routine clinical care and trials.

Methods

The review was reported in accordance to the COSMIN[20] and the Preferred Reporting Items for Systematic and Meta-Analysis reporting guidelines.[26]

Eligibility criteria

Peer-reviewed randomized controlled trials, and prospective and retrospective studies that developed and validated HRQoL PROMs for adults (aged 18 years and older) with IHD (myocardial infarction (MI) and angina included) HF, AF, AS, percutaneous coronary intervention (PCI), or generic CVD were included, from inception to 8 February 2025. Only studies published in English were included. CVD PROMs that were validated in other CVDs were excluded as were review articles, meeting and conference abstracts, secondary analyses, and editorials.

Search strategy

PubMed, Web of Science, and Embase were searched using a structured search strategy that followed the population, phenomenon of interest and outcome framework. Pragmatic keywords such as 'cardiovascular disease' and 'patient reported outcome measure' were included but MeSH words were not included (*Appendix 1*, Supplementary section). To minimize publication bias, targeted keyword searches of gray online literature sources were conducted, in conjunction with hand searching of the reference lists of included studies (pearling).

Study selection

Screening

Two reviewers (MS, MH) independently screened titles and abstracts and selected eligible studies after full-text assessment using the pre-determined eligibility criteria. At the full-text review, reasons for excluding studies were recorded using the Rayyan software.[27] Disagreements between reviewers were resolved through discussion, and a third reviewer (AB) was invited if the disagreement persisted.

Data extraction

Key study and PROM characteristics, and their psychometric measurement properties were extracted by two independent reviewers (TM, ABS). Study characteristics included title, author, year of publication, and design. PROMs characteristics included: number of items, domains, response format, administration methods, and each PROM was categorized according to generic or disease-specific CVD type; IHD (encompassing ACS, PCI, and angina), AF, HF, AS (encompassing TAVI), or generic CVD.

Evaluation of methodological quality

The methodological quality of each study was assessed using the COSMIN risk of bias checklist[20,28] across nine domains: content validity, internal structure [structural validity, internal consistency (IC), cross-cultural, and measurement invariance] reliability, measurement error, construct validity (criterion/convergent and hypothesis testing) and responsiveness.[26] The definition and criteria for good measurement properties in each domain are provided in Supplementary material online, *Table S1*.

The results were categorized accordingly; green as strong, yellow as adequate, and red as inadequate, (*Table 1*). For example, reliability was rated as strong if a study provided evidence that patients were stable, time interval was appropriate, test conditions were similar, inter-class correlation calculated for continuous scores and kappa for dichotomous/nominal/ordinal data (see Supplementary material online, *Table S1*).[20]

Table 1 Criteria for good measurement properties used in this study

Level	Rating	Criteria
Strong evidence in favor or against	Green	Consistent findings in multiple studies of good methodological quality or in one study of excellent methodological quality.
Moderate	Yellow	Consistent findings in multiple studies of fair methodological quality or in one study of good methodological quality.
Limited	Red	One study of fair methodological quality.
Conflicting	Red	Conflicting findings.
Unknown	Red	Only studies of poor methodological quality.

PROM quality assessment

Thereafter, the psychometric properties of each PROM was assessed and followed the recommended order: content validity, structural validity, and IC then cross-cultural validity/measurement invariance, reliability and measurement error, criterion validity (if applicable), hypotheses testing for construct validity, and responsiveness.[20] One reviewer (TM) performed the qualitative assessment of the PROM developmental articles, and a subsequent independent reviewer completed the qualitative assessment of further validation articles (ABS) following COSMIN guidelines.[20] HRQoL was assessed at domain level where information was available. The psychometric evidence of each PROM measurement property was rated using the updated quality criteria[26,29] (see Supplementary material online, *Table S1*).

Quality of evidence

The overall quality of each measurement property and the PROM quality assessment were combined to give an appraisal of the evidence provided by each validation study. Where multiple studies evaluated a measurement property of a PROM, the results of the studies were summarized to produce an overall rating, in accordance with the 'Grading of Recommendations Assessment, Development and Evaluation' (GRADE) approach.[30] GRADE is considered a transparent and systematic approach for appraising the quality of evidence in literature reviews and clinical practice guidelines.[31] The evidence from the quality appraisal was synthesized to determine which PROM would be best for use in clinical practice for each CVD condition. The full definitions and criteria of good measurement properties were aligned to contemporary COSMIN guidelines.[20] Three categories of recommendations were used for this review:

(1) High-quality evidence, most suitable to be recommended for use within clinical care.
(2) High-quality evidence for some properties, PROM may be recommended for use within clinical care but more validation is required.
(3) Insufficient evidence provided, no recommendations can be made for routine clinical care use.

Data synthesis

The summarized data was described narratively to present the results. Measurement properties of the different PROMs were reported in tables and graphs as appropriate. Additional content validity and comparison of disease-specific PROMs were conducted by examining the items of disease-specific PROMs, domains and comparing them by mapping onto the Cleary and Wilson conceptual model of HRQoL (*Table 2*).[30] The model integrates clinical and psychosocial approaches to health care and links the biological and physiological (objective health) variables to the measure of HRQoL or subjective health construct.[32] The five health concepts described in the model are biological and physiological factors, symptoms status,

functioning, general health perceptions, and overall quality of life (see Supplementary material online, *Figure S1*).

Results

Study selection

In total, 9430 articles were identified; 220 studies were included after full-text assessment (*Figure 1*). Thirty-eight unique PROMs were identified after full-text assessment; the remaining articles were validation studies of PROMs (see Supplementary material online, *Table S2*).

PROMs characteristics and coverage

Of the included studies, most evaluated HF (116 studies; 52%), then IHD (61 studies, 28%), AF (38 studies; 17%), AS (4 studies; 2%), and generic (2 studies; 1%). Of the included CVD PROMs, most evaluated HF ($n = 17$; 45%), then AF ($n = 11$; 29%), IHD ($n = 7$, 18%), generic ($n = 2$; 5%), and AS ($n = 1$; 3%).

The content and domains covered by the different PROMs are presented in *Table 2*.

Heart failure

The Kansas City Cardiomyopathy Questionnaire (KCCQ) was the most frequently evaluated among HF PROMs (33 studies; 28%).[33–35] The HF PROMs included between 7 and 86 items, 2 to 18 domains and all used a Likert scale as a response format for the items. The domains covered by the HF PROMs domains include the domains in the HRQoL Wilson and Cleary model (physical health, mental health, social health, emotional, symptom burden, and overall global health) and life satisfaction (*Tables 2* and *3*).

Ischaemic heart disease

The MacNew Heart disease HRQoL questionnaire was the most frequently evaluated among the IHD related PROMs (25 studies; 41%).[36] Of the included PROMs, four were MI related (MIDAS,[37] MacNew,[36] QLMI_1,[38] QLMI_2[39]) which included between 26 and 35 items and used a Likert scale as a response format. The MacNew[36] and QLMI[38,39] cover three domains (physical, emotional, and social function), whereas the MIDAS[37] is more comprehensive by including: physical activity, insecurity, emotional reaction, dependency, diet, concerns over medication and side effects. The two Angina-PROMs were both versions of the Seattle Angina Questionnaire: SAQ-7[40] and SAQ-19.[41] SAQ-7 covers seven items, five domains and recall time is up to four weeks. The SAQ-19 has 19 items, 5 domains, and completion time is less than 5 min.

