JACC: HEART FAILURE
VOL. 4, NO. 10, 2016

ISSN 2213-1779/$36.00
PUBLISHED BY ELSEVIER
http://dx.doi.org/10.1016/j.jchf.2016.04.010

CLINICAL RESEARCH

# Patient-Reported Outcomes in Chronic Heart Failure

## Applicability for Regulatory Approval

Mitchell A. Psotka, MD, PhD,[a] Robyn von Maltzahn, PhD,[b] Milena Anatchkova, PhD,[c] Irene Agodoa, MD,[d] Dina Chau, PharmD,[d] Fady I. Malik, MD, PhD,[e] Donald L. Patrick, MPH, PhD,[f] John A. Spertus, MPH, MD,[g] Ingela Wiklund, PhD,[b] John R. Teerlink, MD[a]

## ABSTRACT

**OBJECTIVES** The study sought to review the characteristics of existing patient-reported outcome (PRO) instruments used with chronic heart failure (HF) patients and evaluate their potential to support an approved U.S. Food and Drug Administration (FDA) product label claim.

**BACKGROUND** PROs, including symptoms and their associated functional limitations, contribute substantially to HF patient morbidity. PRO measurements capture the patient perspective and can be systematically assessed with structured questionnaires, however rigorous recommendations have been set by the FDA regarding the acceptability of PRO measures as a basis for product label claims.

**METHODS** Extensive searches of databases and specialty guidelines identified PRO instruments used in patients with chronic HF. Information on critical properties recommended by the FDA guidance were systematically extracted and used to evaluate the selected PRO instruments.

**RESULTS** Nineteen PRO instruments used with chronic HF patients were identified. The Kansas City Cardiomyopathy Questionnaire and Minnesota Living with Heart Failure Questionnaire were the most extensively evaluated and validated in studies of this population. However, judged by criteria listed in the FDA PRO guidance, no existing PRO measure met all of the criteria to support a product label claim in the United States.

**CONCLUSIONS** Currently available chronic HF PRO measures do not fulfill all the recommendations provided in the FDA PRO guidance and therefore may not support an FDA-approved product label claim. Future investigations are merited to develop a PRO measure for use in patients with chronic HF in accordance with the FDA guidance.
(J Am Coll Cardiol HF 2016;4:791–804) © 2016 by the American College of Cardiology Foundation.

From the [a]School of Medicine, University of California San Francisco and Section of Cardiology, San Francisco Veterans Affairs Medical Center, San Francisco, California; [b]Evidera, London, United Kingdom; [c]Evidera, Bethesda, Maryland; [d]Amgen Inc., Thousand Oaks, California; [e]Cytokinetics, Inc., South San Francisco, California; [f]Department of Health Services, University of Washington, Seattle, Washington; and the [g]Saint Luke's Mid America Heart Institute and Department of Biomedical and Health Informatics, University of Missouri, Kansas City, Missouri. This study was sponsored by Amgen Inc. Dr. von Maltzahn is a former employee and Drs. Anatchkova and Wiklund are current employees of Evidera, which received funding from Amgen Inc. to conduct this study. Dr. Agodoa is a former employee and shareholder of Amgen, Inc. Dr. Chau is a current employee and shareholder of Amgen Inc. Dr. Malik is an employee and shareholder of Cytokinetics, Inc., which has collaborated with Amgen Inc. in the development of omecamtiv mecarbil. Dr. Patrick is a consultant to Amgen Inc. Dr. Spertus served as a consultant to Amgen Inc. for this study; is a consultant to Janssen, Amgen Inc., Novartis Pharmaceuticals Corporation, Bayer Healthcare, and Regeneron Pharmaceuticals, Inc.; has received grant support from Gilead Sciences, Inc., Eli Lilly and Co., the National Institutes of Health, American Heart Association, the Patient-Centered Outcomes Research Institute, and the Aetna Foundation; has equity interest in Health Outcomes Sciences; and owns the copyright to the Seattle Angina Questionnaire, Peripheral Artery

**ABBREVIATIONS AND ACRONYMS**

**EMA** = European Medicines Agency

**FDA** = U.S. Food and Drug Administration

**HF** = heart failure

**HRQoL** = health-related quality of life

**KCCQ** = Kansas City Cardiomyopathy Questionnaire

**MLHFQ** = Minnesota Living with Heart Failure Questionnaire

**PRO** = patient-reported outcome

The primary goals of heart failure (HF) treatment are to prevent hospitalization and maximize survival as well as health status: patients' symptoms, function, and well-being, sometimes referred to as health-related quality of life (HRQoL) (1). Evaluating these patient-reported outcomes (PROs) requires collecting data directly from patients themselves and is incorporated into the most recent HF performance measures and scientific statements (1–5). Ascertaining HF patient health status is complementary to other endpoints in clinical management and research trials and has been encouraged by regulatory agencies when evaluating novel treatments (6–8).

A PRO is defined as "any report of the status of a patient's health condition that comes directly from the patient, without interpretation of the patient's response by a clinician or anyone else" (8). Patients are best able to judge the severity of their symptoms, the impact of those symptoms on their physical, social, and psychological function, and how those symptoms and limitations impair their HRQoL (1). PROs can be quantified through the use of structured questionnaires termed PRO instruments and these tools should be appropriately designed and demonstrate adequate psychometric properties that provide an accurate representation of a patient's health status.

SEE PAGE 805

Three primary uses of chronic HF PRO instruments include: 1) in clinical research, to evaluate treatment benefit; 2) for regulatory approval, to support a product label claim; and 3) in clinical practice, to monitor disease progression and assist in treatment decisions. Partially validated PRO measures may be suitable for some clinical practice or research settings, whereas regulatory approval requires fulfillment of rigorous recommendations (5–8). PRO instruments can be generic and widely applied or used to compare subjects across multiple medical conditions (9). However, disease-specific PRO measures are more sensitive to clinical change and more interpretable because they quantify symptoms specific to the disease of interest (10–12).

Because many PRO tools exist for use in patients with chronic HF, and their use is increasing, there is a need to delineate the properties of available instruments so providers and researchers can select those appropriately aligned with their needs (13–16). We therefore performed a targeted evaluation of available PRO measures used for the assessment of patients with chronic HF. We used the comprehensive recommendations from the U.S. Food and Drug Administration (FDA) guidance described in **Table 1** to compare the PRO instruments, catalogued the incorporated symptoms and symptom impacts, and described the measurement properties that guide their use for regulatory approval, research, and clinical practice. Because the FDA guidance views PRO instruments as a mechanism to quantify treatment benefits in medical product clinical trials, we focused on more sensitive and interpretable HF-specific PRO measures rather than generic instruments (8).

## METHODS

**INSTRUMENT SEARCH STRATEGY.** EMBASE and MEDLINE were searched in the period between January 2006 (the publication year for the draft FDA guidance) and May 2015 in iterative fashion (**Figure 1**), using search terms listed in Online Appendix, to identify PRO instruments administered to patients with chronic HF and their psychometric properties. Societal conferences and guidelines listed in the Online Appendix were also investigated. Manual searches were completed in the PROQoLID database and clinicaltrials.gov, and for specific missing information using Google.

**INSTRUMENT ASSESSMENT.** Identified titles and abstracts were screened according to predetermined eligibility criteria by at least 2 independent reviewers (Online Appendix). Eligible articles reported instruments that were administered to patients with acute or chronic HF and included information on symptoms, symptom impacts, or therapeutic effect on symptoms. Articles were excluded that reported studies in other populations, did not mention HF symptoms or symptom-impacts, reported no new data, were nonempirical, failed to specify methodology, or were not published in English.

Using the FDA guidance as an evaluative framework, we extracted relevant development and validation information from eligible articles (8). Captured

Questionnaire, and Kansas City Cardiomyopathy Questionnaire. Dr. Teerlink has received grant/research support and is a consultant to Amgen Inc., Bayer, Cytokinetics, Inc., Novartis Pharmaceuticals Corp., Relypsa, and Trevena Inc. All other authors have reported that they have no relationships relevant to the contents of this paper to disclose.

