


The development and initial validation of the PROMIS®+HF-27 and PROMIS+HF-10 profiles

Faraz S. Ahmad^{1,2*} , Kathryn L. Jackson³, Susan E. Yount³, Nan E. Rothrock³, Michael A. Kallen³, Leilani Lacson³, Karl Y. Bilimoria⁴, Abel N. Kho², Raja Kannan Mutharasan¹, Peter A. McCullough⁵, Jeffrey Bruckel⁶, Savitri Fedson^{7,8}, Stephen E. Kimmel⁹, David T. Eton^{10,11}, Kathleen L. Grady^{1,12}, Clyde W. Yancy¹ and David Cella^{3,13}

¹Division of Cardiology, Department of Medicine, Northwestern University Feinberg School of Medicine, 676 North Saint Clair Street, Suite 600, Chicago, IL 60611, USA; ²The Center for Health Information Partnerships (CHIP), Institute of Public Health & Medicine, Northwestern University Feinberg School of Medicine, Chicago, IL, USA;

³Department of Medical Social Sciences, Northwestern University Feinberg School of Medicine, Chicago, IL, USA; ⁴Surgical Outcomes and Quality Improvement Center (SOQIC), Department of Surgery and Center for Healthcare Studies, Northwestern University Feinberg School of Medicine, Chicago, IL, USA; ⁵Truth for Health Foundation, Tucson, AZ; ⁶Division of Cardiology, University of Rochester Medical Center, Rochester, NY, USA; ⁷Section of Cardiology, Michael E DeBakey Veterans Administration Medical Center, Houston, TX, USA; ⁸Center for Medical Ethics and Health Policy, Baylor College of Medicine, Houston, TX, USA; ⁹Department of Epidemiology, University of Florida College of Public Health and Health Professions and College of Medicine, Gainesville, FL, USA; ¹⁰Robert D. and Patricia E. Kern Center for the Science of Health Care Delivery, Mayo Clinic, Rochester, MN, USA; ¹¹Division of Health Care Delivery Research, Mayo Clinic, Rochester, MN, USA; ¹²Division of Cardiac Surgery, Department of Surgery, Northwestern University Feinberg School of Medicine, Chicago, IL, USA; and ¹³Center for Patient Centered Outcomes, Institute of Public Health & Medicine, Northwestern University Feinberg School of Medicine, Chicago, IL, USA

Abstract

Aims Heart failure (HF) is a common and morbid condition impacting multiple health domains. We previously reported the development of the PROMIS®-Plus-HF (PROMIS+HF) profile measure, including universal and HF-specific items. To facilitate use, we developed shorter, PROMIS+HF profiles intended for research and clinical use.

Methods and results Candidate items were selected based on psychometric properties and symptom range coverage. HF clinicians ($n = 43$) rated item importance and clinical actionability. Based on these results, we developed the PROMIS+HF-27 and PROMIS+HF-10 profiles with summary scores (0–100) for overall, physical, mental, and social health. In a cross-sectional sample ($n = 600$), we measured internal consistency reliability (Cronbach's alpha and Spearman–Brown), test–retest reliability (intraclass coefficient; $n = 100$), known-groups validity via New York Heart Association (NYHA) class, and convergent validity with Kansas City Cardiomyopathy Questionnaire (KCCQ) scores. In a longitudinal sample ($n = 75$), we evaluated responsiveness of baseline/follow-up scores by calculating mean differences and Cohen's d and comparing with paired t -tests. Internal consistency was good to excellent (α 0.82–0.94) for all PROMIS+HF-27 scores and acceptable to good (α /Spearman–Brown 0.60–0.85) for PROMIS+HF-10 scores. Test–retest intraclass coefficients were acceptable to excellent (0.75–0.97). Both profiles demonstrated known-groups validity for the overall and physical health summary scores based on NYHA class, and convergent validity for nearly all scores compared with KCCQ scores. In the longitudinal sample, we demonstrated responsiveness for PROMIS+HF-27 and PROMIS+HF-10 overall and physical summary scores. For the PROMIS+HF overall summary scores, a group-based increase of 7.6–8.3 points represented a small to medium change (Cohen's $d = 0.40$ –0.42). For the PROMIS+HF physical summary scores, a group-based increase of 5.0–5.9 points represented a small to medium change (Cohen's $d = 0.29$ –0.35).

Conclusions The PROMIS+HF-27 and PROMIS+HF-10 profiles demonstrated good psychometric characteristics with evidence of responsiveness for overall and physical health. These new measures can facilitate patient-centred research and clinical care, such as improving care quality through symptom monitoring, facilitating shared decision-making, evaluating quality of care, assessing new interventions, and monitoring during the initiation and titration of guideline-directed medical therapy.

Keywords Heart failure; Outcomes research; Quality of life; Patient-reported outcomes; Health status

Received: 18 March 2022; Revised: 11 June 2022; Accepted: 27 June 2022

*Correspondence to: Faraz S. Ahmad, Division of Cardiology, Department of Medicine, Northwestern University Feinberg School of Medicine, 676 North Saint Clair Street, Suite 600, Chicago, IL 60611, USA. Tel: 312-695-4965; Fax: 312-926-8620. Email: faraz.ahmad@northwestern.edu

Introduction

Patient-reported outcome measures (PROMs) are used in clinical studies and in routine clinical care to capture health status directly from patients.^{1–5} Collection of data from PROMs may be particularly informative for patients with heart failure (HF), a prevalent, chronic condition that affects an estimated 64.3 million adults worldwide manifesting diverse symptoms and having broad impact on health.⁶ Beyond experiencing symptoms of shortness of breath, fatigue, exercise intolerance, and oedema, patients with HF experience a wide range of physical, mental, and social health effects both from HF itself and from common, non-cardiovascular comorbid conditions, such as chronic obstructive pulmonary disease, diabetes, depression, and anxiety.^{7–10}

Applications of PROMs in HF include assessing new interventions, improving clinical care and shared decision-making, evaluating quality of care, and monitoring during the initiation and titration of guideline-directed medical therapy. Although PROMs are being included in HF clinical trials to assess outcomes, their use in clinical care has been limited, in part due to electronic health record (EHR), health system, clinician, and patient-level barriers.^{8,11,12} The rapid adoption of telehealth visits during the COVID-19 pandemic as well as studies illustrating the association of lower patient-reported health status with mortality in the general population and changes in pulmonary artery pressures and risk of HF hospitalization in patients with HF have further highlighted the potential utility of PROMs for remote monitoring and enhancing tele-visits.^{13–16} The 2020 American College of Cardiology/American Heart Association report on HF performance measures included two quality measures based on PROMs: the first on measuring health status and the second on avoidance of worsening health status.¹⁷ Thus, capturing multiple domains of health for patients with HF in a single, brief instrument may increase usage of PROMs in clinical care.

The Patient-Reported Outcomes Measurement Information System (PROMIS®) is a set of largely person-centred measures that can be used in the general population and in those with chronic conditions. Major strengths of PROMIS include extensive prior testing, the ability to compare PROMIS T-scores from PROMIS domains (such as fatigue, dyspnoea, and depression) across different populations, and integration into research studies and in multiple clinical settings via an application programming interface (API) operated by HealthMeasures, a non-profit organization that manages PROMIS and other National Institutes of Health-funded measurement systems.

We previously reported the development and initial validation of the PROMIS-Plus-HF (PROMIS+HF) profile measure, an HF-specific measure that combines existing, universal PROMIS items with new HF-specific items.¹⁸ The measure includes items identified as important to HF by both patients and clinicians from focus groups and interviews. It exhibited

good psychometric characteristics and select evidence of responsiveness. The PROMIS+HF profile measure comprises 86 items across 18 domains of physical, mental, and social health. Clinicians or researchers are free to select those items or domains that are most relevant to a specific context. However, given that shorter measures are often more practical to end users by reducing respondent burden, to further encourage uptake of the PROMIS+HF, we sought to develop two, briefer profile measures: one tailored for research (PROMIS+HF-27) and the other for clinical use (PROMIS+HF-10).