PCI and TAVI

The CROQ questionnaire is the only PROM that evaluates patient outcomes after PCI and/or coronary artery bypass grafting (CABG) and has 32 items using a 3–6-point Likert scale responses and has been evaluated in nine studies.[42] The only PROM originally validated for AS and TAVI was the Toronto aortic stenosis quality of life questionnaire (TASQ) which was validated in four studies.[43]

Atrial fibrillation

The Arrhythmia-Specific questionnaire in Tachycardia and Arrhythmia (ASTA) questionnaire is the most frequently evaluated AF PROM (nine studies; 24%).[44] The 11 AF-PROMs (AF impact,[45] AFSS, CCS-SAF,[46] AFQoL,[47] AFEQT,[48] Toronto AF symptoms,[49] AF-6,[50] QLAF,[51] PPAQ,[52] ASTA,[44] AFHLQ[53]) included between 6 and 22 items, used a Likert scale as a response format. The range of domains covered by AF-PROMs include physical health, emotional, sleep, vitality, symptoms, treatment concerns, treatment satisfaction, quality of life, and frequency and severity of symptoms (*Table 2*).

Generic cardiovascular disease

There were two HRQoL PROMs that were originally validated for any CVD: the Multidimensional Index of Life Quality (MILQ)[54] and the Canadian Cardiovascular Society (CCS) questionnaire,[55] and included between 5 and 35 items, used a Likert scale as a response format (*Table 2*).

Methodological quality of the studies

Content validity

Most PROMs demonstrated strong evidence for content validity (28 PROMs; 74%). The studies that were rated as very good provided a clear description of the methodology that was used to assess relevance, comprehensibility (i.e. use of skilled trainers, appropriate methods to analyse data, rewording of interviews, and verbatim transcription) following COSMIN guidelines.

Structural/construct validity

Most CVD PROMs demonstrated strong evidence for structural validity (27 PROMs; 71%). The methodological quality for structural validity of the included studies ranged from very good to inadequate (*Table 3*) and was evaluated using exploratory factor analysis or confirmatory factor analysis, item response theory (IRT) models. Item response theory models were less common with only two studies (4%) using these models. Overall, the quality for structural validity of 19 studies (37%) were rated as very good because they used classical test theory or IRT models, 10 were rated as adequate (19%), and 7 as inadequate (13%) due to small sample sizes or use of inappropriate methods to evaluate structural validity (*Table 3*).

Internal consistency

Most PROMs provided strong evidence for IC (30 PROM, 79%). Methods to assess this included the Cronbach alpha. Among the PROMs that fulfilled the prerequisite of one-dimensionality, with limited evidence provided for IC was demonstrated for eight PROMs (21%).

Cross-cultural validity and measurement invariance

This was assessed using multiple group factor analysis and differential item function across groups according to the COSMIN criteria (see Supplementary material online, *Table S1* for further details). Half of included CVD PROMs provided evidence for cross-cultural validity of very good or adequate methodological quality (19 PROMs; 50%, *Table 3*).

Reliability and measurement error

Most PROMs evaluated provided adequate evidence for reliability (27 PROMs; 71%). The main method used to assess test-retest reliability was intra-cluster correlation coefficient a few studies used Pearson or Spearman correlation coefficients. Measurement error was one of the COSMIN domains least evaluated by CVD PROMs with only 10 PROMs demonstrating adequate evidence on evaluation (26%, *Table 3*).

Hypothesis testing for construct validity (convergent and divergent validity)

Most PROMs provided evidence for convergent/criterion validity (32 PROMs; 84%) by comparing the correlations of the PROM with a

Table 2 Overview of the cardiovascular disease specific PROMs identified by the scoping review literature and their characteristics search

PROM	Measure, items, and domains	Response format	Recall period	Administration method
MINNESOTA LIVING WITH HEART FAILURE (MLHF)[1]	21 items measuring 2 domains: Physical and emotional. Swelling, shortness of breath, fatigue, poor memory/ concentration, and depression. Some items are not included in sub-scores	Likert scale	4 weeks	Paper, interviewer, self-administered
CHRONIC HEART FAILURE ASSESSMENT TOOL (CHAT)[2]	46 items measuring 4 domains: symptoms, activity levels, psychosocial aspects, and emotions	Likert scale	2 weeks	Self-administered
KANSAS CITY CARDIOMYOPATHY QUESTIONNAIRE (KCCQ) SHORT VERSION[3]	20 items measuring 3 domains: Dyspnoea, fatigue, emotional function	Likert scale	2 weeks	Paper, interviewer, self-administered
KCCQ[4]	23 items measuring 7 domains of patients' HF-related health status: Physical Limitation, Symptom Stability, symptom frequency, symptom burden, self-efficacy, quality of life, and social Limitations. Item responses are coded sequentially (1, 2, 3, etc.) from worst to best status. Scores are generated for each domain and scaled from 0 to 100, with 0 denoting the worst and 100 the best possible status.	Likert scale	2 weeks	Paper, self-administered
QUALITY OF LIFE IN SEVERE HEART FAILURE QUESTIONNAIRE (QLQ-SHF)[5]	26 items, Physical, psychological, symptoms, life satisfaction patients with severe heart failure (EF < 35%).	Likert scale	1 week	Interviewer, self-administered
LEFT VENTRICULAR DYSFUNCTION QUESTIONNAIRE (LVD-36)[6]	36 item questionnaires for patients with left ventricular dysfunction. Responses are dichotomous (true or false). True responses are summed, and the sum is expressed as a percentage, so that 100 is the worst possible score and 0 the best possible score.	True/false	1 week	Self-administered, telephone
PATIENT REPORTED OUTCOMES MEASUREMENT INFORMATION SYSTEMS PLUS HF PROFILE (PROMIS PLUS HF PROFILE)[7]	Is a library of measures and items; users can select subsets of domains and items and create customized short-form versions based on the clinical need or research question. Contains 86 items across 18 domains consisting of physical, emotional, social, psychological symptoms.	Likert scale	1 week	Paper, self-administered
HF SOMATIC PERCEPTION SCALE -18[8]	Assesses a patient's perception of deteriorating HF symptoms and its effect on quality of life.	5-point Likert scale	1 week	Paper, self-administered
SELF-CARE OF HF INDEX	Covers 2 main domains: a patient's self-care maintenance	Likert scale	12 weeks	Paper, REDCAP