Manuscript received January 25, 2016; revised manuscript received March 14, 2016, accepted April 27, 2016.

data for each PRO instrument included the published reference(s), study design, details of the subject population, symptoms and impacts measured, domains covered, administration format, recall period, and scoring methods. Validation variables included the methods used in instrument development such as qualitative interviews, literature reviews and stakeholder input for content validity. Captured psychometric properties included reliability, internal consistency, construct validity, responsiveness, responder definitions, and the minimal clinically important difference. Fulfillment by each PRO instrument of recommendations made in the FDA guidance was agreed on by at least 2 independent reviewers, with disagreements settled by the author group.

## RESULTS

Nineteen PRO instruments used with chronic HF patients were identified by screening 2,552 articles and 2,334 conference abstracts (**Figure 1**). All PRO measures identified in conference abstracts and guidelines were also found by literature review. Every PRO tool was either initiated or completely developed prior to the release of the draft or finalized FDA guidance in 2006 and 2009, respectively (8).

Most of the selected PRO measures were developed with patients documented to have chronic HF (13 tools), whereas fewer were generated using patients with coronary artery disease (4 tools) or any cardiovascular disease (2 tools) who may also have had HF (**Table 2**). Most were validated employing HF patients with reduced left ventricular ejection fraction, however the Kansas City Cardiomyopathy Questionnaire (KCCQ) was also tested in patients with preserved left ventricular ejection fraction and advanced HF (17,18). The recall periods for identified PROs ranged from 1 day to 12 weeks. Six measures had ambiguous or undefined recall periods, and 2 utilized an FDA-preferred short (1 day) recall period for symptoms and not more than 1-week recall period for impacts, the European Heart Failure Self-care Behavior Scale (EHFScBS) and the Shortness of Breath in Heart Failure (SOB-HF) (19,20). All 19 PRO tools were administered on paper; 13 of the tools were self-administered and 5 could also be given by an interviewer. Estimated completion times for each PRO instrument ranged between 5 and 15 min. Administration times on the high end of the spectrum were found for the Chronic Heart Failure Questionnaire (CHQ), the only included PRO that required an interviewer, and those tools containing large numbers of response items such as the Chronic Heart Failure Assessment Tool (CHAT) (21,22).

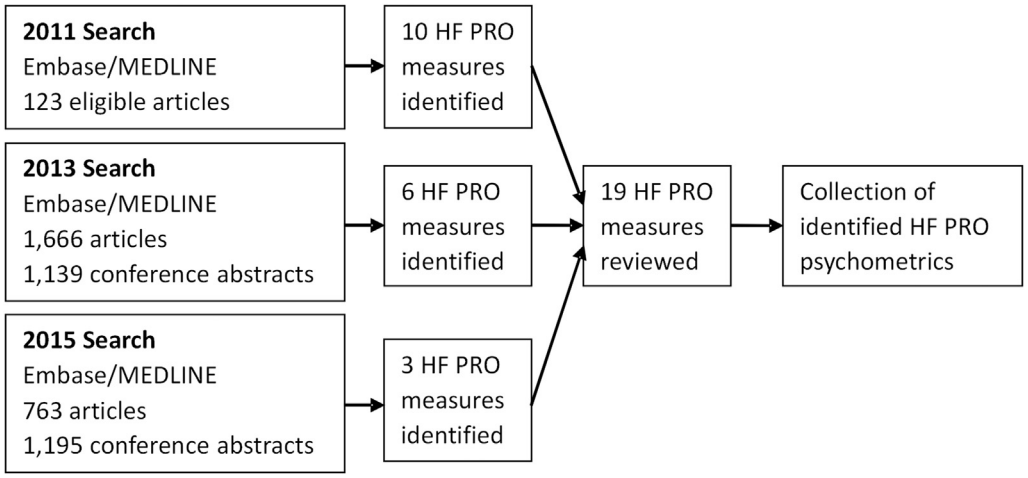
**TABLE 1 Critical Elements for PRO Instruments From the 2009 Food and Drug Administration Guidance**

| Component | Description |
|---|---|
| Documented content validity | Content validity is evidence that the PRO instrument measures the concept of interest including documentation from qualitative studies (open-ended heart failure patient interviews) that all the items and domains of an instrument are appropriate and comprehensive relative to its intended measurement concept, population, and use. Of note, PRO instrument items are often generated by literature review and expert opinion and then tested qualitatively utilizing patient input. |
| Developed for heart failure | Content validity is specific to the population, condition and treatment to be studied by the PRO instrument, thus input for PRO development should come from the target patient population, in this case patients with chronic heart failure. |
| Preferred administration mode | Self-administration or patient interview are both acceptable PRO administration modes, however if multiple administration modes are available they must produce comparable data. |
| Recommended recall period | PRO instruments that call for patients to rely on memory, especially if they must recall over a long period of time, compare their current state with an earlier period, or average their response over a period of time, may undermine content validity because responses are influenced by the patient's state at the time of recall. Items with short recall periods or items that ask patients to describe their current or recent state are preferred. |
| Demonstrated reproducibility | Test-retest reliability (and intrainterviewer as well as interinterviewer reliability if interviewer administered) as demonstrated by stability of PRO instrument scores over time if no change is expected in the concept of interest, appropriately measured by the intraclass correlation coefficient; often reported by a correlation coefficient. |
| Proven internal consistency | The extent to which items comprising a PRO instrument measure the same concept, measured by Cronbach's alpha. Poor internal consistency increases the likelihood of producing false negative results. |
| Demonstrated construct validity | Evidence that relationships among items, domains, and concepts conform to a priori hypotheses concerning logical relationships that should exist with measures of related concepts or scores produced in similar or diverse patient groups; meaning that the PRO instrument scores should predictably correlate with similar alternative constructs established for heart failure. |
| Proven ability to detect change | Evidence that a PRO instrument can identify differences in scores over time in individuals or groups who have changed with respect to the measurement concept. |
| Available responder definition | The empirically determined individual patient PRO instrument score change over a predetermined time period that should be interpreted as a treatment benefit, derived using anchor-based methods. The anchor chosen should be easier to interpret than the PRO measure itself. |

Adapted from the 2009 Food and Drug Administration guidance (8).
PRO = patient-reported outcome.

The most commonly incorporated symptoms were shortness of breath, swelling, fatigue, and chest pain (**Table 2**). Less frequent, but used in more than 1 instrument, were nausea, lightheadedness, concentration or memory deficits, and weakness. Recurrent proximal symptom-impacts included physical activity and patient self-efficacy, and some measures focused on symptom-impacts rather than symptoms. The distal impact on HRQoL was distinctly measured in 4 assessments, the Cardiac Health Profile (CHP), KCCQ, Quality of Life Index (QLI), and Quality of Life Questionnaire for Severe Heart Failure (QLQ-SHF). The most common quantification system

**FIGURE 1 HF PRO Search Strategy and Outcomes**




HF = heart failure; PRO = patient-reported outcome.

(14 measures) was a Likert scale, rarely combined with a visual analog scale. The Left Ventricular Dysfunction Questionnaire (LVD-36) used dichotomous true/false elements, the Daily Activity Questionnaire in Heart Failure (DAQIHF) captured activity duration, and the San Diego Heart Failure Questionnaire (SDHF) used multiple response and true/false questions. Scoring methods were heterogeneous.

At least 1 of the major components of FDA-defined content validity, including literature review, or expert or patient interviews, was not mentioned by 9 of the 19 PRO measures; although 8 purported to have performed all these components, they lacked documentation including the demographics and number of subjects included (**Table 3**). Among the PRO instruments reviewed, only 2 surveys had content validity likely acceptable to the FDA, the EHFScBS and SOB-HF. In contrast, measurement properties such as reliability and validity were more commonly reported. Test-retest reliability was reported for 11 measures assessed between 1 and 26 weeks after initial testing. Internal consistency was reported for 16 instruments. Some evidence of convergent validity was reported for all but 3 instruments, the Heart Failure Needs Assessment Questionnaire (HFNAQ), SOB-HF, and QLI, and was most often demonstrated by comparison to generic PRO measures. The most widely validated were the KCCQ and the Minnesota Living with Heart Failure Questionnaire (MLHFQ), which also exhibited independent predictive validity for hospitalization and death (23,24). Adequate responsiveness to a clinically significant change was recorded for 13 of the tools; however only 6 disclosed an anchor-based minimal clinically important difference, and only the KCCQ and MLHFQ delineated the anchored responder definition preferred by the FDA. The KCCQ equally detected both clinical worsening and improvement, however other PRO instruments were better at determining improvement rather than worsening (25). No identified PRO measure fulfilled all the FDA guidance criteria (**Figure 2**).