Methods

Data sets for PROMIS+HF-27 and PROMIS+HF-10 development and validation

The development and initial validation of the long-form PROMIS+HF profile measure was previously reported.¹⁸ The 86-item PROMIS+HF profile includes 64 existing PROMIS items and 22 new items to fill content-coverage gaps identified in patient and clinician focus groups and interviews. Items have existing or new calibration estimates derived from item response theory.

This study utilized two archival, de-identified data sets from the development and initial validation studies for the PROMIS+HF profile measure. All participants in both data sets had been administered the PROMIS+HF profile measure and comparison measures described in the succeeding text.

A cross-sectional sample of 600 participants with self-reported HF was identified and recruited by the online panel company Opinions4Good (Op4G; Portsmouth, NH). The sample included quotas for age, gender, race, ethnicity, and self-reported functional status modelled on the New York Heart Association (NYHA) classes. For reproducibility testing, a subsample of 100 respondents was retested 3–7 days following baseline survey administration.

A longitudinal sample was recruited from six US health systems. Patients with HF were enrolled if they fit into one of the following five treatment categories that would likely show improvement in HF-related symptoms and health status between baseline and 3 month follow-up: (i) initiation of guideline-directed medical therapy after a new HF diagnosis or first hospitalization; (ii) cardiac rehabilitation for chronic stable HF; (iii) initiation of cardiac resynchronization therapy; (iv) implantation of a left ventricular assist device; and (v) recent discharge after hospitalization primarily for HF. Participants completed the instruments via paper forms or a tablet. The final sample included 185 participants with data at baseline and a subset of 75 participants with follow-up data at 3 months. The sample criteria are displayed in Supporting Information, *Table S1*, sample characteristics are summarized in *Table 1*, and a flowsheet

Table 1 Overview of cross-sectional and longitudinal samples

	Cross-sectional sample (N = 600)	Longitudinal sample (N = 75)
Age (years), mean (SD)	54 (14)	58 (12)
Sex, N (%)		
Female	270 (45)	35 (47)
Male	330 (55)	40 (53)
Race, N (%)		
American Indian or Alaska Native	17 (3)	1 (1)
Asian	39 (7)	1 (1)
Black or African American	115 (19)	32 (43)
Native Hawaiian or Pacific Islander	1 (<1)	0 (0)
White	401 (67)	38 (51)
Some other race	20 (3)	0 (0)
More than one race	7 (1)	0 (0)
Unknown or not reported	0 (0)	3 (4)
Ethnicity, N (%)		
Hispanic or Latino	171 (29)	2 (3)
Not Hispanic or Latino	429 (72)	73 (97)
Region of enrolment, N (%)		
Midwest	—	14 (19)
North-east	—	8 (11)
Pacific	—	15 (20)
South	—	38 (51)
Education level, N (%)		
Did not complete high school	4 (0.7)	9 (12)
High school diploma or equivalent	73 (12)	23 (31)
Some college	154 (26)	21 (28)
Graduated college or higher	369 (61.5)	22 (29)
Categories, N (%)		
New diagnosis or first hospitalization	—	11 (15)
Cardiac rehabilitation	—	5 (7)
Cardiac resynchronization therapy	—	5 (7)
Left ventricular assist device	—	11 (15)
Hospitalization primarily for HF	—	43 (57)
Clinical characteristics, N (%)		
Diabetes	80 (13)	28 (37)
Chronic obstructive lung disease	40 (7)	15 (20)
Depression	59 (10)	13 (17)
Chronic kidney disease	16 (3)	27 (36)
NYHA class ^a		
Class I	60 (10)	4 (5)
Class II	248 (41)	14 (19)
Class III	245 (41)	27 (36)
Class IV	47 (8)	6 (8)
Not available	—	24 (32)
Left ventricular ejection fraction <50%	—	54 (72)

HF, heart failure; NYHA, New York Heart Association; SD, standard deviation.

Adapted from Ahmad *et al.*, with permissions from Wolters Kluwer Health, Inc. Copyright 2019, American Heart Association.

^aNYHA class was measured via self-report in the cross-sectional sample and via chart review in the longitudinal sample.

description of reasons for participant loss at time of follow-up in Supporting Information, *Figure S1*.

Comparison measures

In addition to the PROMIS+HF profile measure, we used three comparator measures. NYHA class was measured via self-report in the cross-sectional, 600-person sample and chart review in the longitudinal sample. The PROMIS Global Health Scale and the Kansas City Cardiomyopathy Questionnaire (KCCQ) were administered in both the cross-sectional and longitudinal samples.^{19–22} The PROMIS Global Physical Health and Mental Health scores and the KCCQ overall summary,

clinical summary, quality of life, and social limitation scores were used as comparators. The KCCQ is a responsive measure with a threshold of 5 points suggested as a meaningful, clinically important difference for the overall summary and clinical summary scores.²³ Additional details on all measures are included in Supporting Information, Methods.

PROMIS+HF-27 profile and PROMIS+HF-10 profile development

Initial item selection

A group from the research team ($n = 4$) with expertise in measurement science, psychometrics, and HF identified

candidate items from the long-form PROMIS+HF measure for the two, abbreviated profile measures. Items from given domains (e.g. fatigue and depression) were reviewed together. The team prioritized selection of PROMIS items based on prior psychometric testing by PROMIS investigators, wide use in other research studies across multiple settings and populations, and coverage a range of content and intensity/difficulty that may have relevance to patients with HF. For example, all items from the Health Behavior Outcomes domain were excluded because these items measure behaviours and not health effects. All items from the domains of anger, cognitive abilities, satisfaction with social roles and activities, and social isolation were excluded because of overlap of content with other domains and/or lack of wider use in other research studies. A total of 31 candidate items from 13 domains were selected for further evaluation for inclusion in the two profile measures (Supporting Information, *Table S2*).

Heart failure clinician survey

Because the intended users of the abbreviated PROMIS+HF profile were clinicians, we designed a web-based survey for clinicians (physicians, nurse practitioners, registered nurses, physician assistants, and psychologists) who take care of patients with HF for input on the 31 candidate items. In the survey, clinicians rated each item on clinical actionability ('Not at all', 'A little bit', 'Quite a bit', and 'Very much'). From a group of items from one to three related domains, they also selected 'the most important questions to ask patients with heart failure'.

From December 2019 to April 2020, we contacted a convenience sample of 50 HF clinicians across the US via email and asked that they complete the survey and forward the invitation on to other colleagues who take care of patients with HF. A total of 43 clinicians, which include the primary recipients of the email survey and colleagues to whom the invitation was forwarded, completed the survey. Of the 43 participants, there were 23 physicians, 17 registered nurses/nurse practitioners/physician assistants, and 3 psychologists. Fifty-eight per cent were female, and the group had an average experience of 9.1 years (SD ± 7.1) taking care of patients with HF. The results were summarized using standard descriptive statistics (counts, mean, and median) and reviewed by the research team (Supporting Information, *Table S3*). The team removed four items based on survey responses (i.e. rated relatively low by respondents) and to reduce respondent burden and finalized the PROMIS+HF-27 profile. From the PROMIS+HF-27 items, a group from the research team selected items for the PROMIS+HF-10 profile that were the most highly rated on clinical actionability and importance by the survey respondents while maximizing coverage of the range of content. *Figure 1* and Supporting Information, *Table S4* summarize the PROMIS+HF-27 and PROMIS+HF-10

profiles. The full item content is available in Supporting Information, *Tables S5* and *S6*.