Continued

Table 2 *Continued*

PROM	Measure, items, and domains	Response format	Recall period	Administration method
SELF-CARE MANAGEMENT SCALE-22[9]	(daily weights, daily activity, keeping to appointments) and a patient's self-care treatment (symptom recognition, symptom evaluation, quality of life etc.)			
SYMPTOM STATUS QUESTIONNAIRE-HEART FAILURE[10]	Covers 7 main items related symptoms: physical and quality of life	4-point Likert scale	4 weeks	Interviewer, self-administered
CARE-RELATED QUALITY OF LIFE SURVEY FOR CHRONIC HEART FAILURE (CAREQOL CHF)[11]	20 items and three scales: social and emotional problems, physical limitations, and being in safe hands.	5-point scale	4 weeks	Interviewer, self-administered
CHRONIC HEART FAILURE HEALTH-RELATED QUALITY OF LIFE QUESTIONNAIRE (CHFQOLQ-20)[12]	20 items, measuring 4 domains Physical, emotional, psychological, life satisfaction.	5-point Likert scale	4 weeks	Interviewer, self-administered
HEART FAILURE-DAILY SYMPTOM DIARY[13]	HF symptoms cover 3 main domains; physiological symptoms (ankle oedema, shortness of breath etc.), physical functioning (walking, climbing stairs, sleeping etc.) and psychological symptoms (irritability, fear, and anxiety etc.)	Free text	Daily	Interviewer, self-administered
HEARTQOL[14]	Physical, emotional, global QoL. The HeartQoL questionnaire comprises 14-items with 10-item physical and 4-item emotional subscales which are scored from 0 (poor HRQL) to 3 (better HRQL) with a global score if needed.	4 point, 0 to 3 Likert scale	4 weeks	Paper, self-administered
MD ANDERSON SYMPTOM INVENTORY-HEART FAILURE (MDASI-HF)[15]	Covers two main domains: physical and psychological. The physiologic symptoms include abdominal bloating, ankle swelling, chest pain, difficulty sleeping with head of bed flat (orthopnoea), dizziness, fatigue, loss of appetite, lack of energy, limitation in physical activity, lower extremity swelling, nausea, rapid heartbeat (palpitations), nighttime cough, shortness of breath, sleep problems, thirst, urinary incontinence, headache, waking up at night because of shortness of breath (paroxysmal nocturnal dyspnoea), waking up at night to urinate, and sudden weight gain. The psychologic symptoms include anxiety, confusion, depression, fear	11-point scale	Daily	Interviewer, self-administered

Continued

Table 2 *Continued*

PROM	Measure, items, and domains	Response format	Recall period	Administration method
	of disability, fear of sudden death, fear of loss of control, fear of loss of independence, forgetfulness, and mood disturbances. Validated for patients with concurrent cancer and heart failure.			
CARDIAC HEALTH PROFILE CONGESTIVE HEART FAILURE[16]	The Cardiac Health Profile is made up of three parts: Part I assesses subjectively reported functional ability, part II measures generic QoL for patients with heart disease in general by using a visual analog scale; and part III measures disease-specific QoL such as specific symptoms in heart failure.	Likert scale	Not stated	Self-administered
HF SELF-MONITORING TOOL[17]	A self-administered scale comprises 2 domains and covers 38 items. Domain 1 deals with 'awareness' and 'measurement' of aspects of self-monitoring, Domain 2 with 'interpretation' of aspects of self-monitoring.	Likert scale	Daily	Self-administered
MULTIDIMENSIONAL INDEX OF LIFE QUALITY (MILQ)[18]	Physical function, physical health, social, emotional, psychological, mental domains covered.	4-point importance scale and a 7-point satisfaction scale.	Not stated	Telephone, interviewer, self-administered
CANADIAN CARDIOVASCULAR SOCIETY QUESTIONNAIRES[19]	The CCS grading system employs four grades from I (without limitation of physical activity) to IV (inability to carry out any physical without discomfort) to assess health-related quality of life	4 point categories	Not stated	Paper and self-administered
MYOCARDIAL INFARCTION DIMENSIONAL ASSESSMENT SCALE(MIDAS)[20]	The MIDAS contains 35 questions measuring seven areas of health status: physical activity, insecurity, emotional reaction, dependency, diet, concerns over medication and side effects.	Likert	1 week	Paper, electronic
MAC NEW HEART DISEASE HEALTH RELATED QUALITY OF LIFE (MACNEW)[21]	27 items which fall into 3 domains: a 13-item physical limitations domain scale, a 14-item emotional function domain scale, and a 13-item social function domain scale.	Likert	2 weeks	Paper, self-administered
THE QUALITY OF LIFE AFTER MYOCARDIAL INFARCTION QUESTIONNAIRE (QLMI_V1)[22]	In total 26 items measuring 2 domains: limitations (including symptoms and restrictions) and emotions (including emotional function, confidence and self-esteem)	Likert	12 months	Paper, self-administered
THE QUALITY OF LIFE AFTER MYOCARDIAL INFARCTION QUESTIONNAIRE (QLMI_2)[23]	18 items measuring, three domains vitality, emotional distress, and sleep.	Likert	6 months	Paper, self-administered

Continued

Table 2 *Continued*

PROM	Measure, items, and domains	Response format	Recall period	Administration method
AF IMPACT[24]	20 items measuring, Symptoms, treatment concerns, treatment satisfaction, daily activities, overall QoL	Likert	1 week	Self-administered
TORONTO ATRIAL FIBRILLATION SEVERITY SCALE OF QUALITY OF LIFE (AFSS)[25]	A bedside scale that ranges from class 0 to 4, from no effect on functional quality of life to a severe effect on life quality.	Likert 4 point	4 weeks	Self-administered
CANADIAN CARDIOVASCULAR SOCIETY SEVERITY IN ATRIAL FIBRILLATION SCALE (CCS-AF)[26]	AF-QoL-7 items deal with the psychological domain while those of the AF-QoL-11 deal with physical activity. The questionnaire comprising the AF-QoL-7, and the AF-QoL-11 domains is identified as AFQoL-18.	Likert 4-point scale	Not stated	Bed side scale
QUALITY OF LIFE QUESTIONNAIRE FOR PATIENTS WITH ATRIAL FIBRILLATION (AF-QOL)[27]	4 conceptual domains (Symptoms, Daily Activities, Treatment Concern, and Treatment Satisfaction) from which individual domain and global scores can be calculated.	Likert	4 weeks	Paper, self-administered
ATRIAL FIBRILLATION EFFECT ON QUALITY-OF-LIFE (AFEQT)[28]	Covers 5 main domains: symptoms, patient perception of AF, quality of life with the disease, social limitation, and functional health.	Likert	8 weeks	Self-administered
TORONTO AF SYMPTOMS CHECKLIST-16 (SCL)[28]	Covers 6 domains in a short questionnaire: Physical, emotional, psychological, daily activities, fatigue.	Likert, 5-point scale	1 week	Self-administered
AF-6 SYMPTOMS SCALE[29]	Symptoms, global QoL,	Likert, 0 to 10	1 week	Self-administered
ATRIAL FIBRILLATION-SPECIFIC MEASURE OF PATIENT-REPORTED HEALTH-RELATED QUALITY OF LIFE (QLAF)[30]	The QLAF questionnaire consisted of seven domains, 22 numbered questions and 83 items, or simple questions which made up the domains. Domains were numbered sequentially (I–VII) covering the main clinical manifestations (palpitation, breathlessness, chest pain, and dizziness) and therapeutic interventions (drugs, direct-current cardioversion, and ablation)	Yes/No	12 weeks	Self-administered
PATIENT PERCEPTION OF ARRHYTHMIA QUESTIONNAIRE (PPAQ)[31]	PPAQ measures frequency and duration of episodes, symptoms, impact on daily activities, and restricted activity days validated in patients with any arrhythmia.	Yes/No	Daily	Interviewer, self-administered
ARRHYTHMIA-SPECIFIC QUESTIONNAIRE IN TACHYCARDIA AND ARRHYTHMIA (ASTA)[32]	Covers 9 Symptoms and the effect of symptoms on daily activities validated in patients with any arrhythmia.	4-point Likert scale	12 weeks	Interviewer, self-administered
ATRIAL FIBRILLATION	47 items measuring health literacy in these dimensions:	(yes, no, don't know)	Not stated	Interviewer, self-administered