## DISCUSSION

Although many PRO instruments exist for use in patients with chronic HF and despite their potential utility for clinical care, all those investigated lacked 1 or more pieces of psychometric evidence, test characteristic, or sufficient documentation of content validity suggested by the FDA (16). Thus, researchers interested in a product label claim approval are advised that current PRO measures may not satisfy FDA requirements, even though they might have clinically significant associations or be acceptable for other regulatory bodies. These conclusions are on the basis of extensive literature from the field of psychometrics and qualitative instrument development (1,7,8,26). Common deficiencies included incomplete or inadequately documented content validity, longer recall periods than desired by the FDA, and lack of responder definitions (8).

There are many PRO guidelines published by national and international organizations, including the

European Medicines Agency and the International Society for Pharmacoeconomics and Outcomes Research; however, the FDA guidance is the most rigorous in terms of patient involvement to establish content validity and interpretability using responder definitions (5–8). The draft FDA guidance was released in 2006 and finalized in 2009, and from 2006 to 2010 sponsors of 45% of products applied for labels with data from PRO endpoints (27). Of new products approved by both the EMA and FDA from 2006 to 2010, PRO label claims were approved for 47% and 19%, respectively (28). All PRO claims approved by the FDA were also granted by the EMA, however tolvaptan was the only HF medication to have an application submitted during this period and it received a PRO label claim only from the EMA. The FDA appeared to prefer symptom and impact-related claims whereas the EMA endorsed claims on the basis of HRQoL and global assessments (28). These differences suggest that sponsors seeking international label claims should either utilize tools with both lower and higher order concepts or follow the FDA guidance with primarily symptom and impact assessment.

Understanding the inadequacies of the PRO instruments described here may be useful to help generate novel measures suitable for the FDA. Such measures could be useful for clinical practitioners, researchers, and for regulatory approval. To address the listed deficiencies a PRO tool should include appropriate responder definitions as well as incorporate and document the components required for content validity and saturation. Based on the current analysis concept saturation would likely include shortness of breath, swelling, and fatigue (**Table 2**), items also identified in standalone qualitative interviews of 63 chronic HF patients, along with the key impacts of physical and emotional function (29). Because of the presence of comorbid causes of pain such as ischemia, it remains unclear whether chest pain is an essential manifestation of HF (**Table 2**). The FDA prefers short recall periods because of concern for recall bias, as retrospective evaluation can underestimate prior health status and correlates best with current health status assessment (8,26,30). However, the construct validity of HF PRO tools with varied recall periods argues that this element of the FDA guidance may benefit from revision (**Tables 2** and **3**). Moreover, a recent comparison of the Seattle Angina Questionnaire with daily diaries suggested high agreement between a 4-week recall period and the daily records of patients' angina (31).

Nevertheless, the strengths of currently available PRO instruments should not be lost to the decrees of the regulatory review process. There are few reasonably validated and feasible HF PRO measures, best represented by the KCCQ and MLHFQ, and nonregulatory use of these instruments should continue in clinical and research settings (3,22). The KCCQ and MLHFQ have demonstrated excellent construct validity, responsiveness to change, strong correlations with chronic HF symptoms, and after multivariable adjustment have reliable independent predictive power for traditional outcomes, including costs of care (10,23,32–34) (**Table 3**). The KCCQ and MLHFQ can measure efficacy of an intervention and these associations suggest they capture and quantify aspects of the disease process not otherwise accessed by clinical evaluation. Additionally, prediction of hospital readmission may rely as much or more on PRO-captured elements as on traditional clinical variables (3,10,35). Clinicians and researchers can employ these HF PRO measures to assist clinical decision making and risk prediction, determine efficacy, refine outcome probabilities, facilitate shared decision-making, identify deficits in self-care behavior, assess health disparities, and meaningfully assist care providers and patients to understand and improve their experience (36,37).

Validated and dependable instruments should replace the methods commonly used in clinical practice and research that poorly appraise HF patients' health status (13,15). For example, perhaps because almost one-half of patients do not recognize dyspnea as a symptom of HF, PRO elements are incompletely reported without a formal PRO instrument (38). In contrast to evaluation by a PRO instrument, physician-determined New York Heart Association functional classification is inconsistent and incongruent with patient self-assessment, and though most cardiologists routinely query how far patients say they can walk, responses do not correlate with cardiopulmonary performance (39–41). Finally, only 16% of modern cardiovascular clinical trials utilized a PRO measure, 11% of which were ad hoc and lacked evidence of validity (15). An individual or group seeking to employ a PRO in practice or investigation can survey the compiled tables and determine which measure will best assess the desired concept in the subject or population of interest.

Still, despite their expansive evaluation as descriptive devices in multiple diverse settings, the role of PRO measurement as a clinical management tool remains undefined (16). Given that PRO instruments can distinguish patients at elevated risk for traditional outcomes, it is logical that they could be used to identify those patients who merit increased attention, more intensive therapy, or other

**TABLE 2 PRO Measure Attributes**

| Measure (Abbreviation) | Intended Population | Recall Period | Format | Domains (Items) | Symptoms | Impacts | Scoring (Range) | Ref. #* |
|---|---|---|---|---|---|---|---|---|
| Cardiac Health Profile (CHP) | Angina | Not stated | Self-administer | Angina (1)<br>QoL (16)<br>Possible intervention (2) | Chest pain | HRQoL, concentration, memory, depression, anxiety, fear, sleep, social relationships, sexual activity, treatment decisions | 19 items<br>Angina 5-point Likert (unscored)<br>Total score: (VAS millimeters)/Items (1–100) Lower is better | (46) |
| Cardiac Self-Efficacy Questionnaire (CSEQ) | Coronary heart disease | Not stated | Self-administer | Maintain function (5)<br>Control symptoms (8)<br>Some items not included in subscores. | Chest pain, breathlessness | Self-efficacy, social activities, physical activities, sexual activity | 16 items<br>5-pont Likert scale<br>Total and domain scores<br>(not reported) | (47) |
| Chronic Heart Failure Assessment Tool (CHAT) | Heart failure | 2 weeks | Self-administer | Symptoms, activity levels, psychosocial aspects, emotions<br>(item numbers incompletely published) | Shortness of breath with and without activity, fatigue, fear | Physical activities, family activities, social activities, sleep, sexual relationships, anger, fear, tension, concern | 46 items<br>5-point Likert scale or 11-point numerical rating<br>Total and domain scores<br>(not reported) | (48) |
| Chronic Heart Failure Questionnaire (CHQ) | Chronic heart failure | 2 weeks | Interviewer administer | Fatigue (4)<br>Emotional function (7)<br>Dyspnea (5) | Shortness of breath, dyspnea, fatigue, frustration, depression, anxiety | Physical activities, emotional function | 16 items (4 added later)<br>7-point Likert scale<br>Total and domain scores<br>(16–112) higher is better | (49,50) |
| Daily Activity Questionnaire in Heart Failure (DAQIHF) | Heart failure | 1 week | Self- or interviewer administer | Sleep (5)<br>Basic activity (10)<br>Housework activity (14)<br>Leisure time activity (26)<br>Work/social activity (6)<br>Transportation (7)<br>Miscellaneous (2)<br>Autonomy (12) | Unpublished | Unpublished | 82 items<br>Activity duration<br>Daily or domain energy expenditure = intensity × duration<br>(not reported)<br>higher is better | (51) |
| European Heart Failure Self-care Behavior Scale (EHFScBS) | Heart failure inpatients and outpatients | 1 day | Self-administer | Regimen compliance (6)<br>Asking for help (4)<br>Adapting activities (2) | Shortness of breath, leg/feet swelling, weight gain, fatigue | Self-care behavior | 12 items<br>5-point Likert scale<br>Total score, no domains<br>(0–12) lower is better | (19) |
| Heart Failure Needs Assessment Questionnaire (HFNAQ) | Heart failure | 4 weeks | Self-administer | Physical (10)<br>Psychological (9)<br>Social (8)<br>Existential (3) | Thirst, chest pain, depression, shortness of breath, fatigue, nausea, leg swelling, anxiety, sleep | Difficulty to bath or toilet, coping ability, physical activities, sexual activity | 30 items<br>5-point Likert scale<br>Total and domain scores<br>(30–150) lower is better | (52) |
| Heart Failure Somatic Perception Scale (HFSPS) | Heart failure | Not stated | Self-administer | Acute symptoms<br>Peripheral edema<br>MI-like symptoms<br>Early HF symptoms | Breathlessness (at night, lying down), chest pain, nausea, rapid heartbeat, cough, fatigue, feet swelling, weight gain | Usual daily activities, dressing, eating | 12 items (6 added later)<br>4-point Likert scale<br>Total score<br>(0–36, after additions 0–90) higher is better | (53,54) |
| HeartQoL (HeartQoL) | Coronary heart disease (angina, MI, HF) | 4 weeks | Self- or interviewer administer | Physical (10)<br>Emotional (4) | Fatigue, shortness of breath, worry, frustration, depression, relaxation | Physical activities | 14 items<br>4-point Likert scale<br>Global and subscale: mean<br>(0–3) higher is better | (44,55) |
| Kansas City Cardiomyopathy Questionnaire (KCCQ) | Acute or chronic HF, including reduced and preserved LVEF and valve disease | 2 weeks | Self-administer | Physical (6)<br>Symptoms (7)<br>Change (1)<br>Self-efficacy (2)<br>Social (1), QoL (7) | Shortness of breath, fatigue, ankle swelling | Physical limitation, self-efficacy, social interference, QoL | 23 items,<br>5, 6, or 7-point Likert scale<br>Overall and domain scores<br>(0–100) higher is better | (17,56) |