As data were de-identified and the clinician survey was part of measure development, the development and initial validation of the PROMIS+HF-27 and PROMIS+HF-10 profiles received a Not Human Subject Research Determination from the study team's Institutional Review Board. The data that support the findings of this study are available from the corresponding author upon reasonable request.

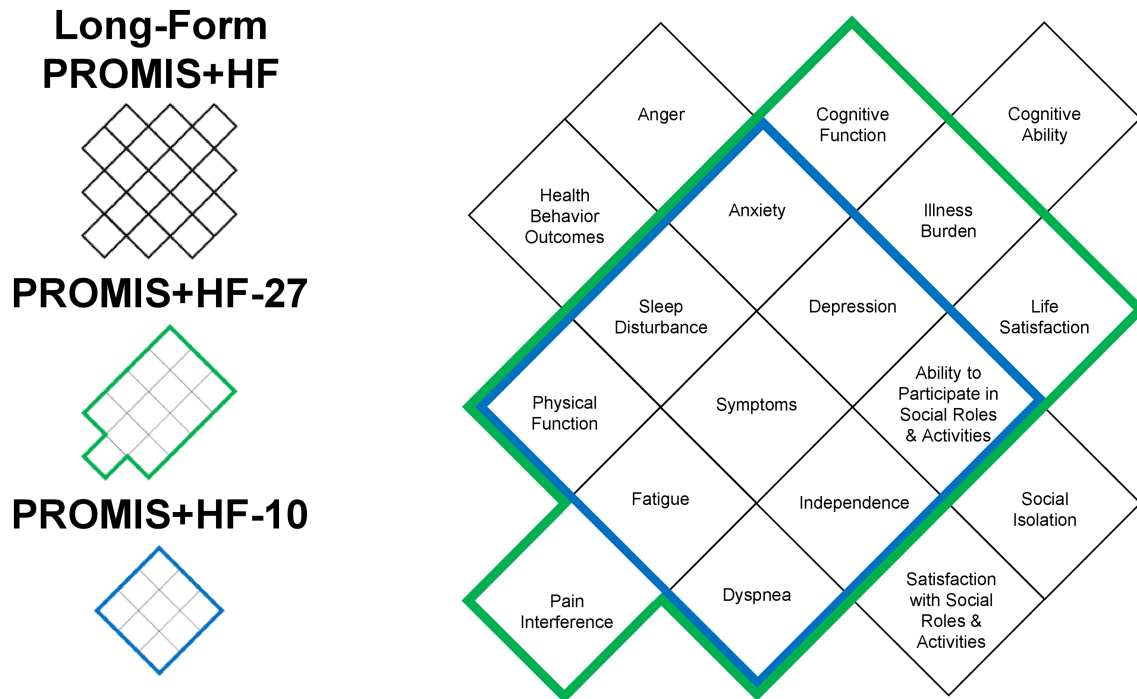
Scoring of the PROMIS+HF-27 and PROMIS+HF-10 profiles

For the PROMIS+HF-27 profile, we generated domain-level, raw average scores as well as summary scores for physical, mental, social, and overall health. Raw response scores (range 1–5) were averaged to create domain-level scores that were then summed and transformed to produce physical, mental, and social health summary scores (range 0–100). Higher score represents better health. The overall health summary score was calculated as an average of physical health summary score, mental health summary score, and social health summary score and weighted such that the physical health summary score receives twice the weight of the mental and social health summary scores. The physical health summary was weighted more heavily than mental and social scores because physical health items comprise more than half of all PROMIS+HF-27 items and to reflect the weight of emphasis on physical vs. other symptoms in HF clinical practice and research. Domains were grouped under physical, mental, and social health based on the existing conceptual health framework used across PROMIS measures.

In the PROMIS+HF-10, physical, mental, and social health summary scores were calculated by averaging the items in these domains. As with the PROMIS+HF-27, the overall health summary score was calculated as an average of the physical, mental, and social health scores, with the physical health score receiving twice the weight of the mental and social scores.

In addition to the 0–100 range summary health scores, for the PROMIS+HF-27 and PROMIS+HF-10, T-scores for existing PROMIS domains can be generated for the following domains: dyspnoea, fatigue, physical function, sleep disturbance, depression, anxiety, and ability to participate in social roles. Additionally, from the PROMIS+HF-27, PROMIS T-scores can be generated for pain interference and cognitive function. The precision of T-scores increases with more items per domain; thus, the standard error for T-scores generated using the PROMIS+HF-10 is larger than with the PROMIS+HF-27.

Figure 1 Comparison of domains covered in each of the PROMIS+HF profiles. This figure compares the domains covered in the long-form PROMIS+HF, PROMIS+HF-27, and PROMIS+HF-10 profiles. All domains in the PROMIS+HF-10 (in blue) are also in the PROMIS+HF-27 and long-form PROMIS+HF profiles. All domains in the PROMIS+HF-27 (in green) are on the long-form PROMIS+HF profile. The long-form PROMIS+HF profile has five domains (in black) not included in the abbreviated profiles. The number of items in each domain varies by profile as detailed in the Supporting Information.



Additional details on scoring and handling missing data are available in the scoring guide included in the Supporting Information.

Statistical and psychometric analysis

Initial psychometric assessment and reliability testing

Using the cross-sectional sample, we generated item-level frequencies for each item to evaluate skewness and floor and ceiling effects. We summarized results by physical, mental, social, and overall health summary scores using standard descriptive statistics (mean, median, min, max, kurtosis, and graphic score distributions) and classical item analyses (inter-item correlation and item-total correlation).²⁴ Internal consistency reliability of the PROMIS+HF-27 and PROMIS+HF-10 physical, mental, social, and overall health summary scores was assessed using Cronbach's alpha for scales with greater than two items and Spearman-Brown coefficients for scales with two items.^{25,26} We calculated intraclass correlation coefficients (ICCs) to evaluate test-retest reliability in the subset of 100 participants retested at 3–7 days.

Validity testing

We evaluated known-groups validity for the PROMIS+HF-27 and PROMIS+HF-10 using clinical groupings of PROMIS Global Physical Health and Mental Health scores and self-reported NYHA class symptoms. Participants' overall and physical health summary scores were calculated and grouped into 'low', 'average', or 'high' categories based on PROMIS Global Physical Health tertiles. Participants' overall, mental, and social health summary scores were calculated and grouped into 'low', 'average', or 'high' categories based on PROMIS Global Mental Health tertiles. We evaluated group differences with analysis of variance (ANOVA). We used ANOVA to examine difference in overall and physical health summary scores by the NYHA classification (I through IV) in the cross-sectional sample.

Convergent validity was evaluated by comparing health summary scores of the PROMIS+HF-27 and PROMIS+HF-10 to KCCQ. We calculated Pearson's r and Spearman's ρ (to account for skewed score distributions) correlation coefficients comparing the physical, mental, social, and overall health summary scores of the PROMIS+HF-27 and PROMIS+HF-10 profiles with the KCCQ clinical summary score, quality of life score, social limitation score, and overall summary score, respectively. Convergent validity was defined as r or

$\rho > 0.60$.^{27,28} We repeated the internal consistency reliability and validity testing in the baseline population of participants in the longitudinal clinical sample.