Continued

Table 2 *Continued*

PROM	Measure, items, and domains	Response format	Recall period	Administration method
HEALTH LITERACY QUESTIONNAIRE (AFHLQ)[33]	What is AF, Symptoms of AF, Why do people get AF, Management of AF, What measures can slow or prevent the progression of AF.			
CORONARY REVASCULARISATION OUTCOME QUESTIONNAIRE[34]	The CROQ has four versions (CROQ-CABG_Pre, CROQ-PTCA_Pre, CROQ-CABG_Post, CROQ-PTCA_Post) contain 32 core evaluative items and one descriptive item that is not included in scale scores. The post-revascularisation versions of the CROQ (CROQ-CABG_Post 52 items, CROQ-PTCA_Post 47 items) contain these 33 core items plus additional evaluative items about adverse effects and satisfaction with outcome and two descriptive items.	3–6-point Likert	4 weeks	Self-administered
TORONTO AORTIC STENOSIS QUALITY OF LIFE QUESTIONNAIRE[35]	The CROQ is scored to produce 6 domains as follows: symptoms (7 items), physical functioning (8 items), psychosocial functioning (14 items), cognitive functioning (3 items), satisfaction (6 items), and adverse effects (11 or 6 items).	Likert	4 weeks	Self-administered
SEATTLE ANGINA QUESTIONNAIRE SHORT VERSION[36]	TASQ is a 16-item self-administered questionnaire that assesses AS-specific QoL across five domains: physical symptoms; physical limitations; emotional impact; social limitations, and health expectations.	Likert	4 weeks	Self-administered
SEATTLE ANGINA QUESTIONNAIRE[37]	Physical, Symptoms. 5 domains. Physical limitation, anginal stability, anginal frequency treatment satisfaction and disease perception. Chest pain, chest tightness, and angina.	Likert, 1 to 5/6	4 weeks	Self-administered

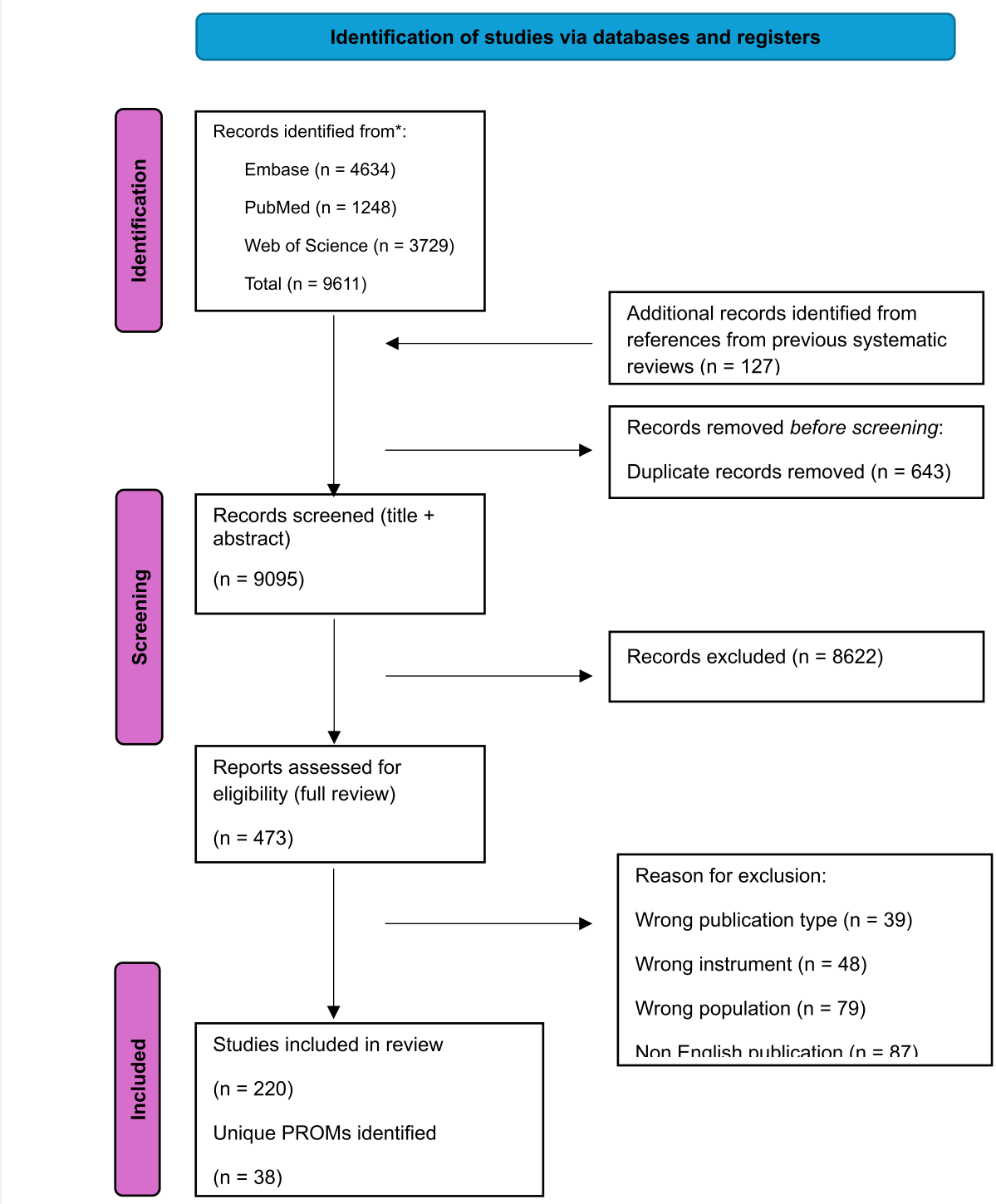


Figure 1 Preferred Reporting Items for Systematic and Meta-Analysis chart of included articles.

gold standard biomarker or other questionnaire such as the Short Form 36.[56] The main method used was comparing the PROMs with other measurement scales that measure similar constructs. The main statistical methods used include correlations using Pearson and Spearman correlation coefficient and multitrait–multimethod analysis. After applying the criteria for good measurement properties, strong positive evidence was found for most PROMs (*Table 3*).

Hypothesis testing for construct validity (known group, discrimination)

The majority of PROMs provided evidence for hypothesis testing for construct validity/discrimination/known group validity (22 PROMs; 58%) by comparing it with a gold standard biomarker or other questionnaire.[56] The main methods for demonstrating discriminant or

Table 3 Overall levels of evidence per measurement property and PROM

PROM	Content	Reliability	Internal consistency	Structural	Criterion	Cross-cultural	Measurement error	Discrimination	Responsiveness	Grade
MLHF[1]										B
CHAT[2]										B
KCCQ-12[3]										A
KCCQ-23[4]										A
QLQ-SHF[5]										B
LVD-36[6]										A
PROMIS plus HF profile[7]										A
HF Somatic Perception Scale -18[8]										B
Self-care of HF Index self-care management score-22[9]										A
Symptom Status Questionnaire-Heart Failure[10]										B
CaReQoL CHF[11]										B
CHFQOLQ-20[12]										B
Heart Failure-Daily Symptom Diary[13]										C
HeartQoL[14]										A
MDASI-HF[15]										B
CHPchf[16]										B
HF self-monitoring tool[17]										B
MILQ[18]										C
CCS questionnaires[19]										C
MIDAS[20]										B
MacNew[21]										A
QLMI_v1[22]										B
QLMI_v2[23]										C
AF Impact[24]										B
AFSS[25]										C
CCS-AF[26]										C
AFQoL[27]										B
AFEQT[28]										B
Toronto AF symptom checklist[28]										B
AF-6 symptoms scale[29]										B
QLAF[30]										B