*Continued on the next page*

**TABLE 2 Continued**

| Measure (Abbreviation) | Intended Population | Recall Period | Format | Domains (Items) | Symptoms | Impacts | Scoring (Range) | Ref. #* |
|---|---|---|---|---|---|---|---|---|
| Left Ventricular Dysfunction Questionnaire (LVD-36) | Heart failure, reduced LVEF | "These days" | Self-administer | No subdomains identified | Leg swelling, leg fatigue, nausea, shortness of breath, fatigue, weakness, temperature, night wakening, fragility | Social function, emotional function, daily physical activity | 36 items<br>True/false<br>Total score: true %<br>(0–100) lower is better | (57,58) |
| MacNew Heart Disease HRQL Questionnaire (MacNew) | Coronary heart disease | 2 weeks | Self-administer | Physical (13)<br>Emotional (14)<br>Social (13)<br>Global (27) | Fatigue, chest pain, leg ache, restlessness, shortness of breath, lightheadedness | Frustration, depression, confidence, personal happiness, social activity, fear, physical activity, sexual activity | 27 items<br>7-point Likert scale<br>Global or domain score: mean<br>(1–7) higher is better | (59) |
| Minnesota Living with Heart Failure Questionnaire (MLHFQ) | Heart failure NYHA functional class I-III, reduced LVEF | 4 weeks | Self-administer | Physical (8)<br>Emotional (5)<br>Some items not included in subscores. | Ankle/leg swelling, shortness of breath, fatigue, poor memory or concentration, depression | Physical activity, sleep, sexual activity, financial difficulty, leisure activity, eating | 21 items,<br>6-point Likert scale<br>Overall and domain scores<br>(0–105) lower is better | (60) |
| Quality of Life Index – Cardiac Version (QLI) | Cardiovascular disease | Not stated | Self-administer | Health and functioning<br>Socioeconomic<br>Psychological/spiritual<br>Family/relationships | Chest pain, shortness of breath, fatigue | Overall health, self-care, self-efficacy, family, happiness, finances, emotional support, anxiety, work | 35 items<br>6-point Likert scale<br>Total and domain scores weighted by importance<br>(0–30) higher is better | (61) |
| Quality of Life Questionnaire for Severe Heart Failure (QLQ-SHF) | Heart failure NYHA functional class II-IV | 1 week | Self-administer | Psychological (7)<br>Physical activity (7)<br>Somatic symptoms (7)<br>Life satisfaction (5) | Shortness of breath at rest, breathlessness, chest pain, fatigue/ exhaustion, weakness, dullness, somatic symptoms | Psychological complaints, life satisfaction, physical activity | 26 items<br>6-point Likert scale and VAS<br>(0–130) lower is better | (62) |
| San Diego Heart Failure Questionnaire (SDHF) | Cardiomyopathy | "Since last visit" | Self-administer | Functional Capacity (11)<br>Disability (1)<br>Some items not included in subscores. | Shortness of breath, fatigue with exertion, chest pain, edema, lightheadedness, orthopnea | Alcohol consumption, weight gain, physical activities (shopping, stairs, housework) | 28 items,<br>Each scored between 0 and 5<br>Total score, no domains<br>(0–66) lower is better | (63,64) |
| Seattle Angina Questionnaire (SAQ) | Coronary heart disease, stress testing, angioplasty | 4 weeks | Self-Administer | Physical (9)<br>Angina frequency (2)<br>Angina stability (1)<br>Treatment Satisfaction (4)<br>Disease perception/QoL (3) | Chest pain, chest tightness, angina | Physical limitation, treatment satisfaction, disease perception | 19 items<br>5- or 6-point Likert scale<br>Each scale transformed to 0–100 scale<br>(0–100) higher is better | (65,66) |
| Self-Care Heart Failure Index (SCHFI) | Heart failure | 12 weeks | Self- or interviewer administer | Self-maintenance (10)<br>Self-management (6)<br>Self-confidence (6) | Shortness of breath, ankle swelling | Remedy use<br>Recognition of symptoms, self-efficacy | 22 items<br>4-point Likert scale<br>Domain scores transformed to 0–100<br>(0–300) higher is better | (38,67) |
| Shortness of Breath in Heart Failure (SOB-HF) | Heart failure | 1 day | Self- or interviewer administer | Breathlessness (11) | Shortness of breath (suffocating, needing air, breathless), chest tightness, rapid breathing | None | 11 items<br>5-point Likert scale<br>Total score: mean × 100<br>(0–100) lower is better | (20) |

*See the Online Appendix for references 46–80.

HF = heart failure; LVEF = left ventricular ejection fraction; MI = myocardial infarction; NYHA = New York Heart Association; QoL = quality of life; VAS = visual analog scale.