Responsiveness and estimates of group-based clinically important differences

To evaluate within-person change across time, we conducted paired *t*-tests of baseline vs. follow-up of PROMIS+HF-27 and PROMIS+HF-10 physical, mental, social, and overall health summary scores and calculated effect size (Cohen’s *d*). Several different methods have been described to establish a threshold for clinically important differences (CIDs) for PROMs, and CID thresholds can vary by use case, population, and context.^{29,30} To provide an estimate for group-based CID for the PROMIS+HF overall and physical summary scores, we used the mean difference between baseline and follow-up and Cohen’s *d* in the longitudinal sample with cut-offs for a small, medium, and large effect sizes of 0.2, 0.5, and 0.8, respectively.²⁷

Responsiveness to change was analysed using change in KCCQ scores as a clinical reference. Differences in KCCQ scores from baseline to follow-up were calculated and categorized as ‘improved’ (change of KCCQ ≥ 5), ‘no clinical change’ (change of KCCQ between -4 and 4), and ‘worsened’ (change of KCCQ ≤ -5) for KCCQ clinical summary and overall summary scores. Average change in scores across time for the PROMIS+HF-27 and PROMIS+HF-10 physical health summary and overall health summary scores by each KCCQ group (‘improved’, ‘no clinical change’, and ‘worsened’) were assessed using mixed-effects linear regression.

Comparison of scores from PROMIS+HF-27 and PROMIS+HF-10 profiles

We used Spearman’s ρ to calculate the correlation of the physical, mental, social, and overall health scores between the PROMIS+HF-27 and PROMIS+HF-10 profiles.

Results

Initial psychometric assessment and reliability testing

In the 600-person sample, the physical, mental, social, and overall health summary scores for the PROMIS+HF-27 and PROMIS+HF-10 were normally distributed with no evidence of skewness (Table 2). For internal consistency reliability testing (Table 3), Cronbach’s alphas were good (≥ 0.80) to excellent (≥ 0.90) for the overall health summary scores for the PROMIS+HF-27 (0.94) and PROMIS+HF-10 (0.85) and for the PROMIS+HF-27 physical, mental, and social health summary scores (0.90, 0.90, and 0.82, respectively). Given the small number of items, Cronbach’s alpha and Spearman–Brown coefficients were fair (≥ 0.60) to acceptable (≥ 0.70) for PROMIS+HF-10 physical health, mental health, and social health summary scores (0.76, 0.60, and 0.64, respectively). The average inter-item correlations across summary scores ranged from 0.37 to 0.50 in the PROMIS+HF-27 and from 0.34 to 0.44 for the PROMIS+HF-10, which indicates that the items are measuring the same construct without being overly repetitive. The item-total correlations (i.e. the correlation of an individual item with the summary score excluding the item of interest) were moderate to high (0.33–0.86) across both profiles.

The test–retest reliability (Supporting Information, Table S7) was excellent for nearly all PROMIS+HF-27 and PROMIS+HF-10 summary scores (ICC > 0.90) except for the PROMIS+HF-10 social health summary score, which was acceptable (0.75).

Validity testing

As shown in Supporting Information, Table S8, for the Global Physical Health known-groups validity comparisons, members

Table 2 Measure statistics of the PROMIS+HF-27 and PROMIS+HF-10 profiles in 600-participant cross-sectional sample

PROMIS+HF-27 profile								
Summary score	Mean	Median	SD	Skewness	Kurtosis	Observed		
						Min	Max	
Physical health	54.9	53.9	17.2	0.08	0.27	6.3	97.5	
Mental health	56.3	55.0	19.8	0.10	-0.03	0	100	
Social health	55.4	56.3	21.0	-0.12	0.13	0	100	
Overall health	55.4	54.2	16.9	0.04	0.27	3.5	98.8	
PROMIS+HF-10 profile								
Summary score	Mean	Median	SD	Skewness	Kurtosis	Observed		
						Min	Max	
Physical health	54.7	54.2	16.6	0.24	0.22	5	95.8	
Mental health	57.4	50.0	23.1	-0.13	-0.38	0	100	
Social health	57.3	62.5	21.8	-0.34	0.16	0	100	
Overall health	56.0	55.2	16.9	-0.01	0.25	7.3	97.9	

SD, standard deviation.

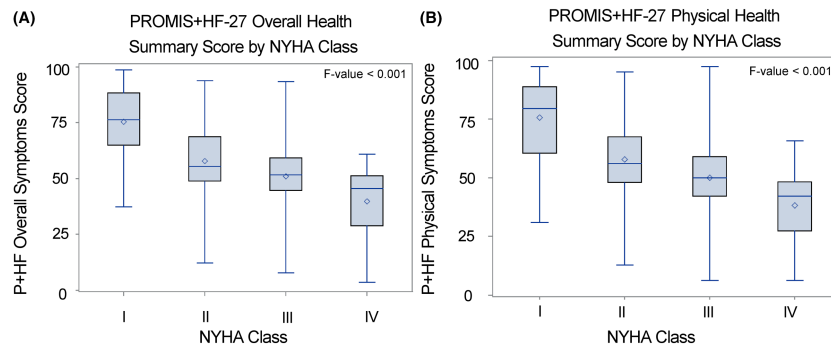
Table 3 Internal consistency reliability of the PROMIS+HF-27 and PROMIS+HF-10 profiles in 600-participant cross-sectional sample

PROMIS+HF-27 profile							
Summary score	No. of items	Alpha	Inter-item correlation			Item-total correlation	
			Average	Min	Max	Min	Max
Physical health	14	0.90	0.37	0.10	0.60	0.42	0.73
Mental health	9	0.90	0.49	0.27	0.76	0.62	0.79
Social health	4	0.82	0.50	0.43	0.55	0.75	0.83
Overall health	27	0.94	0.37	0.10	0.75	0.33	0.74
PROMIS+HF-10 profile							
Summary score	No. of items	Alpha	Inter-item correlation			Item-total correlation	
			Average	Min	Max	Min	Max
Physical health	6	0.76	0.34	0.14	0.50	0.47	0.71
Mental health	2	0.60 ^a	0.43	NA	NA	0.84	0.84
Social health	2	0.64 ^a	0.44	NA	NA	0.81	0.86
Overall health	10	0.85	0.35	0.14	0.50	0.41	0.73

Alpha, Cronbach's alpha.

^aSpearman-Brown coefficient due to two-item summary score.

Figure 2 Box plot for known-groups validity for PROMIS+HF-27 profile by New York Heart Association (NYHA) class. With each increase in self-reported NYHA class, there was a significant graded increase in mean PROMIS+HF-27 overall (A) and physical health (B) summary scores (F -value < 0.001). The solid line represents median score, and the circle represents mean score. The box represents the inter-quartile range of scores and whiskers the min and max values observed within each class. HF, heart failure; P+HF, PROMIS+HF; PROMIS, Patient-Reported Outcomes Measurement Information System.



of the high-tertile score group had significantly higher PROMIS+HF-27 and PROMIS+HF-10 overall and physical health summary scores than members of the middle-tertile and low-tertile groups. Members of the middle-tertile group had better overall and physical health summary scores than members of the low-tertile groups. Similarly, for the Global Mental Health known-groups validity comparison, members of the high-tertile score group had significantly better PROMIS+HF-27 and PROMIS+HF-10 overall, mental, and social health summary scores than members of the middle-tertile and low-tertile groups. Members of the middle-tertile groups had significantly better overall, mental, and social health summary scores than members of the low-tertile groups. All differences between groups were statistically significant at $P \leq 0.001$, with the exception of the difference in PROMIS+HF-10 social health summary score

between the highest and middle mental health groups ($P = 0.05$).

Across self-reported NYHA class, there were significantly lower (worse) overall and physical health summary scores (F -value < 0.001) reported on both PROMIS+HF-27 and PROMIS+HF-10 for persons reporting a higher (worse) NYHA grade (Figures 2 and 3).