Continued

Table 3 *Continued*

PROM	Content	Reliability	Internal consistency	Structural	Criterion	Cross-cultural	Measurement error	Discrimination	Responsiveness	Grade
PPAQ[31]										B
ASTA[32]										A
AFHLQ[33]										C
CROQ[34]										A
TASQ[35]										C
SAQ-7[36]										A
SAQ-19[37]										A

Level of evidence rating: Green: Strong; Yellow: Adequate; Red: Doubtful or Inadequate.
COSMIN recommendation category: A—Most suitable to be recommended, B—may have potential to be recommended but further validation studies are required, C—Not suitable to be recommended.
Abbreviations: MLHF, Minnesota Living with Heart Failure; KCCQ, Kansas City Cardiomyopathy Questionnaire; LVD, Left Ventricular Dysfunction questionnaire; PROMIS, Patient Reported Outcome Measurement Information Systems Plus Heart Failure; CHFQOLQ-20, Chronic heart failure health-related quality of life questionnaire; CHAT, Chronic heart failure assessment tool; MDASI-HF, MD Anderson Symptom Inventory-Heart Failure; CaReQoL CHF, Care-Related Quality of Life survey for Chronic Heart Failure; CHPchf, Cardiac Health Profile congestive heart failure; QLQ-SHF, Quality of life questionnaire in severe heart failure; MILQ, Multidimensional Index of Life Quality; CCS, Canadian Cardiovascular Society; MIDAS, Myocardial infarction dimensional assessment scale; MacNew, MacNew Heart disease health related quality of life questionnaire; QLMI, Quality of Life after Myocardial Infarction Questionnaire; AF, Atrial Fibrillation; AFSS, Atrial Fibrillation Severity Scale of quality of life-20; AFQoL-18, Quality of Life questionnaire for Patients with Atrial Fibrillation; AFEQT, Atrial Fibrillation Effect on QualiTy-of-life; ASTA, Arrhythmia-Specific questionnaire in Tachycardia and Arrhythmia; PPAQ, Patient Perception of Arrhythmia Questionnaire; AFHLQ, Atrial Fibrillation Health Literacy Questionnaire; QLAF, AF specific health related quality of life; CROQ, Coronary Revascularisation Outcome Questionnaire; SAQ, Seattle Angina Questionnaire; TASQ, Toronto aortic stenosis quality of life questionnaire.

known group validity include using analysis of variance comparing scores of known groups and multitrait–multimethod analysis or predictive models using regression analysis. The known severity groups were categorized using mostly the New York Heart Association (NYHA) and compared severity PROM scores across the four NYHA severity groups.[12] Other studies determined the predictive validity of the PROM using logistic regression or Cox proportional hazards models and reported the area under the curve.

Responsiveness

Most PROMs did not have adequate evidence of responsiveness in their validation studies (20 PROMs; 53%). The main methods for demonstrating responsiveness were based on hypothesis testing comparing changes on the PROM and a gold standard, or change scores of pre and post treatments, baseline and follow up, standardized response mean, effect sizes, or a clinically important change/difference. Cohen effect size criteria of $c > 0.80$ large, 0.2 poor, and 0.5 moderate effect size were used.

Recommendations

Of the 38 PROMs reviewed in this study, 11 (29%) were given an overall rating score of A. These were; the KCCQ questionnaires,[33,34] HeartQOL,[57,58] LVD-36,[59] PROMIS HF profile,[60] Self-care for HF Index,[61] MacNew,[62] CROQ,[42] SAQ questionnaires,[40,41] and the ASTA questionnaire.[44] Most evaluated HF (six PROMs), four PROMs evaluated IHD and one evaluated AF. An overall rating of B was given to 19 PROMs (50%) and a C rating was given to 8 PROMs (21%, *Table 3*).

Discussion

This scoping review and COSMIN analysis of 38 PROMs from 220 studies for the evaluation of HRQoL in individuals across a range of CVD found that 11 instruments (29%) had excellent psychometric properties across all nine COSMIN criteria with most PROMs (50%) requiring further validation prior to recommending their routine use within cardiology. The psychometric properties that were prioritized were content validity, reliability, IC, discrimination, and structural validity due to its clinical implications. The quality of patient data generated from such instruments, therefore, is reduced potentially limiting its ability to inform clinical care and the generalisability of trial results.

Similar to previous reviews[23,25] we found that the majority of HRQoL CVD PROMs (71%) available did not satisfy all nine domains of the COSMIN checklist for robust psychometric properties.[28] One reason may be that some PROMs were developed and validated before the COSMIN guidelines were developed. For systematic reviews with COSMIN analysis for disease-specific instruments such as AF[23] most instruments were rated as good, in line with our findings, with a specific focus on cross-cultural validity, measurement error and responsiveness for further validation. We found that most HF questionnaires were still advised to undergo further validation.[22]

However, our study differed by placing emphasis on a comprehensive psychometric evaluation using the COSMIN analysis whereas others used the Evaluating the Measurement of Patient-Reported Outcomes tool[22] which places more emphasis on administrative burden as well as psychometric properties of questionnaires over COSMIN. This could therefore cause the results to differ.[63] Furthermore, other studies focused on disease-specific PROMs[21–23] whereas we concentrated on common cardiovascular conditions. One review[21] occurred before the development and validation of the TASQ questionnaire[43] for patients with AS treated with TAVI and therefore our review offers a more contemporary insight. One review investigated adherence of a range cardiovascular conditions, including congenital heart disease, to the US Food and Drug Authority regulatory

criteria.[64] Only two PROMs fulfilled all the COSMIN criteria according to the previous review (KCCQ-23[33] and MacNew[36]) whereas ours identified others which were well validated.[60] The COSMIN analysis does contain an element of subjectivity and previously demonstrated a low inter-rater reliability which could explain the difference in the analyses. This suggests the need for additional training for experts and an independent reviewer to ratify results.[65]

We graded 19 (50%) PROMs as a B and recommended that they may be used in research and clinical practice but require further validation on cross-cultural validity/measurement invariance and measurement error. For example, limited evidence of cross-cultural invariance was provided in under half of HRQoL CVD PROMs (16 PROMs, 42%) whereas most questionnaires presented limited evidence for measurement invariance (28 PROMs, 74%). A notable exception was the PROMIS Plus HF questionnaire which was shown to be measurement invariant by sex, age, and education level.[60]

The heart specific PROMs that adhered to all nine COSMIN criteria were: both the 12 and 23 item KCCQ,[33,34] HeartQOL,[57,58] LVD-36,[59] PROMIS HF profile,[60] Self-care for HF Index self-care management score,[61] MacNew,[62] CROQ,[42] SAQ questionnaires,[40,41] and the ASTA questionnaire.[44] These findings are in broad agreement with other studies[23,25] which rated these PROMs adequate using the COSMIN criteria.