**TABLE 3 PRO Measure Psychometric Characteristics**

| Measure (Abbreviation) | Content Validity | Test-Retest Reliability | Internal Consistency (Cronbach's α) | Construct Validity Population | Responsiveness Change in Status | MID (Responder) | Ref. #* Year |
|---|---|---|---|---|---|---|---|
| Cardiac Health Profile (CHP) | Patient interviews† | 4 weeks: (n = 80) r = 0.927, p = 0.0009 | Each item (0.87-0.89) QoL (0.89) | Outpatient angina (n = 80) Total score: Nottingham Health Score (r = 0.753, p = 0.0001) | CABG (n = 22) mean change: 5.9 (p = 0.02); Angina (n = 76) or Control (n = 51) mean difference: 13.2 (p < 0.0001) | Not reported (Not reported) | (46) 1996 |
| Cardiac Self-Efficacy Questionnaire (CSEQ) | Not reported | Not reported | Control symptoms (0.90) Maintain function (0.87) Total (0.77) | Outpatients following catheterization (n = 198) Maintain function: SF36 Physical (Wald = 3.25, p < 0.001) | Outpatients stable angina or unstable angina (n = 214): (F = 3.24, p < 0.05) | Not reported (Not reported) | (47,68) 1998 |
| Chronic Heart Failure Assessment Tool (CHAT) | Patient interviews, 3 phases (n = 302) | Not reported | Symptoms (0.93) Activity levels (0.92) Psychosocial (0.86) Emotions (0.84) | Outpatient HF (n = 68) Total score: SF36 total (r = 0.79, CI: 0.68-0.87) Total score: MLHFQ total (r = 0.57, CI: 0.39-0.71) | Not reported | Not reported (Not reported) | (48) 2007 |
| Chronic Heart Failure Questionnaire (CHQ) | Literature 88 Patient interviews† Expert review† Items = 67% of variance (n = 211) | Incompletely reported. Coefficient of variance: (n = 25) Dyspnea: 14% Fatigue: 18% Emotional: 18% | Dyspnea (0.86) Fatigue (0.86) Emotional function (0.92) Total (0.93) | Outpatient HF (n = 211) Total score: MLHFQ (r = -0.81) Dyspnea score: MLHFQ (r = -0.63) Total score: NYHA (ANOVA p = 0.0001) Outpatient HF (n = 88) Total score: 6MWT (r = 0.60, p < 0.05) | Improvement: (n = 29) Dyspnea (0.8, p = 0.004) Fatigue (1.11, p = 0.002) Emotional (0.43, p = 0.006). Worsening: (n = 37) Dyspnea (0.53, p = 0.001) Fatigue (0.91, p = 0.001) Emotional (0.42, p = 0.004). | 0.5 points/item = MID, 1.0 points/item = substantial Total Score: 0.21-2 Dyspnea: 0.11-3 Fatigue: 0.38-2 Emotional: 0.19-4 | (12,21,26,49,50) 1988 |
| Daily Activity Questionnaire in Heart Failure (DAQIHF) | Literature Patient interviews† Expert review† | 6 weeks: (n = 24) Variation: 1.37% R = 0.989 (ns) | Not reported | Outpatient HFrEF (n = 105) Total score: peak $Vo_2$ (r = 0.72, p < 0.0001) Awake score: peak $Vo_2$ (r = 0.72, p < 0.0001) | 2-month rehabilitation (n = 21): Total score: change in peak Vo2 (R = 0.88, p < 0.0001) | Not reported (Not reported) | (51) 2004 |
| European Heart Failure Self-care Behavior Scale (EHFScBS) | Literature Patient interviews Expert review | 2 weeks: median change 1 point (p = 0.06-0.99) | Total (0.69-0.93) | Outpatient and inpatient HF (n = 140) Unable to significantly differentiate extra HF care (t = 2.6, p = 0.09) Outpatient HF (n = 159) Not convergent with MLHFQ or SCHFI Total Scores | Not reported | Not reported (Not reported) | (19,69) 2002 |
| Heart Failure Needs Assessment Questionnaire (HFNAQ) | Literature Patient interviews† Expert review† | Not reported | Physical (0.70) Psychological (0.70) Social (0.71) Existential (0.77) Total (0.80) | Unpublished | Not reported | Not reported (Not reported) | (52,70) 2004 |
| Heart Failure Somatic Perception Scale (HFSPS) | Literature Expert review | Not reported | Total score (0.71-0.83) | Hospitalized HF (n = 75) Total score: 6-month hospitalization ($r_p$ = 0.42, p < 0.001) Total score: NYHA functional class (r = 0.35, p = 0.002) | Not reported | Not reported (Not reported) | (53,54,71) 2006 |

*Continued on the next page*

**TABLE 3 Continued**

| Measure (Abbreviation) | Content Validity | Test-Retest Reliability | Internal Consistency (Cronbach's α) | Construct Validity Population | Responsiveness Change in Status | MID (Responder) | Ref. #* Year |
|---|---|---|---|---|---|---|---|
| HeartQoL (HeartQoL) | Patient interviews (n = 6384) from MacNew, MLHFQ, and SAQ | Not reported | Physical score (0.89–0.90)<br>Emotional score (0.80–0.82)<br>Global score (0.90–0.91) | Ischemic heart disease (n = 6,384)<br>Physical score: SF36 Physical (r = 0.68, p < 0.001)<br>Emotional score: SF36 Mental (r = 0.60, p < 0.001)<br>Global score: NYHA functional class II or III/IV (p < 0.001)<br>Outpatient HF (n = 1,922)<br>Physical score: SF36 Physical (r = 0.67, p < 0.001)<br>Emotional score: SF36 Mental (r = 0.60, p < 0.001) | Angioplasty (n = 398)<br>Mean change:<br>global: 0.3 ± 0.7<br>(p < 0.001)<br>Rehabilitation (n = 383)<br>Mean change:<br>global: 0.4 ± 0.5<br>(p < 0.001) | Not reported<br>(Not reported) | (44,55)<br>2014 |
| Kansas City Cardiomyopathy Questionnaire (KCCQ) | Literature<br>Patient interviews†<br>Expert review† | 3 months:<br>(n = 39) Summary score range = 0.8–4 (p > 0.05)<br>6 weeks:<br>(n = 79)<br>Physical $r_i$ = 0.85<br>Symptoms $r_i$ = 0.83<br>QoL $r_i$ = 0.76<br>Social $r_i$ = 0.86<br>Summary $r_i$ = 0.91 | Physical (0.87–0.90)<br>Symptoms (0.88–0.89)<br>QoL (0.78–0.82)<br>Social (0.86–0.88)<br>Self-efficacy (0.62–0.64)<br>Functional status (0.92–0.93)<br>Summary (0.94–0.95)<br>For HFpEF all domains (>0.69)<br>Summary (0.96) | Decompensated HFrEF (n = 129)<br>Physical score: NYHA functional class (r = −0.65, p < 0.001)<br>Physical score: 6MWT (r = 0.48, p < 0.001)<br>QoL score: MLHFQ Emotional (r = 0.62, p < 0.001)<br>Summary score: NYHA functional class (F = 41.9, p < 0.001)<br>Outpatient HFrEF (n = 505)<br>Summary score <25: 1-year Mortality or HF Hospitalization (HR: 2.77, p = 0.02)<br>Outpatient HFpEF (n = 200)<br>Summary score: NYHA functional class (r = −0.62, p < 0.001)<br>Summary score: death or hospitalization (log rank p < 0.001)<br>HF after MI (n = 1,358)<br>5-point decrease summary score: 2-yr mortality (HR: 1.11) | 6 weeks: (n = 298)<br>Summary scores outperformed EQ-5D and RAND12;<br>c-statistic:<br>0.77 (small change)<br>0.90 (large change)<br>3 months: (n = 39)<br>Mean change of summary score on hospital admission 24.3 (p < 0.001), outperformed MLHFQ and SF36 | Summary score change:<br>5 = mild<br>10 = medium<br>15 = large<br>(Summary score change ≥5) | (17,23,25,36,56,72,73)<br>1999 |
| Left Ventricular Dysfunction Questionnaire (LVD-36) | Literature<br>Patient interviews†<br>65% stated: Saturated.<br>Expert review† | 1 week:<br>(n = 52)<br>$r_i$ = 0.95 | Kuder-Richardson coefficient:<br>(Cronbach's α for dichotomous variables)<br>Total score 0.95 | Outpatient HFrEF (n = 60)<br>Total score: NYHA functional class (F = 0.46, p < 0.0001)<br>Total score: SF36 physical (r = −0.75, p < 0.0001)<br>Total score: LVEF (r = 0.22, ns)<br>Outpatient HFrEF (n = 38)<br>Total score: $Vo_2$ (r = 0.52, p < 0.001)<br>Total score: treadmill duration (r = 0.45, p < 0.01) | Compared to global change at 6 months (n = 49):<br>(F = 5.7, p < 0.001)<br>Outperformed MLHFQ | Total score mild decrease: 10 points<br>(Not reported) | (58)<br>1998 |
| MacNew Heart Disease HRQL Questionnaire (MacNew) | Literature<br>Patient interviews†<br>Expert review† | 3–6 months:<br>(n = 68)<br>$r_i$ = 0.79–0.88 | Physical (0.89–0.93)<br>Emotional (0.94–0.95)<br>Social (0.84–0.95)<br>Global (>0.80) | Patients requiring pacemakers (n = 68)<br>Physical score: SF36 Physical (r = 0.85, p < 0.01)<br>Emotional score: SF36 Mental (r = 0.82, p < 0.01) | Pacemaker (n = 68)<br>Global ES: 0.86<br>Emotional ES: 0.83<br>Physical ES 0.76<br>Social ES: 0.67 | Global or Domain score change: 0.5<br>(Not reported) | (45,74,75)<br>1993 |