For convergent validity testing, the PROMIS+HF-27 and PROMIS+HF-10 overall health and physical health summary scores had expected correlations with the corresponding scores of KCCQ (all $r/\rho > 0.60$). Some of the correlations of the mental health and social health summary scores of the two PROMIS profiles with the corresponding KCCQ scores (quality of life and social limitations, respectively) were slightly below the 0.60 threshold for convergence but were still relatively high (i.e. ≥ 0.50) (Table 4).

Figure 3 Box plot for known-groups validity for PROMIS+HF-10 profile by New York Heart Association (NYHA) class. With each increase in self-reported NYHA class, there was a significant graded increase in mean PROMIS+HF-10 overall (A) and physical health (B) summary scores (F -value < 0.001). The solid line represents median score, and the circle represents mean score. The box represents the inter-quartile range of scores and whiskers the min and max values observed within each class. HF, heart failure; P+HF, PROMIS+HF; PROMIS, Patient-Reported Outcomes Measurement Information System.

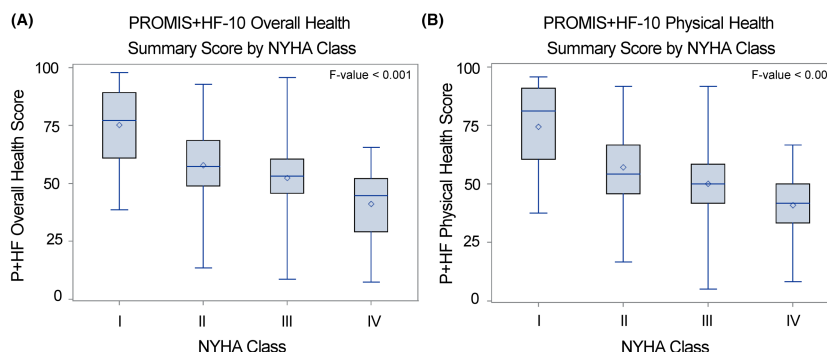


Table 4 Convergent validity of PROMIS+HF summary scores and KCCQ scores in 600-participant cross-sectional sample

PROMIS+HF-27 profile		
Measure score comparison	Pearson's r coefficient	Spearman ρ coefficient
P+HF physical health summary and KCCQ clinical summary	0.79	0.69
P+HF mental health summary and KCCQ quality of life	0.68	0.61
P+HF social health summary and KCCQ social limitation	0.64	0.57
P+HF overall health summary and KCCQ overall summary	0.82	0.72
PROMIS+HF-10 profile		
Measure score comparison	Pearson's r coefficient	Spearman ρ coefficient
P+HF physical health summary and KCCQ clinical summary	0.78	0.69
P+HF mental health summary and KCCQ quality of life	0.62	0.56
P+HF social health summary and KCCQ social limitation	0.57	0.50
P+HF overall health summary and KCCQ overall summary	0.80	0.70

KCCQ, Kansas City Cardiomyopathy Questionnaire; P+HF, PROMIS-Plus-Heart Failure.

Confirmatory reliability and validity testing

The findings of the confirmatory reliability and validity testing in the baseline population of participants on the longitudinal clinical sample ($n = 185$)—including known-groups validity testing with NYHA class abstracted from chart review—were overall similar to those in the 600-participant cross-sectional sample (data not shown).

Responsiveness and estimates of group-based clinically important differences

In the longitudinal sample of 75 participants from five treatment categories, there was evidence of responsiveness with a statistically significant improvement for the PROMIS+HF-27 overall, physical, and mental health summary scores and the PROMIS+HF-10 overall and physical health summary scores, with small to medium effect sizes (Table 5). The mean

change for the PROMIS+HF-27 social health and PROMIS+HF-10 mental and social health summary scores reflected improvement in health but did not achieve statistical significance. For the PROMIS+HF-27 and PROMIS+HF-10 overall summary scores, a group-based increase of 8.3 (Cohen's $d = 0.42$) and 7.6 points (Cohen's $d = 0.40$), respectively, represented a small to medium change. For the PROMIS+HF-27 and PROMIS+HF-10 physical summary scores, a group-based increase of 5.9 (Cohen's $d = 0.35$) and 5.0 points (Cohen's $d = 0.29$), respectively, represented a small to medium change.

More than half of the 75 participants ($n = 39$ – 42) had an improvement of ≥ 5 points in the KCCQ clinical summary and overall summary scores (Supporting Information, Tables S9 and S10). In the linear mixed regression analysis of change by KCCQ group, improvement in PROMIS+HF-27 physical and overall health summary scores and PROMIS+HF-10 physical summary scores was associated with a ≥ 5 -point improvement of KCCQ comparator score. Association of change in PROMIS

Table 5 Paired *t*-test of PROMIS+HF-27 and PROMIS+HF-10 summary scores and KCCQ scores at baseline and follow-up in longitudinal sample

	Time	N	Mean	SD	Mean difference	Effect size (Cohen's <i>d</i>)	<i>t</i> -test <i>P</i> value	Result interpretation ^a
PROMIS+HF-27 profile								
Physical health summary score	T1	75	45.4	20.9	8.3	0.42	0.0007	T2 = better domain status
	T2	75	53.7	23.6				
Mental health summary score	T1	75	60.1	23.1	4.6	0.25	0.0334	T2 = better domain status
	T2	75	64.7	22.7				
Social health summary score	T1	75	45.9	25.0	2.5	0.10	0.4005	No significant time difference
	T2	75	48.4	28.3				
Overall health summary score	T1	75	49.2	20.4	5.9	0.35	0.0036	T2 = better domain status
	T2	75	55.1	22.2				
PROMIS+HF-10 profile								
Physical health summary score	T1	75	47.0	20.9	7.6	0.40	0.0009	T2 = better domain status
	T2	75	54.6	22.0				
Mental health summary score	T1	75	60.7	23.3	4.0	0.18	0.1221	No significant time difference
	T2	75	64.7	24.4				
Social health summary score	T1	75	46.3	26.3	0.7	0.02	0.8422	No significant time difference
	T2	75	47.0	29.3				
Overall health summary score	T1	75	50.2	19.5	5.0	0.29	0.0159	T2 = better domain status
	T2	75	55.2	20.9				

SD, standard deviation.

^aInterpretation of results of statistical *t*-test.

+HF-10 overall health summary score with KCCQ overall summary score was directionally consistent with improvement but did not achieve statistical significance ($P = 0.06$). The sample sizes for participants without clinically important change or worsening KCCQ score of ≥ 5 points were small (n range 11–15), which limits the interpretability of the regression model in those groups.

Comparison of scores from PROMIS+HF-27 and PROMIS+HF-10 profiles

In the cross-sectional sample, the means for the PROMIS+HF-27 and PROMIS+HF-10 overall, physical, mental, and social health summary scores, respectively, were highly correlated (0.82–0.95; Supporting Information, *Table S11*).

Discussion

Using a combination of team expertise and input 43 HF clinicians, we developed two abbreviated profiles from the 86-item PROMIS+HF profile measure: the PROMIS+HF-27 and PROMIS+HF-10 profiles, each producing summary scores for overall health, physical health, mental health, and social health. The PROMIS+HF-27 is primarily intended for research purposes as it provides a broader coverage of health domains. The PROMIS+HF-10 is primarily for clinical purposes given its brevity. We performed initial validation of these new profiles using a cross-sectional sample of 600 participants with HF and a 75-person clinical, longitudinal sample. More than half of items in each profile comprise existing PROMIS items: 17 of 27 items in the PROMIS+HF-27 profile

and 6 of 10 items in the PROMIS+HF-10 profile were existing PROMIS items.