The commonly used methods for evaluating the structural validity of a PROM were exploratory factor analysis and confirmatory analysis, and only two studies used IRT models. IRT is specified for use in the COSMIN criteria and advantages to using IRT include providing insight into cross-cultural validity using measurement invariance/differential item function, item back and computerized adaptive testing that can assess measurement error.[66] There was no evidence found for IRT for most PROMs in this review. The methodological criteria for assessing responsiveness used the criterion approach by comparing the PROM with a gold standard using statistical measures such as correlations or area under the curve, sensitivity and specificity depending on the data types (continuous, or categorical), or determining the minimum important change for a PROM. For most of the PROMs this validation using these statistical methods was not conducted. The minimal important change that mattered to patients was not established hence there is no known threshold of improvement that is clinically relevant.[26]

Clinical and research implications

While there are many PROMS available for the measurement of HRQoL in CVD, their psychometric properties vary within and across the disease states. There are few with cross-cultural validity and the majority provide limited evidence for measurement error. Good practice dictates that an instrument should be translated and culturally sensitive to the target population[67] as the results from a poorly understood questionnaire are less reliable and valid.[26] Measurement error refers to a change in score from an instrument that is not due to random error[28] and is especially important given the subjective nature of PROMs. This can obscure the effect of an intervention due to noise which contributes to type II errors.[68] This is further exacerbated by some instrument's inadequate rating for content validity, reliability, and IC.[69] A recent review found that over half of HF randomized control trials published in highly cited journals utilized a PROM, hence weaknesses in a PROM's measurement error, for example, may obscure the safety and efficacy of evaluated treatments.[70] Patient outcomes generated from inadequately validated questionnaires may directly impact patient care as major organizations, such as the ESC, have advocated for their increased uptake within routine care for AF, HF, and TAVI.[13–16] Understandably this has repercussion for the use of PROMs in clinical care and research especially.

We found only one disease-specific HRQoL instrument for AS. TAVI is expanding to wider populations given recent safety and efficacy data[71] and using PROMs during the assessment of patients with AS has become a marker of good clinical care.[13] However, TASQ was ranked as having inadequate psychometric properties, therefore its patient reported data should be used with caution. A generic PROM may be used until further validation studies have been performed.

As a minimum, the PROMs rated B in our study require further validation in, where not available, cross-cultural/measurement invariance, measurement error, responsiveness, and a minimum clinically important difference. We propose that the PROMs that were rated C require validation work before these are used in research or clinical practice.

Strengths and limitations

The methodological quality of the PROMs was assessed in validation studies from database inception to February 2025 using the COSMIN checklist. However, we note the limitations of our work. First, only PROMs measuring health, health-related states, quality of life, and symptoms were assessed but no other forms of PROMs such as utility tools or PROMs that assess functional and mental health. Also, only PROMs available in English were included, hence we may have missed PROMs validated in other languages and therefore the cross-cultural validity of some instruments may be underreported. We did not evaluate if a PROM originally validated for one CVD could be validated for another CVD.

Conclusion

In this scoping review and analysis of the psychometric properties of 38 PROMs for the evaluation of HRQoL in individuals with common CVDs, only 11 PROMs met all nine requirements of the COSMIN criteria for robust psychometric properties. The main deficient areas were lack of evidence on cross-cultural validity/measurement invariance, measurement error, and information on responsiveness. Therefore, there should be caution in implementing most PROMS that evaluate HRQoL in individuals with CVDs within routine care and trials. As a minimum, most available PROMs require additional validation work, and for some CVDs there is a limited selection of suitable PROMs for HRQoL measurement.

Supplementary material

Supplementary material is available at *European Journal of Cardiovascular Nursing* online.

Author contributions

A.B. updated the search results, performed a targeted search for additional studies, cleaned the data, and drafted the manuscript and amended it after co-author comments. A.B.S. performed the second COSMIN analysis for all studies, M.S. devised the search strategy and assisted in including and excluding articles, M.H. assisted with the search strategy, C.W. provided supervision for the project and gave critical feedback, and C.P.G. and T.M. both devised the project and provided critical feedback to the manuscript with T.M. performing the initial COSMIN analysis of all developmental studies.

Funding

No funding received for this project.

Conflict of interest: There are no declarations with respect to this manuscript. A.B., A.B.S., M.S., M.H., and T.M. report no conflicts of interest. C.W. reports grants from British Heart Foundation and National Institute for Health Research and has unpaid roles with EHJ Quality of Care and Clinical Outcomes (Associate Editor), and NICE Indicator Advisory Committee. C.P.G. has received grants from Alan Turing Institute, British

Heart Foundation, National Institute for Health Research, Horizon 2020, Abbott Diabetes, Bristol Myers Squibb, and European Society of Cardiology and reports consulting fees from AI Nexus, AstraZeneca, Amgen, Bayer, Bristol Myers Squibb, Boehrinher-Ingleheim, CardioMatics, Chiesi, Daiichi Sankyo, GPRI Research B.V., Menarini, Novartis, iRhythm, Organon, and The Phoenix Group and speaker's fees from AstraZeneca, Boston Scientific, Menarini, Novartis, Raisio Group, Wondr Medical, and Zydus. C.P.G. is also a Deputy Editor: EHJ Quality of Care and Clinical Outcomes, NICE Indicator Advisory Committee member, Chair ESC Quality Indicator Committee member and participated on a Data Safety Monitoring Board or Advisory Board for DANBLCOK trial and TARGET CTCA trial.

Data availability

The data underlying this article will be shared on reasonable request. Please email t.munyombwe@leeds.ac.uk for consideration.