*Continued on the next page*

**TABLE 3 Continued**

| Measure (Abbreviation) | Content Validity | Test-Retest Reliability | Internal Consistency (Cronbach's α) | Construct Validity<br>Population | Responsiveness<br>Change in Status | MID (Responder) | Ref. #*<br>Year |
|---|---|---|---|---|---|---|---|
| Minnesota Living with Heart Failure Questionnaire (MLHFQ) | Literature<br>Patient interviews†<br>Expert review†<br>Items = 70% of Variance (n = 211) | 1-3 weeks:<br>(n = 83)<br>Mean change: -1 (-11 to 5),<br>weighted κ = 0.84<br>4 weeks:<br>Total r = 0.79<br>Physical r = 0.74<br>Emotional r = 0.75 | Physical (0.94)<br>Emotional (0.88-0.90)<br>Total (0.94-0.95) | Outpatient pulmonary hypertension (n = 83)<br>Total score: 6MWT (r = -0.54, p = 0.01)<br>Total score: NYHA functional class (r = 0.54, p = 0.01)<br>Physical score: 6MWT (r = -0.58, p = 0.01)<br>Physical score: NYHA functional class (r = 0.63, p = 0.01)<br>Inpatient and outpatient HF (n = 83)<br>Total score: NYHA functional class (r = 0.60, p < 0.01)<br>Outpatient HF (n = 211)<br>Total score: NYHA functional class (ANOVA p = 0.0001)<br>Outpatient HF (n = 1,151)<br>Total score: 1-, 3-, 5-year mortality (HR: 1.012, p < 0.001)<br>Outpatient HFrEF (n = 5,025)<br>Total score: 1.5-year mortality (RR: 1.205, p < 0.001)<br>Total score: 1.5-year hospitalization (RR: 1.188, p < 0.001)<br>Outpatient HF (n = 198)<br>Total score: Naughton treadmill (r = 0.33, p < 0.01) | (n = 173)<br>Mean change of total score 3 months after hospitalization: 13.3 with further total score change from 3-6 months: 2.2 (p < 0.001) | Total score change: 4.84<br>Physical: 2.56<br>Emotional: 0.98<br>(Not reported) | (12,21,24,32,60,76-78)<br>1987 |
| Quality of Life Index - Cardiac Version (QLI) | Literature<br>Patient interviews† | Not reported | Health and functioning (0.87)<br>Socioeconomic (0.73)<br>Psychological/ spiritual (0.88)<br>Family/relationships (0.63)<br>Total (0.89-0.96) | Not reported<br>(published for alternative versions) | 6-week rehabilitation: (n = 16) insensitive for all domains;<br>6-month goal setting (n = 60)<br>F = 4.6, p < 0.01 | Not reported<br>(Not reported) | (61,79)<br>1985 |
| Quality of Life Questionnaire for Severe Heart Failure (QLQ-SHF) | Literature<br>Patient interviews†<br>Expert review† | 1 week:<br>Total score: $r_i$ = 0.82<br>Subdomains: $r_i$ = 0.75-0.85 | Overall (0.88) | Outpatient HF (n = 51)<br>Domain score: presence of angina (r = 0.42) | ACE inhibitor (meta-analysis, hypothesized):<br>Somatic ES: 0.3<br>Psychological ES: 0.2<br>Physical ES: 0.3<br>Total ES: 0.2-0.3 | Not reported<br>(Not reported) | (43,62)<br>1987 |
| San Diego Heart Failure Questionnaire (SDHF) | Not reported | Not reported | Not reported | Outpatient HFrEF (n = 2,708)<br>Total score: MLHFQ total (r = 0.57, p < 0.0001)<br>Total score: all-cause Mortality (HR: 1.6, p < 0.0001)<br>Total score: mortality or HF hospitalization (HR: 1.66, p < 0.0001) | Bucindolol therapy (n = 2,708):<br>Mean change 0.01,<br>Total score Wilcoxon rank sum p = 0.69 (ns) | Not reported<br>(Not reported) | (63,64)<br>1983 |

*Continued on the next page*

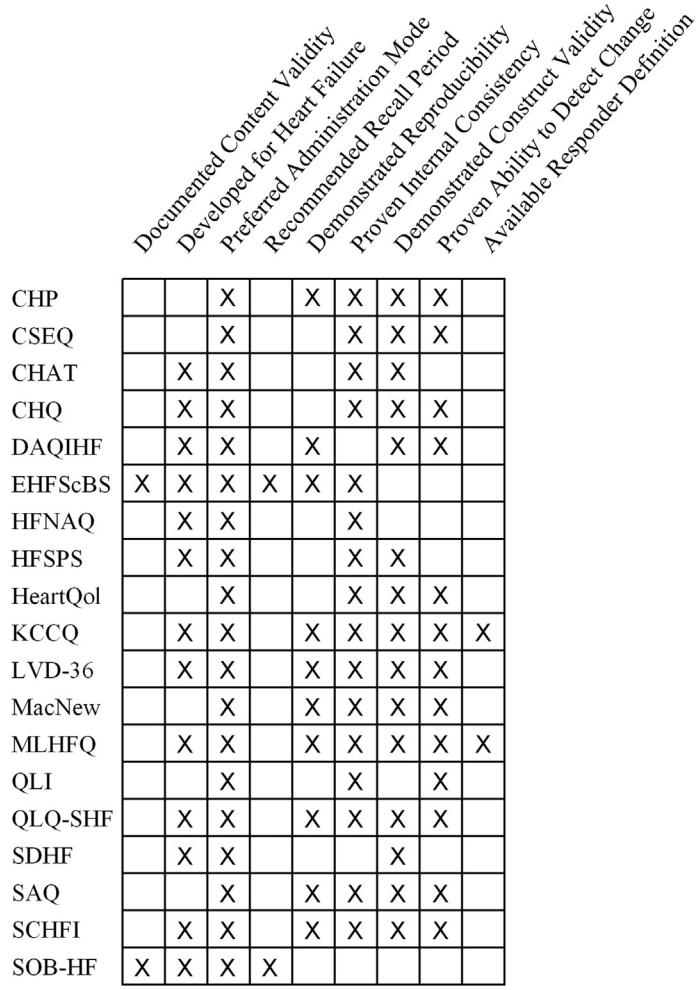
**TABLE 3 Continued**

| Measure (Abbreviation) | Content Validity | Test-Retest Reliability | Internal Consistency (Cronbach's α) | Construct Validity Population | Responsiveness Change in Status | MID (Responder) | Ref. #* Year |
|---|---|---|---|---|---|---|---|
| Seattle Angina Questionnaire (SAQ) | Literature | 3 months: (n = 117)<br>Domain score change 0.5–3.8 ($p \geq 0.1$ for all)<br>$r_i$ = 0.24–0.83 | Physical (0.89–0.91)<br>Angina stability (not calculated)<br>Angina frequency (0.69–0.87)<br>Treatment satisfaction (0.72–0.77)<br>Disease perception (0.66–0.67)<br>Summary scale (0.86) | Bruce treadmill test takers (n = 70)<br>Physical score: test duration (r = 0.42, p = 0.001)<br>Coronary artery disease (n = 45)<br>Physical score: angioplasty ($p < 0.0001$)<br>Angina stability score: Angioplasty ($p < 0.0001$)<br>Angina frequency score: angioplasty ($p < 0.0001$)<br>Disease perception score: angioplasty ($p < 0.0001$)<br>Coronary artery disease (n = 192)<br>Disease perception score: SF36 (r = 0.60, $p < 0.0001$) | Angioplasty (n = 45)<br>Mean change:<br>Physical: 17.9<br>Angina stability: 46.3<br>Angina frequency: 33.3<br>QoL: 36<br>(All $p < 0.0001$)<br>Treatment satisfaction: -1.5 (p = 0.66, NS) | Any subscore: change ≥10 points<br>(Not reported) | (65,66)<br>1995 |
| Self-Care Heart Failure Index (SCHFI) | Literature<br>Patient interview† | 15 days:<br>$r_i$ = 0.64–0.89 | Self-maintenance (0.55–0.56)<br>Self-management (0.60–0.70)<br>Self-confidence (0.82–0.83)<br>Total (0.76) | Inpatient and outpatient HF (n = 34)<br>Self-maintenance: EHFScBS (r = -0.65, $p < 0.001$) | Prior HF diagnosis <2 months or >2 months (n = 93) Total score mean change: 20<br>(*t* test p = 0.04) | Domain score ≥70 = self-care adequacy, no MID for change<br>(Not anchored) | (38,67,80)<br>2000 |
| Shortness of Breath in Heart Failure (SOB-HF) | Literature<br>Patient interviews<br>Expert review | Not reported | Not reported | Not reported | Not reported | Not reported<br>(Not reported) | (20)<br>2011 |

*See Online Appendix for references 46–80. †Incomplete documentation.