The PROMIS+HF-27 and PROMIS+HF-10 exhibited good to excellent internal consistency reliability and excellent test–retest reliability for each of the summary scores. Comparisons with PROMIS Global Physical and Mental Health measures, NYHA class, and KCCQ scores all supported the validity of the PROMIS+HF-27 and PROMIS+HF-10 profile summary scores. The confirmatory reliability and validity findings were, as anticipated, overall similar using the baseline data in the 185-person longitudinal cohort from six health systems. The PROMIS+HF-27 and PROMIS+HF-10 mental health and social health summary scores were modestly correlated with the KCCQ quality of life and social limitation scores, which may reflect the different aspects of mental and social health covered by the different measures or the difference in framing items with and without HF-specific language.

In the analysis of longitudinal data of 75 participants, there was some evidence of responsiveness, in particular for the PROMIS+HF-27 and PROMIS+HF-10 overall and physical health summary scores, both for the entire sample and when using the group with improvement of comparator KCCQ scores of ≥ 5 points as an anchor. The estimated range of group-based, small to medium clinical improvement for the PROMIS+HF overall and physical summary scores ranged from 5.0 to 8.3. PROMIS+HF mental and social health summary scores did not show consistent evidence of responsiveness, which may reflect limitations in the sample, the overall less responsive nature of mental and social health to specific clinical interventions in the 3 month period, and differences in content between the PROMIS+HF mental and social health items and KCCQ quality of life and social limitation items. Mean summary scores from the PROMIS+HF-27 and

PROMIS+HF-10 were highly correlated, which is expected given their overlap in content. The PROMIS+HF-10 was less precise than the PROMIS+HF-27, likely due to its relative brevity and the absence of items on low intensity physical activities. However, by covering key domains important to patients with HF in a very brief profile, the PROMIS+HF-10 scores can be the starting point for a more extensive conversation on patients' health status, goals of care, and care plan.

Items for the PROMIS+HF-10 were selected based on feedback from HF clinicians on the items they would find mostly clinically informative and important. If used in clinical care, the administration of the 10-item profile, including mental and social health items, on a regular interval may be helpful to track changes in HF patients over longer periods of time, such as every 6 months. The combination of physical, mental, and social health information, including measurement of loss of independence, may enhance the comprehensive management of patients with HF. This may include referral to palliative care specialists or mental health clinicians or implementing strategies to support caregivers. In other clinical scenarios, such as during the transitional period after HF hospitalization or for guideline-directed medical therapy titration in patients with HF, the PROMIS+HF-10 physical health items and summary score may be particularly informative.

Integration of the PROMIS+HF-10 into routine care may be facilitated by existing pathways to integrate PROMs into EHRs via APIs or native functionality. Logical Observation Identifiers Names and Codes (LOINC[®]) codes exist for PROMIS items and are in the review process for HF-specific items, which will facilitate mapping into common data models for networks like PCORnet and the Observational Health Data Sciences and Informatics.^{31,32}

Several PROMIS measures are currently being used in clinical practices in a wide range of institutions and settings, including primary care, orthopaedics, rheumatology, paediatrics, and oncology.^{33–37} Deployment of the PROMIS+HF-10 profile into clinical practice is a future area of investigation.

The PROMIS+HF-27 and PROMIS+HF-10 profiles extend the prior work of the long-form PROMIS+HF profile measure by generating overall, physical, mental health, and social health summary scores in concise instruments that cover key aspects of health in patients with HF. For a particular research case, additional items from the long-form PROMIS+HF may be useful. For example, if a researcher wanted to evaluate domains not included in the PROMIS+HF-27, such as anger or illness burden, or generate a more precise T-score for an existing PROMIS domain, such as fatigue, dyspnoea, or depression, additional items from the long-form PROMIS+HF profile could be administered.

These instruments add to the library of HF PROMs, each of which has strengths and limitations, depending on intended use case.¹⁰ The KCCQ and the Minnesota Living with Heart Failure Questionnaire have been used in clinical research, including regulatory review for use as medical device develop-

ment tools, and they are widely used in clinical studies.^{8,10}

Strengths of the PROMIS+HF-27 profile include the ability to comprehensively capture different components of mental and social health and leverage previously tested and widely used PROMIS items. Strengths of the PROMIS+HF-10 profile include concisely capturing key areas related to physical, mental, and social health and its modular form such that as few as six questions could be administered to produce the physical health summary score. Although, similar to other measures, the items do not capture all potential symptoms experienced by patients with HF, the abbreviated profiles capture key symptoms identified as most important by patients during the initial development of the measure and clinicians during the initial development of the long-form and abbreviated PROMIS+HF profiles. The PROMIS+HF-10 and the KCCQ-12 have several similarities, including comparable length and coverage of physical, mental, and social health. There are some differences in the specific content. For example, the PROMIS+HF-10 has questions to evaluate anxiety and independence, and the KCCQ-12 includes a more specific orthopnoea question and frames many of its questions with an HF attribution. Lastly, the PROMIS+HF-10 includes questions related to anxiety and independence and the ability to generate PROMIS T-scores on six PROMIS domains (dyspnoea, fatigue, physical function, sleep disturbance, depression, and ability to participate in social activities) that enable comparisons with the general population.

The PROMIS+HF-27 and PROMIS+HF-10 profiles add to the HF PRO assessment library and expand options for specific research and clinical use cases. For example, in a specific research or clinical use case, the lookback period for symptoms may be particularly important, and this period can vary by measure. The lookback period when specified is 7 days for PROMIS+HF items, whereas it is 2 weeks for KCCQ items and 24 h for items in the recently developed HF Symptom Tracker, which was designed for daily use.^{22,38} The researcher or clinician has the option to choose the measure that is best suited for the specific use case. In addition, health systems that have adopted a PROMIS approach to assessing its population may wish to remain in the PROMIS system when building out an HF-specific programme, affording some efficiencies and compatibility across service lines.

This study has several limitations. The PROMIS+HF-27 and PROMIS+HF-10 were developed using the same data set as the PROMIS+HF profile measure. Limitations of that data set include the relatively small number of patients in the clinical sample with longitudinal follow-up data and the lack of objective measures of exercise capacity and physical function to correlate with the PROMIS+HF physical health summary score. Although we demonstrated reliability, validity, and some evidence of responsiveness and report preliminary estimated of group-based CIDs, future work in prospective cohorts will provide additional data on CIDs, responsiveness, and the utility of the PROMIS+HF-27 and PROMIS+HF-10 pro-

files. The PROMIS+HF-10 profile has 10 items with two questions comprising mental and social health scores. The small number of questions limits psychometric testing of the PROMIS+HF-10 profile. Lastly, T-scores that enable comparisons to the general population or patients with HF were not developed for all domains as part of this initial study due to different calibrations for existing PROMIS items and HF-specific items. However, the PROMIS+HF-27 can generate T-scores anchored to the general population for the seven, pre-existing PROMIS domains with pre-existing PROMIS items (dyspnoea, fatigue, physical function, sleep disturbance, depression, cognitive function, and ability to participate in social roles and activities). Lastly, baseline characteristics, such as HF aetiology, HF subtype, and medication regimen, were not available in the study samples.

In summary, we developed and performed an initial validation of the PROMIS+HF-27 and PROMIS+HF-10 profiles, which produce overall health, physical health, mental health, and social health summary scores. Beyond additional research on the prospective validation, clinically important difference thresholds, and the prognostic value of the profiles and scores, several questions related to the broader implementation of PROMs in patients with HF remain. Although several studies have described barriers and facilitators for the integration of PROMs into routine care, more work remains to identify best practices to increase patient completion rates and clinician adoption, including streamlining EHR and workflow integration, and evaluating the impact of PROM administration on outcomes.^{39–44} Many questions remain about the frequency, instrument selection, and mode of PROM assessment in clinical practice, administration of PROMs during key aspects in HF care, such as post-hospitalization, medication titration, the use of PROMs as part of an evaluation for advanced therapies and/or palliative care, and the integration of PROMs with other data sources, including EHR data and implantable and non-invasive sensors. The PROMIS+HF-27 and PROMIS+HF-10 profiles can be used to facilitate patient-centred care and research, including the study of optimal PROM implementation in routine care.