References

1. Wilkinson C, Bhatty A, Smith AB, Dwight J, Sanders J, Gale CP. Embracing the promise of patient reported outcome measures in cardiology. *Eur Heart J Qual Care Clin Outcomes* 2024;**10**:651–652.
2. Black N. Patient reported outcome measures could help transform healthcare. *BMJ* 2013;**346**:f167–f167.
3. Munyombwe T, Dondo TB, Aktaa S, Wilkinson C, Hall M, Hurdus B, *et al.* Association of multimorbidity and changes in health-related quality of life following myocardial infarction: a UK multicentre longitudinal patient-reported outcomes study. *BMC Med* 2021;**19**:227.
4. Palermo TM, Long AC, Lewandowski AS, Drotar D, Quittner AL, Walker LS. Evidence-based assessment of health-related quality of life and functional impairment in pediatric psychology. *J Pediatr Psychol* 2008;**33**:983–996.
5. Naughton MJ, Shumaker SA. The case for domains of function in quality of life assessment. *Qual Life Res* 2003;**12**:73–80.
6. U.S. Department of Health and Human Services FDA Center for Drug Evaluation and Research, U.S. Department of Health and Human Services FDA Center for Biologics Evaluation and Research, U.S. Department of Health and Human Services FDA Center for Devices and Radiological Health. Guidance for industry: patient-reported outcome measures: use in medical product development to support labeling claims: draft guidance. *Health Qual Life Outcomes* 2006;**4**:79.
7. Soloveva A, Gale CP, Naung Tun H, Hurdus B, Aktaa S, Palin V, *et al.* Associations of health-related quality of life with major adverse cardiovascular and cerebrovascular events for individuals with ischaemic heart disease: systematic review, meta-analysis and evidence mapping. *Open Heart* 2023;**10**:e002452.
8. Velikova G, Booth L, Smith AB, Brown PM, Lynch P, Brown JM, *et al.* Measuring quality of life in routine oncology practice improves communication and patient well-being: a randomized controlled trial. *J Clin Oncol* 2004;**22**:714–724.
9. Thompson DR, Yu C-M. Quality of life in patients with coronary heart disease-I: assessment tools. *Health Qual Life Outcomes* 2003;**1**:42.
10. Dondo TB, Munyombwe T, Hurdus B, Aktaa S, Hall M, Soloveva A, *et al.* Association of baseline and changes in health-related quality of life with mortality following myocardial infarction: multicentre longitudinal linked cohort study. *Eur Heart J Qual Care Clin Outcomes* 2025;**11**:730–738.
11. Kemp I, Appleby C, Lane S, Lisboa P, Stables RH. A comparison of angina symptoms reported by clinicians and patients, pre and post revascularisation: insights from the stent or surgery trial. *Int J Cardiol* 2019;**293**:25–31.
12. Caraballo C, Desai NR, Mulder H, Alhanti B, Wilson FP, Fiuzat M, *et al.* Clinical implications of the New York heart association classification. *J Am Heart Assoc* 2019;**8**:e014240.
13. Ali N, Aktaa S, Younsi T, Beska B, Batra G, Blackman DJ, *et al.* European society of cardiology quality indicators for the care and outcomes of adults undergoing transcatheter aortic valve implantation. *Eur Heart J Qual Care Clin Outcomes* 2024;**10**:723–736.
14. McDonagh TA, Metra M, Adamo M, Gardner RS, Baumbach A, Böhm M, *et al.* 2021 ESC guidelines for the diagnosis and treatment of acute and chronic heart failure: developed by the task force for the diagnosis and treatment of acute and chronic heart failure of the European Society of Cardiology (ESC) with the special contribution of the heart failure association (HFA) of the ESC. *Eur Heart J* 2021;**42**:3599–3726.
15. Abdin A, Wilkinson C, Aktaa S, Böhm M, Polovina M, Rosano G, *et al.* European society of cardiology quality indicators update for the care and outcomes of adults with heart failure. The heart failure association of the ESC. *Eur J Heart Fail* 2024;**26**:1867–1875.
16. Van Gelder IC, Rienstra M, Bunting KV, Casado-Arroyo R, Caso V, Crijns HJGM, *et al.* 2024 ESC guidelines for the management of atrial fibrillation developed in collaboration with the European association for cardio-thoracic surgery (EACTS): developed by the task force for the management of atrial fibrillation of the European Society of Cardiology (ESC), with the special contribution of the European heart rhythm association (EHRA) of the ESC. Endorsed by the European stroke organisation (ESO). *Eur Heart J* 2024;**45**:3314–3414.
17. European Medicines Agency. Committee for medicinal products for human use. Reflection paper on the regulatory guidance for the use of health-related quality of life (HRQL) measures in the evaluation of medicinal products. 2005.
18. The PROTEUS guide to implementing patient-reported outcomes in clinical practice: a synthesis of resources. 2023.
19. Moons P, Norekvål TM, Arbelo E, Borregaard B, Casadei B, Cosyns B, *et al.* Placing patient-reported outcomes at the centre of cardiovascular clinical practice: implications for quality of care and management: a statement of the ESC association of cardiovascular nursing and allied professions (ACNAP), the association for acute CardioVascular care (ACVC), European association of percutaneous cardiovascular interventions (EAPCI), European association of preventive cardiology (EAPC), heart failure association (HFA), European heart rhythm association (EHRA), European association of cardiovascular imaging (EACVI), ESC regulatory affairs committee, ESC advocacy committee, ESC digital health committee, ESC education committee, and the ESC patient forum. *Eur Heart J* 2023;**44**:3405–3422.
20. Mokkink LB, Elsman EBM, Terwee CB. COSMIN guideline for systematic reviews of patient-reported outcome measures version 2.0. *Qual Life Res* 2024;**33**:2929–2939.
21. de Heer F, Gökalp AL, Kluin J, Takkenberg JJ. Measuring what matters to the patient: health related quality of life after aortic valve and thoracic aortic surgery. *Gen Thorac Cardiovasc Surg* 2019;**67**:37–43.
22. Garin O, Herdman M, Vilagut G, Ferrer M, Ribera A, Rajmil L, *et al.* Assessing health-related quality of life in patients with heart failure: a systematic, standardized comparison of available measures. *Heart Fail Rev* 2014;**19**:359–367.
23. Sale A, Yu J. Quality of life instruments in atrial fibrillation: a systematic review of measurement properties. *Health Qual Life Outcomes* 2022;**20**:143.
24. Moons P, Norekvål TM, Arbelo E, Borregaard B, Casadei B, Cosyns B, *et al.* Placing patient-reported outcomes at the centre of cardiovascular clinical practice: implications for quality of care and management. *Eur Heart J* 2023;**44**:3405–3422.
25. Chew DS, Whitelaw S, Vaduganathan M, Mark DB, Van Spall HGC. Patient-reported outcome measures in cardiovascular disease: an evidence map of the psychometric properties of health status instruments. *Ann Intern Med* 2022;**175**:1431–1439.
26. Prinsen CAC, Mokkink LB, Bouter LM, Alonso J, Patrick DL, de Vet HCW, *et al.* COSMIN guideline for systematic reviews of patient-reported outcome measures. *Qual Life Res* 2018;**27**:1147–1157.
27. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan—a web and mobile app for systematic reviews. *Syst Rev* 2016;**5**:210.
28. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, *et al.* The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol* 2010;**63**:737–745.
29. Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, *et al.* Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol* 2007;**60**:34–42.
30. Schünemann H, Brożek J, Guyatt G, Oxman A. GRADE handbook, introduction to GRADE handbook. Handbook for grading the quality of evidence and the strength of recommendations using the GRADE approach. Updated October 2013. 2013.
31. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, *et al.* GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 2008;**336**:924–926.
32. Wilson IB, Cleary PD. Linking clinical variables with health-related quality of life. A conceptual model of patient outcomes. *JAMA* 1995;**273**:59–65.
33. Green CP, Porter CB, Bresnahan DR, Spertus JA. Development and evaluation of the Kansas city cardiomyopathy questionnaire: a new health status measure for heart failure. *J Am Coll Cardiol* 2000;**35**:1245–1255.
34. Spertus JA, Jones PG. Development and validation of a short version of the Kansas city cardiomyopathy questionnaire. *Circ Cardiovasc Qual Outcomes* 2015;**8**:469–476.
35. Watanabe-Fujinuma E, Origasa H, Bamber L, Roessig L, Toyoda T, Haga Y, *et al.* Psychometric properties of the Japanese version of the Kansas city cardiomyopathy questionnaire in Japanese patients with chronic heart failure. *Health Qual Life Outcomes* 2020;**18**:236.
36. Höfer S, Lim L, Guyatt G, Oldridge N. The MacNew heart disease health-related quality of life instrument: a summary. *Health Qual Life Outcomes* 2004;**2**:3.
37. Thompson DR, Jenkinson C, Roebuck A, Lewin RJ, Boyle RM, Chandola T. Development and validation of a short measure of health status for individuals with acute myocardial infarction: the myocardial infarction dimensional assessment scale (MIDAS). *Qual Life Res* 2002;**11**:535–543.
38. Lim LL, Valenti LA, Knapp JC, Dobson AJ, Plotnikoff R, Higginbotham N, *et al.* A self-administered quality-of-life questionnaire after acute myocardial infarction. *J Clin Epidemiol* 1993;**46**:1249–1256.
39. Valenti L, Lim L, Heller R, Knapp J. An improved questionnaire for assessing quality of life after acute myocardial infarction. *Qual Life Res* 1996;**5**:151–161.