6MWT = 6-min walk test; ACE = angiotensin-converting enzyme; ANOVA = analysis of variance; CABG = coronary artery bypass graft; CI = confidence interval; ED-5D = EuroQol-5D; EHFScBS = European Heart Failure Self-care Behavior Scale; ES = effect size; HFpEF = heart failure with preserved left ventricular ejection fraction; HFrEF = heart failure with reduced left ventricular ejection fraction; HR = hazard ratio; MLHFQ = Minnesota Living with Heart Failure Questionnaire; MI = myocardial infarction; MID = minimal clinically important difference; ns = nonsignificant; NYHA = New York Heart Association; QoL = quality of life; RR = risk ratio; r = Spearman correlation coefficient; $r_i$ = intraclass correlation coefficient; $r_p$ = Pearson correlation coefficient; SF36 = short form 36.

**FIGURE 2 Summary of Identified Chronic HF PRO Measures and Suitability to Support a Product Label Claim by FDA Guidance Criteria**




| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| CHP | | | X | | X | X | X | X | |
| CSEQ | | | X | | | X | X | X | |
| CHAT | | X | X | | | X | X | | |
| CHQ | | X | X | | | X | X | X | |
| DAQIHF | | X | X | | X | | X | X | |
| EHFScBS | X | X | X | X | X | X | | | |
| HFNAQ | | X | X | | | X | | | |
| HFSPS | | X | X | | | X | X | | |
| HeartQol | | | X | | | X | X | X | |
| KCCQ | | X | X | | X | X | X | X | X |
| LVD-36 | | X | X | | X | X | X | X | |
| MacNew | | | X | | X | X | X | X | |
| MLHFQ | | X | X | | X | X | X | X | X |
| QLI | | | X | | | X | | X | |
| QLQ-SHF | | X | X | | X | X | X | X | |
| SDHF | | X | X | | | | X | | |
| SAQ | | | X | | X | X | X | X | |
| SCHFI | | X | X | | X | X | X | X | |
| SOB-HF | X | X | X | X | | | | | |

CHP = Cardiac Health Profile; CHAT = Chronic Heart Failure Assessment Tool; CHQ = Chronic Heart Failure Questionnaire; CSEQ = Cardiac Self-Efficacy Questionnaire; DAQIHF = Daily Activity Questionnaire in Heart Failure; EHFScBS = European Heart Failure Self-care Behavior Scale; HF = heart failure; HFNAQ = Heart Failure Needs Assessment Questionnaire; HFSPS = Heart Failure Somatic Perception Scale; KCCQ = Kansas City Cardiomyopathy Questionnaire; LVD-36 = Left Ventricular Dysfunction Questionnaire; MacNew = MacNew Heart Disease HRQL Questionnaire; MLHFQ = Minnesota Living with Heart Failure Questionnaire; PRO = patient-reported outcome; QLI = Quality of Life Index-Cardiac Version; QLQ-SHF = Quality of Life Questionnaire for Severe Heart Failure; SAQ = Seattle Angina Questionnaire; SCHFI = Self-Care Heart Failure Index; SDHF = San Diego Heart Failure Questionnaire; SOB-HF = Shortness of Breath in Heart Failure.

change in management as part of a clinical trial. Such a trial might establish another way these PRO measures fit into the range of clinical practice and demonstrate the clinical utility of a low-cost intervention.

The present evaluation complements other assessments of the chronic HF PRO literature. A structured evaluation of HF PROs for the measurement of population-level health status identified the KCCQ and MLHFQ as the strongest HF-specific PRO instruments (42). A systematic review and meta-analysis of HF PRO measures between 1996 and 2006 included only the CHQ, KCCQ, LVD-36, MLHFQ, and QLQ-SHF, and did not evaluate suitability for regulatory approval (43). In that analysis the MLHFQ was the most validated option and the KCCQ had yet to be extensively investigated. Meta-analyses demonstrated that interventions with small expected effects such as exercise programs produced small score changes and interventions expected to produce large effects such as medications produced large score changes. An updated qualitative appraisal of the HF PRO literature was limited to only the CHAT, CHP, CHQ, KCCQ, LVD-36, MLHFQ, and QLQ-SHF (22). Using a standardized tool for evaluating PRO instruments the most highly rated were the KCCQ followed by the MLHFQ and CHQ. The CHQ was felt to be conceptually superior, LVD-36 the most reliable, and the KCCQ the most widely validated and interpretable with low administrator burden. Finally, a recent systematic assessment of HF-specific PRO instruments for implementation into clinical care concluded that the KCCQ and MLHFQ were the best suited because of their favorable psychometric properties, ease-of-use, and prognostic value (16).

**STUDY LIMITATIONS.** Although this was an extensive review, it may have failed to identify evidence in alternative databases or languages, PRO instruments published prior to the selected time frame, or attributes of PRO measures not published in manuscripts or conference abstracts. Nevertheless, given the multiple organized search strategies it is unlikely that well-described HF-specific PRO tools were missed. We purposefully did not include generic PRO instruments due to their decreased sensitivity to clinical change. Due to space limitation, some PRO measures with extensive publication histories such as the KCCQ and MLHFQ were not comprehensively explicated here. Aspects of these PRO instruments like degree of translation into non-English languages were not systematically assessed, however many have been successfully translated and validated in multiple language formats, including the MLHFQ, KCCQ, and MacNew (43–45). Finally, other similar searches in multiple languages did not identify additional PRO measures (43).

## CONCLUSIONS

No disease-specific PRO instrument utilized with chronic HF patients appears to currently satisfy all

the FDA guidance, and none of those identified may merit an approved label claim. The FDA criteria and the findings here should guide development of future PRO measures that would facilitate comparison of results across clinical trials and practices, help evaluate treatment benefits, stratify patients, and support product label claims. Additionally, the clinical utility of PRO measurement as a management tool should be assessed in a clinical trial.

**ACKNOWLEDGMENT** The authors thank Julia R. Gage, PhD (on behalf of Amgen Inc.), for her assistance in preparation of this manuscript.

**REPRINT REQUESTS AND CORRESPONDENCE:** Dr. John R. Teerlink, San Francisco VA Medical Center, Cardiology, 111C; Building 203, Room 2A-49, 4150 Clement Street, San Francisco, California 94121-1545. E-mail: john.teerlink@ucsf.edu.

## PERSPECTIVES

**COMPETENCY IN MEDICAL KNOWLEDGE:** Standardized PRO tools allow quantification of the chronic HF patient experience and multiple validated options exist for this purpose. Choice of specific PRO measure depends on the desired use.

**TRANSLATIONAL OUTLOOK 1:** Future studies are required to develop a PRO tool for use in chronic HF acceptable to the FDA and will require focus on the details of content validity and recall periods.

**TRANSLATIONAL OUTLOOK 2:** To further develop the field, a clinical trial using a PRO measure as the intervention is suggested to demonstrate utility as a therapeutic intervention.

## REFERENCES

**1.** Rumsfeld JS, Alexander KP, Goff DC Jr., et al. Cardiovascular health: the importance of measuring patient-reported health status: a scientific statement from the American Heart Association. Circulation 2013;127:2233–49.

**2.** Bonow RO, Ganiats TG, Beam CT, et al. ACCF/AHA/AMA-PCPI 2011 performance measures for adults with heart failure: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Performance Measures and the American Medical Association-Physician Consortium for Performance Improvement. J Am Coll Cardiol 2012;59:1812–32.

**3.** Zannad F, Garcia AA, Anker SD, et al. Clinical outcome endpoints in heart failure trials: a European Society of Cardiology Heart Failure Association consensus document. Eur J Heart Fail 2013;15:1082–94.

**4.** Lindenfeld J, Albert NM, Boehmer JP, et al. HFSA 2010 Comprehensive Heart Failure Practice Guideline. J Card Fail 2010;16:e1–194.

**5.** Frank L, Basch E, Selby JV, Patient-Centered Outcomes Research I. The PCORI perspective on patient-centered outcomes research. JAMA 2014;312:1513–4.