Acknowledgements

We would like to acknowledge to the contributions of Elliott S. Fisher, Karen E. Schifferdecker, and Kathleen Carluzzo in

the initial development and testing of the PROMIS+HF profile measure.

Conflict of interest

None declared.

Funding

F.S.A. was supported by grants from the Agency for Healthcare Research and Quality (K12HS026385); National Institutes of Health, National Heart, Lung, and Blood Institute (K23HL155970); and the American Heart Association (AHA Number 856917).

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Table S1. Inclusion and exclusion criteria by sample.

Table S2. 31 candidate items for the abbreviated PROMIS+HF profiles.

Table S3. Clinician Survey Results by Domain.

Table S4. Overview of domains and items comprising summary scores for the long-form PROMIS+HF, PROMIS+HF-27, and PROMIS+HF-10 Profiles.

Table S5. The PROMIS+HF-27 Profile

Table S6. The PROMIS+HF-10 Profile.

Table S7. Test-retest reliability of PROMIS+HF-27 and PROMIS+HF-10 Profiles in 100-participant subsample of cross-sectional data.

Table S8. Known Groups Validity testing of PROMIS+HF summary scores by PROMIS Global Physical Health (Low vs Average vs High) using cross-sectional data.

Table S9. Responsiveness of PROMIS+HF-27 Profile with KCCQ scores as an anchor.

Table S10. Responsiveness of PROMIS+HF-10 Profile with KCCQ scores as an anchor.

Table S11. Comparison of Scores from PROMIS+HF-27 and PROMIS+HF-10 Profiles.

Figure S1. Flow diagram for longitudinal sample.

PROMIS+HF-27 Profile Scoring Guide

References

- Rumsfeld JS, Alexander KP, Goff DC Jr, Graham MM, Ho PM, Masoudi FA, Moser DK, Roger VL, Slaughter MS, Smolderen KG, Spertus JA, Sullivan MD, Treat-Jacobson D, Zerwic JJ, American Heart Association Council on Quality of Care and Outcomes Research, Council on Cardiovascular and Stroke Nursing, Council on Epidemiology and Prevention, Council on Peripheral Vascular Disease, and Stroke Council. Cardiovascular health: the importance of mea-

- asuring patient-reported health status: a scientific statement from the American Heart Association. *Circulation*. 2013; **127**: 2233–2249.
2. Institute of Medicine (US) Committee on Quality of Health Care in America. *Crossing the Quality Chasm: A New Health System for the 21st Century*. Washington (DC): National Academies Press (US); 2001.
 3. Patrick DL, Burke LB, Powers JH, Scott JA, Rock EP, Dawisha S, O'Neill R, Kennedy DL. Patient-reported outcomes to support medical product labeling claims: FDA perspective. *Value Health*. 2007; **10**: S125–S137.
 4. Zannad F, Garcia AA, Anker SD, Armstrong PW, Calvo G, Cleland JG, Cohn JN, Dickstein K, Domanski MJ, Ekman I, Filippatos GS, Gheorghiadu M, Hernandez AF, Jaarsma T, Koglin J, Konstam M, Kupfer S, Maggioni AP, Mebazaa A, Metra M, Nowack C, Pieske B, Pina IL, Pocock SJ, Ponikowski P, Rosano G, Ruilope LM, Ruschitzka F, Severin T, Solomon S, Stein K, Stockbridge NL, Stough WG, Swedberg K, Tavazzi L, Voors AA, Wasserman SM, Woehrle H, Zalewski A, McMurray JJ. Clinical outcome endpoints in heart failure trials: a European Society of Cardiology Heart Failure Association consensus document. *Eur J Heart Fail*. 2013; **15**: 1082–1094.
 5. Anker SD, Agewall S, Borggrefe M, Calvert M, Jaime Caro J, Cowie MR, Ford I, Paty JA, Riley JP, Swedberg K, Tavazzi L, Wiklund I, Kirchhof P. The importance of patient-reported outcomes: a call for their comprehensive integration in cardiovascular clinical trials. *Eur Heart J*. 2014; **35**: 2001–2009.
 6. Bragazzi NL, Zhong W, Shu J, Abu Much A, Lotan D, Grupper A, Younis A, Dai H. Burden of heart failure and underlying causes in 195 countries and territories from 1990 to 2017. *Eur J Prev Cardiol*. 2021; **28**: 1682–1690.
 7. Comin-Colet J, Martin Lorenzo T, Gonzalez-Dominguez A, Oliva J, Jimenez MS. Impact of non-cardiovascular comorbidities on the quality of life of patients with chronic heart failure: a scoping review. *Health Qual Life Outcomes*. 2020; **18**: 329.
 8. Kelkar AA, Spertus J, Pang P, Pierson RF, Cody RJ, Pina IL, Hernandez A, Butler J. Utility of patient-reported outcome instruments in heart failure. *JACC Heart Fail*. 2016; **4**: 165–175.
 9. WRITING COMMITTEE MEMBERS, Yancy CW, Jessup M, Bozkurt B, Butler J, Casey DE Jr, Drazner MH, Fonarow GC, Geraci SA, Horwich T, Januzzi JL, Johnson MR, Kasper EK, Levy WC, Masoudi FA, McBride PE, McMurray JJ, Mitchell JE, Peterson PN, Riegel B, Sam F, Stevenson LW, Tang WH, Tsai EJ, Wilkoff BL, American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines. 2013 ACCF/AHA guideline for the management of heart failure: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines. *Circulation*. 2013; **128**: e240–e327.
 10. Koshy AO, Gallivan ER, McGinlay M, Straw S, Drozd M, Toms AG, Gierula J, Cubbon RM, Kearney MT, Witte KK. Prioritizing symptom management in the treatment of chronic heart failure. *ESC Heart Fail*. 2020; **7**: 2193–2207.
 11. Mondesir FL, Zickmund SL, Yang S, Perry G, Galyean P, Nativi-Nicolau J, Kemeyou L, Spertus JA, Stehlik J. Patient perspectives on the completion and use of patient-reported outcome surveys in routine clinical care for heart failure. *Circ Cardiovasc Qual Outcomes*. 2020; **13**: e007027.
 12. Wohlfahrt P, Zickmund SL, Slager S, Allen LA, Nicolau JN, Kfoury AG, Felker GM, Conte J, Flint K, DeVore AD, Selzman CH, Hess R, Spertus JA, Stehlik J. Provider perspectives on the feasibility and utility of routine patient-reported outcomes assessment in heart failure: a qualitative analysis. *J Am Heart Assoc*. 2020; **9**: e013047.
 13. Koonin LM, Hoots B, Tsang CA, Leroy Z, Farris K, Jolly T, Antall P, McCabe B, Zelis CBR, Tong I, Harris AM. Trends in the use of telehealth during the emergence of the COVID-19 pandemic—United States, January–March 2020. *MMWR Morb Mortal Wkly Rep*. 2020; **69**: 1595–1599.
 14. Vaduganathan M, Claggett BL, McMurray JJV, Solomon SD. Health status trajectories before and after hospitalization for heart failure. *Circulation*. 2022 Epub ahead of print.
 15. Nassif ME, Spertus JA, Tang F, Windsor SL, Jones P, Thomas M, Khariton Y, Brush J, Gordon RA, Jermyn R, Jonsson O, Lamba S, Shavelle DM, Kosiborod MN. Association between change in ambulatory hemodynamic pressures and symptoms of heart failure. *Circ Heart Fail*. 2021; **14**: e008446.
 16. Lorem G, Cook S, Leon DA, Emaus N, Schirmer H. Self-reported health as a predictor of mortality: a cohort study of its relation to other health measurements and observation time. *Sci Rep*. 2020; **10**: 4886.
 17. Heidenreich PA, Fonarow GC, Breathett K, Jurgens CY, Pisani BA, Pozehl BJ, Spertus JA, Taylor KG, Thibodeau JT, Yancy CW, Ziaieian B. 2020 ACC/AHA clinical performance and quality measures for adults with heart failure. *Circ Cardiovasc Qual Outcomes*. 2020; **13**: e000099.
 18. Ahmad FS, Kallen MA, Schifferdecker KE, Carluccio KL, Yount SE, Gelow JM, McCullough PA, Kimmel SE, Fisher ES, Cella D. Development and initial validation of the PROMIS®-Plus-HF profile measure. *Circ Heart Fail*. 2019; **12**: e005751.
 19. Hays RD, Bjorner JB, Revicki DA, Spritzer KL, Cella D. Development of physical and mental health summary scores from the patient-reported outcomes measurement information system (PROMIS) global items. *Qual Life Res*. 2009; **18**: 873–880.
 20. Hays RD, Schalet BD, Spritzer KL, Cella D. Two-item PROMIS® global physical and mental health scales. *J Patient Rep Outc*. 2017; **1**: 2.
 21. Hays RD, Spritzer KL, Thompson WW, Cella D. U.S. general population estimate for “excellent” to “poor” self-rated health item. *J Gen Intern Med*. 2015; **30**: 1511–1516.
 22. 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.
 23. Spertus JA, Jones PG, Sandhu AT, Arnold SV. Interpreting the Kansas City Cardiomyopathy Questionnaire in clinical trials and clinical care: JACC state-of-the-art review. *J Am Coll Cardiol*. 2020; **76**: 2379–2390.
 24. PROMIS Cooperative Group. PROMIS® instrument development and validation scientific standards version 2.0. HealthMeasures Website. http://www.healthmeasures.net/images/PROMIS/PROMISStandards_Vers2.0_Final.pdf. Accessed 11 Jun 2022.
 25. Cronbach L. *Essentials of Psychological Testing*, 2nd ed. New York: Harper; 1960.
 26. Eisinga R, Grotenhuis M, Pelzer B. The reliability of a two-item scale: Pearson, Cronbach, or Spearman-Brown? *Int J Public Health*. 2013; **58**: 637–642.
 27. Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. Hillsdale, NJ: L. Erlbaum Associates; 1988.
 28. Cohen J. A power primer. *Psychol Bull*. 1992; **112**: 155–159.
 29. Shi Q, Mendoza TR, Cleland CS. Interpreting patient-reported outcome scores for clinical research and practice: definition, determination, and application of cutpoints. *Med Care*. 2019; **S8**–S12.
 30. Revicki D, Hays RD, Cella D, Sloan J. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *J Clin Epidemiol*. 2008; **61**: 102–109.
 31. Forrest CB, McTigue KM, Hernandez AF, Cohen LW, Cruz H, Haynes K, Kaushal R, Kho AN, Marsolo KA, Nair VP, Platt R, Puro JE, Rothman RL, Shenkman EA, Waitman LR, Williams NA, Carton TW. PCORnet® 2020: current state, accomplishments, and future directions. *J Clin Epidemiol*. 2021; **129**: 60–67.