40. Chan PS, Jones PG, Arnold SA, Spertus JA. Development and validation of a short version of the Seattle angina questionnaire. *Circ Cardiovasc Qual Outcomes* 2014;**7**:640–647.
41. Spertus JA, Winder JA, Dewhurst TA, Deyo RA, Prodzinski J, McDonell M, *et al.* Development and evaluation of the Seattle angina questionnaire: a new functional status measure for coronary artery disease. *J Am Coll Cardiol* 1995;**25**:333–341.
42. Schroter S, Lamping DL. Coronary revascularisation outcome questionnaire (CROQ): development and validation of a new, patient based measure of outcome in coronary bypass surgery and angioplasty. *Heart* 2004;**90**:1460–1466.
43. Styra R, Dimas M, Svitak K, Kapoor M, Osten M, Ouzounian M, *et al.* Toronto aortic stenosis quality of life questionnaire (TASQ): validation in TAVI patients. *BMC Cardiovasc Disord* 2020;**20**:209.
44. Walfridsson U, Arestedt K, Stromberg A. Development and validation of a new arrhythmia-specific questionnaire in tachycardia and arrhythmia (ASTA) with focus on symptom burden. *Health Qual Life Outcomes* 2012;**10**:44.
45. Coyne KS, Edvardsson N, Rydén A. Development and validation of the AFImpact: an atrial fibrillation-specific measure of patient-reported health-related quality of life. *Value Health* 2017;**20**:1355–1361.
46. Dorian P, Guerra PG, Kerr CR, O'Donnell SS, Crystal E, Gillis AM, *et al.* Validation of a new simple scale to measure symptoms in atrial fibrillation: the Canadian cardiovascular society severity in atrial fibrillation scale. *Circ Arrhythm Electrophysiol* 2009;**2**:218–224.
47. Badia X, Arribas F, Ormaetxe JM, Peinado R, de Los Terreros MS. Development of a questionnaire to measure health-related quality of life (HRQoL) in patients with atrial fibrillation (AF-QoL). *Health Qual Life Outcomes* 2007;**5**:37.
48. Spertus J, Dorian P, Bubien R, Lewis S, Godejohn D, Reynolds MR, *et al.* Development and validation of the atrial fibrillation effect on QualiTy-of-Life (AFEQT) questionnaire in patients with atrial fibrillation. *Circ Arrhythm Electrophysiol* 2011;**4**:15–25.
49. Dorian P, Paquette M, Newman D, Green M, Connolly SJ, Talajic M, *et al.* Quality of life improves with treatment in the Canadian trial of atrial fibrillation. *Am Heart J* 2002;**143**:984–990.
50. Härdén M, Nyström B, Bengtson A, Medin J, Frison L, Edvardsson N. Responsiveness of AF6, a new, short, validated, atrial fibrillation-specific questionnaire–symptomatic benefit of direct current cardioversion. *J Interv Card Electrophysiol* 2010;**28**:185–191.
51. Braganca EO, Filho BL, Maria VH, Levy D, de Paola AA. Validating a new quality of life questionnaire for atrial fibrillation patients. *Int J Cardiol* 2010;**143**:391–398.
52. Wood KA, Stewart AL, Drew BJ, Scheinman MM, Frolicher ES. Development and initial psychometric evaluation of the patient perspective of arrhythmia questionnaire. *Res Nurs Health* 2009;**32**:504–516.
53. McMichael G, Cusack L, Andina Munawar D, Boyd M, Palmer L, Lim HS, *et al.* Atrial fibrillation health literacy questionnaire (AFHLQ): the development of an AF-specific health literacy questionnaire. *Int J Cardiol Heart Vasc* 2024;**50**:101322.
54. Avis NE, Smith KW, Hambleton RK, Feldman HA, Selwyn A, Jacobs A. Development of the multidimensional index of life quality. A quality of life measure for cardiovascular disease. *Med Care* 1996;**34**:1102–1120.
55. Campeau L. The Canadian cardiovascular society grading of angina pectoris revisited 30 years later. *Can J Cardiol* 2002;**18**:371–379.
56. Failde I, Ramos I. Validity and reliability of the SF-36 health survey questionnaire in patients with coronary artery disease. *J Clin Epidemiol* 2000;**53**:359–365.
57. Oldridge N, Hofer S, McGee H, Conroy R, Doyle F, Saner H. The HeartQoL: part I. Development of a new core health-related quality of life questionnaire for patients with ischemic heart disease. *Eur J Prev Cardiol* 2014;**21**:90–97.
58. Oldridge N, Höfer S, McGee H, Conroy R, Doyle F, Saner H. The HeartQoL: part II. Validation of a new core health-related quality of life questionnaire for patients with ischemic heart disease. *Eur J Prev Cardiol* 2014;**21**:98–106.
59. O'Leary CJ, Jones PW. The left ventricular dysfunction questionnaire (LVD-36): reliability, validity, and responsiveness. *Heart* 2000;**83**:634–640.
60. Ahmad FS, Kallen MA, Schifferdecker KE, Carluzzo KL, Yount SE, Gelow JM, *et al.* Development and initial validation of the PROMIS®-plus-HF profile measure. *Circ Heart Fail* 2019;**12**:e005751.
61. Świątoniowska-Lonc N, Polański J, Pilarczyk-Wróblewska I, Jankowska-Polańska B. The revised self-care of heart failure index—a new tool for assessing the self-care of Polish patients with heart failure. *Kardiol Pol* 2021;**79**:841–847.
62. Höfer S, Saleem A, Stone J, Thomas R, Tulloch H, Oldridge N. The MacNew heart disease health-related quality of life questionnaire in patients with angina and patients with ischemic heart failure. *Value Health* 2012;**15**:143–150.
63. Valderas JM, Ferrer M, Mendívil J, Garin O, Rajmil L, Herdman M, *et al.* Development of EMPRO: a tool for the standardized assessment of patient-reported outcome measures. *Value Health* 2008;**11**:700–708.
64. U.S. Department of Health and Human Services US FDA, Center for Drug Evaluation and Research, Center for Biologics Evaluation and Research. Treatment for heart failure: endpoints for drug development. FDA. 2024. Accessed November 2024.
65. Mokkink LB, Terwee CB, Gibbons E, Stratford PW, Alonso J, Patrick DL, *et al.* Inter-rater agreement and reliability of the COSMIN (COnsensus-based standards for the selection of health status measurement instruments) checklist. *BMC Med Res Methodol* 2010;**10**:82.
66. Jabrayilov R, Emons WHM, Sijtsma K. Comparison of classical test theory and item response theory in individual change assessment. *Appl Psychol Meas* 2016;**40**:559–572.
67. Wild D, Grove A, Martin M, Eremenco S, McElroy S, Verjee-Lorenz A, *et al.* Principles of good practice for the translation and cultural adaptation process for patient-reported outcomes (PRO) measures: report of the ISPOR task force for translation and cultural adaptation. *Value Health* 2005;**8**:94–104.
68. Staggs VS, Cramer E. Reliability of pressure ulcer rates: how precisely can we differentiate among hospital units, and does the standard signal-noise reliability measure reflect this precision? *Res Nurs Health* 2016;**39**:298–305.
69. McGee RG. How to include patient-reported outcome measures in clinical trials. *Curr Osteoporos Rep* 2020;**18**:480–485.
70. Eliya Y, Averbuch T, Le N, Xie F, Thabane L, Mamas M, *et al.* Temporal trends in the inclusion of patient-reported outcomes in heart failure randomized trials published in high-impact medical journals: a systematic bibliometric review. *J Am Heart Assoc* 2021;**10**:e022353.
71. Vahanian A, Beyersdorf F, Praz F, Milojevic M, Baldus S, Bauersachs J, *et al.* 2021 ESC/EACTS guidelines for the management of valvular heart disease. *Eur Heart J* 2022;**43**:561–632.