**6.** European Medicines Agency. Reflection Paper on the Regulatory Guidance for the Use of Health-Related Quality of Life (HRQL) Measures in the Evaluation of Medicinal Products. 2005. Available at: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003637.pdf. Accessed June 2, 2016.

**7.** Rothman M, Burke L, Erickson P, Leidy NK, Patrick DL, Petrie CD. Use of existing patient-reported outcome (PRO) instruments and their modification: the ISPOR Good Research Practices for Evaluating and Documenting Content Validity for the Use of Existing Instruments and Their Modification PRO Task Force Report. Value Health 2009;12:1075–83.

**8.** U.S. Department of Health and Human Services. Guidance for Industry: Patient Reported Outcome Measures: Use in Medical Product Development to Support Label Claims. 2009. Available at: http://www.fda.gov/downloads/drugs/guidances/ucm193282.pdf. Accessed June 2, 2016.

**9.** Berry C, McMurray J. A review of quality-of-life evaluations in patients with congestive heart failure. Pharmacoeconomics 1999;16:247–71.

**10.** Mommersteeg PM, Denollet J, Spertus JA, Pedersen SS. Health status as a risk factor in cardiovascular disease: a systematic review of current evidence. Am Heart J 2009;157:208–18.

**11.** Wiebe S, Guyatt G, Weaver B, Matijevic S, Sidwell C. Comparative responsiveness of generic and specific quality-of-life instruments. J Clin Epidemiol 2003;56:52–60.

**12.** Bennett SJ, Oldridge NB, Eckert GJ, et al. Discriminant properties of commonly used quality of life measures in heart failure. Qual Life Res 2002;11:349–59.

**13.** Scoggins JF, Patrick DL. The use of patient-reported outcomes instruments in registered clinical trials: evidence from ClinicalTrials.gov. Contemp Clin Trials 2009;30:289–92.

**14.** Sanders C, Egger M, Donovan J, Tallon D, Frankel S. Reporting on quality of life in randomised controlled trials: bibliographic study. BMJ 1998;317:1191–4.

**15.** Rahimi K, Malhotra A, Banning AP, Jenkinson C. Outcome selection and role of patient reported outcomes in contemporary cardiovascular trials: systematic review. BMJ 2010;341:c5707.

**16.** Kelkar AA, Spertus J, Pang P, et al. Utility of patient-reported outcome instruments in heart failure. J Am Coll Cardiol HF 2016;4:165–75.

**17.** Joseph SM, Novak E, Arnold SV, et al. Comparable performance of the Kansas City Cardiomyopathy Questionnaire in patients with heart failure with preserved and reduced ejection fraction. Circ Heart Fail 2013;6:1139–46.

**18.** Sullivan MD, Levy WC, Russo JE, Crane B, Spertus JA. Summary health status measures in advanced heart failure: relationship to clinical variables and outcome. J Card Fail 2007;13:560–8.

**19.** Jaarsma T, Stromberg A, Martensson J, Dracup K. Development and testing of the European Heart Failure Self-Care Behaviour Scale. Eur J Heart Fail 2003;5:363–70.

**20.** Ekman I, Granger B, Swedberg K, Stenlund H, Boman K. Measuring shortness of breath in heart failure (SOB-HF): development and validation of a new dyspnoea assessment tool. Eur J Heart Fail 2011;13:838–45.

**21.** Bennett SJ, Oldridge NB, Eckert GJ, et al. Comparison of quality of life measures in heart failure. Nurs Res 2003;52:207–16.

**22.** Garin O, Herdman M, Vilagut G, et al. Assessing health-related quality of life in patients with heart failure: a systematic, standardized comparison of available measures. Heart Fail Rev 2014;19:359–67.

**23.** Heidenreich PA, Spertus JA, Jones PG, et al. Health status identifies heart failure outpatients at risk for hospitalization or death. J Am Coll Cardiol 2006;47:752–6.

**24.** Konstam V, Salem D, Pouleur H, et al. Baseline quality of life as a predictor of mortality and hospitalization in 5,025 patients with congestive heart failure. SOLVD Investigations. Studies of Left Ventricular Dysfunction Investigators. Am J Cardiol 1996;78:890–5.

**25.** Eurich DT, Johnson JA, Reid KJ, Spertus JA. Assessing responsiveness of generic and specific health related quality of life measures in heart failure. Health Qual Life Outcomes 2006;4:89.

**26.** Wyrwich KW, Wolinsky FD. Identifying meaningful intra-individual change standards for health-related quality of life measures. J Eval Clin Pract 2000;6:39–49.

**27.** DeMuro C, Clark M, Mordin M, Fehnel S, Copley-Merriman C, Gnanasakthy A. Reasons for rejection of patient-reported outcome label claims: a compilation based on a review of patient-reported outcome use among new molecular entities and biologic license applications, 2006–2010. Value Health 2012;15:443–8.

**28.** DeMuro C, Clark M, Doward L, Evans E, Mordin M, Gnanasakthy A. Assessment of PRO label claims granted by the FDA as compared to the EMA (2006–2010). Value Health 2013;16:1150–5.

**29.** Gwaltney CJ, Slagle AF, Martin M, Ariely R, Brede Y. Hearing the voice of the heart failure patient: key experiences identified in qualitative interviews. Br J Cardiol 2012;19:e1–7.

**30.** Norman GR, Stratford P, Regehr G. Methodological problems in the retrospective computation of responsiveness to change: the lesson of Cronbach. J Clin Epidemiol 1997;50:869–79.

**31.** Arnold SV, Kosiborod M, Li Y, et al. Comparison of the Seattle Angina Questionnaire With Daily Angina Diary in the TERISA Clinical Trial. Circ Cardiovasc Qual Outcomes 2014;7:844–50.

**32.** Lupon J, Gastelurrutia P, de Antonio M, et al. Quality of life monitoring in ambulatory heart failure patients: temporal changes and prognostic value. Eur J Heart Fail 2013;15:103–9.

**33.** Alla F, Briancon S, Guillemin F, et al. Self-rating of quality of life provides additional prognostic information in heart failure. Insights into the EPICAL study. Eur J Heart Fail 2002;4:337–43.

**34.** Chan PS, Soto G, Jones PG, et al. Patient health status and costs in heart failure: insights from the eplerenone post-acute myocardial infarction heart failure efficacy and survival study (EPHESUS). Circulation 2009;119:398–407.

**35.** Ouwerkerk W, Voors AA, Zwinderman AH. Factors influencing the predictive power of models for predicting mortality and/or heart failure hospitalization in patients with heart failure. J Am Coll Cardiol 2014;2:429–36.

**36.** Kosiborod M, Soto GE, Jones PG, et al. Identifying heart failure patients at high risk for near-term cardiovascular events with serial health status assessments. Circulation 2007;115:1975–81.

**37.** Cohn J, Cleland JG, Lubsen J, et al. Unconventional end points in cardiovascular clinical trials: should we be moving away from morbidity and mortality? J Card Fail 2009;15:199–205.

**38.** Riegel B, Carlson B, Moser DK, Sebern M, Hicks FD, Roland V. Psychometric testing of the self-care of heart failure index. J Card Fail 2004;10:350–60.

**39.** Ekman I, Kjork E, Andersson B. Self-assessed symptoms in chronic heart failure–important information for clinical management. Eur J Heart Fail 2007;9:424–8.

**40.** Raphael C, Briscoe C, Davies J, et al. Limitations of the New York Heart Association functional classification system and self-reported walking distances in chronic heart failure. Heart 2007;93:476–82.

**41.** Flynn KE, Lin L, Moe GW, et al. Relationships between changes in patient-reported health status and functional capacity in outpatients with heart failure. Am Heart J 2012;163:88–94.e3.

**42.** Mackintosh A, Gibbons E, Fitzpatrick R. A Structured Review of Patient-Reported Outcome Measures (PROMs) for Heart Failure. Oxford, England: University of Oxford, 2009.

**43.** Garin O, Ferrer M, Pont A, et al. Disease-specific health-related quality of life questionnaires for heart failure: a systematic review with meta-analyses. Qual Life Res 2009;18:71–85.

**44.** 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;21:90–7.

**45.** Hofer 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.

**KEY WORDS** chronic heart failure, clinical management, European Medicines Agency, Food and Drug Administration, patient-reported outcomes, regulatory approval

**APPENDIX** For expanded Methods and Reference sections, please see the online version of this article.