32. Hripcsak G, Duke JD, Shah NH, Reich CG, Huser V, Schuemie MJ, Suchard MA, Park RW, Wong IC, Rijnbeek PR, van der Lei J, Pratt N, Noren GN, Li YC, Stang PE, Madigan D, Ryan PB. Observational Health Data Sciences and Informatics (OHDSI): opportunities for observational researchers. *Stud Health Technol Inform.* 2015; **216**: 574–578.
33. Cox ED, Dobrozsi SK, Forrest CB, Gerhardt WE, Kliems H, Reeve BB, Rothrock NE, Lai J-S, Svenson JM, Thompson LA, Tran TDN, Tucker CA. Considerations to support use of patient-reported outcomes measurement information system pediatric measures in ambulatory clinics. *J Pediatr.* 2021; **230**: 198–206.e2.
34. Wagner LI, Schink J, Bass M, Patel S, Diaz MV, Rothrock N, Pearman T, Gershon R, Penedo FJ, Rosen S, Cella D. Bringing PROMIS to practice: brief and precise symptom screening in ambulatory cancer care. *Cancer.* 2015; **121**: 927–934.
35. Garcia SF, Wortman K, Cella D, Wagner LI, Bass M, Kircher S, Pearman T, Penedo FJ. Implementing electronic health record-integrated screening of patient-reported symptoms and supportive care needs in a comprehensive cancer center. *Cancer.* 2019; **125**: 4059–4068.
36. Seneviratne MG, Bozkurt S, Patel MI, Seto T, Brooks JD, Blayney DW, Kurian AW, Hernandez-Boussard T. Distribution of global health measures from routinely collected PROMIS surveys in patients with breast cancer or prostate cancer. *Cancer.* 2019; **125**: 943–951.
37. Bingham CO, Bartlett SJ, Merkel PA, Mielenz TJ, Pilkonis PA, Edmundson L, Moore E, Sabharwal RK. Using patient-reported outcomes and PROMIS in research and clinical applications: experiences from the PCORI pilot projects. *Qual Life Res.* 2016; **25**: 2109–2116.
38. Lewis EF, Coles TM, Lewis S, Nelson LM, Barrett A, Romano CD, Stull DE, Turner SJ, Chang CG. Development, psychometric evaluation, and initial feasibility assessment of a symptom tracker for use by patients with heart failure (HFaST). *J Patient Rep Outc.* 2019; **3**: 26.
39. Al Sayah F, Lahtinen M, Bonsel GJ, Ohinmaa A, Johnson JA. A multi-level approach for the use of routinely collected patient-reported outcome measures (PROMs) data in healthcare systems. *J Patient Rep Outc.* 2021: 98.
40. Gibbons C, Porter I, Gonçalves-Bradley DC, Stoilov S, Ricci-Cabello I, Tsangaris E, Gangannagaripalli J, Davey A, Gibbons EJ, Kotzeva A, Evans J, van der Wees PJ, Kontopantelis E, Greenhalgh J, Bower P, Alonso J, Valderas JM. Routine provision of feedback from patient-reported outcome measurements to healthcare providers and patients in clinical practice. *Cochrane Database Syst Rev.* 2021; **10**: CD011589.
41. Stover AM, Haverman L, van Oers HA, Greenhalgh J, Potter CM, ISOQOL PROMs/PREMs in Clinical Practice Implementation Science Work Group. Using an implementation science approach to implement and evaluate patient-reported outcome measures (PROM) initiatives in routine care settings. *Qual Life Res.* 2021; **30**: 3015–3033.
42. Snyder C, Wu AW. Users' guide to integrating patient-reported outcomes in electronic health records [Internet]. Johns Hopkins University 2017. <https://www.pcori.org/sites/default/files/PCORI-JHU-Users-Guide-To-Integrating-Patient-Reported-Outcomes-in-Electronic-Health-Records.pdf>. Accessed on 11 Jun 2022.
43. Withers K, Palmer R, Lewis S, Carolan-Rees G. First steps in PROMs and PREMs collection in Wales as part of the prudent and value-based healthcare agenda. *Qual Life Res.* 2021; **30**: 3157–3170.
44. Wu AW, Kharrazi H, Boulware LE, Snyder CF. Measure once, cut twice—adding patient-reported outcome measures to the electronic health record for comparative effectiveness research. *J Clin Epidemiol.* 2013; **66**: S12–S20.