

RESEARCH

Open Access



Psychometric properties of the heart failure caregiver burden questionnaire among Korean family caregivers: a methodological study

Da-Young Kim^{1†} , Heemin Chae^{1†} , Eun-Jung Kim¹ and Youn-Jung Son^{2*}

Abstract

Background Heart failure (HF) can limit the capacity of patients to perform daily activities, increasing their dependence on caregivers and contributing to significant caregiver burden. Accurate assessment of this burden is crucial for elucidating its impact and guiding the development of evidence-based strategies to support both patients and their caregivers. This study aims to translate, culturally adapt, and evaluate the psychometric properties of the heart failure caregiver burden questionnaire (HF-CQ) for application among Korean caregivers of individuals with HF.

Methods The HF-CQ was translated into Korean and culturally adapted in accordance with the World Health Organization guidelines. A total of 328 caregivers of patients with HF were enrolled between June 2021 and December 2023. To evaluate construct validity, a confirmatory factor analysis was used, including the assessments of convergent and discriminant validity of the instrument. Criterion validity and known-group validity were also examined. Reliability was determined by calculating Cronbach's α for internal consistency and the intraclass correlation coefficient (ICC) for test–retest stability.

Results The Korean version of the HF-CQ confirmed a three-factor structure—physical wellbeing, emotional wellbeing, and lifestyle—comprising 20 items. The instrument demonstrated excellent content validity. Construct validity was supported using acceptable model fit, standardized factor loadings and average variance extracted values > 0.50 , composite reliability values > 0.70 , and heterotrait–monotrait ratios < 0.85 . Criterion validity was verified via significant correlations with depressive symptoms ($r = 0.44, p < 0.001$) and with health-related quality of life ($r = -0.23, p < 0.001$). Known-group validity was not demonstrated, with no significant differences in caregiver burden scores caregivers of patients in NYHA Class II and those in Class III/IV ($t = -1.680, p = 0.094$). Internal consistency was excellent (Cronbach's $\alpha = 0.95$), and test–retest reliability was good (ICC = 0.73).

Conclusions The 20-item Korean HF-CQ demonstrated sound psychometric properties, offering a culturally adapted tool for evaluating caregiver burden in HF. Implementation in clinical and community settings can facilitate

[†]Da-Young Kim and Heemin Chae contributed equally to this work.

*Correspondence:
Youn-Jung Son
yjson@cau.ac.kr

Full list of author information is available at the end of the article



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

early detection of caregivers experiencing high levels of burden, thereby allowing prompt recognition and the development of individualized support strategies.

Keywords Caregiver burden, Heart failure, Psychometric, Scale, Reliability, Validity

Background

Heart failure (HF) is a chronic clinical syndrome characterized by impaired cardiac filling or ejection, leading to insufficient blood flow to meet systemic metabolic needs [1]. With an aging global population and advances in treatment extending survival, the prevalence of HF continues to increase [2, 3]. Older individuals with HF face a disproportionate burden, being more prone to cognitive impairment, functional decline, and ongoing symptoms such as dyspnea, peripheral edema, and fatigue [4]. These impairments limit engagement in daily activities and hinder adherence to key self-care behaviors—including medication adherence, promptly recognizing symptom changes, and restricting dietary sodium [5]—all of which are essential to optimizing HF outcomes. As a result, quality of life often deteriorates, and rates of hospital readmissions and mortality rise [6], underscoring the continued necessity for caregiver involvement [7].

Informal caregivers—most often family members or close acquaintances—provide unpaid assistance that includes symptom monitoring, support for self-care, decision-making, and psychosocial care [7]. Prior studies suggest that family caregiving influences the physical and psychological wellbeing of caregivers as well as the health outcomes of patients with HF [8, 9]. Nonetheless, caring for a person with HF often entails prolonged periods of high-intensity support owing to the unpredictable trajectory of the disease, progressive nature, and the advanced age of many patients, leading to sustained periods of high-intensity caregiving [8, 10].

Such demand can lead to sustained physical strain, emotional distress, social restrictions, and financial challenges for caregivers [10, 11], which may persist for years and ultimately influence care quality and patient outcomes [7]. Accordingly, accurate evaluation of caregiver burden is essential before designing interventions to alleviate caregiving burden and enhance care quality for caregivers of patients with HF.

Widely utilized tools for assessing caregiver burden include the Zarit Burden Interview, the Caregiver Burden Inventory, the Caregiver Burden Scale, and the Caregiver Strain Index [12–15]. While these instruments are used to assess multiple dimensions of caregiver burden, they are not developed specifically for HF and may therefore overlook disease-specific challenges such as frequent symptom fluctuations, complex treatment regimens, and evolving care needs. The heart failure caregiver burden questionnaire (HF-CQ) was created to overcome these limitations. It contains 21 items organized into three

domains—physical, emotional/psychological, and lifestyle—and has shown robust validity and reliability in its initial validation studies [16].

In South Korea, the prevalence of HF has steadily increased from 0.77% in 2002 to 2.58% in 2020, alongside rapid population aging [17]. Furthermore, Korean society holds distinct sociocultural expectations that family members, such as spouses or adult children, bear the primary responsibility for patient care, and that caregiving should be performed within the family. These characteristics mean that the burden patterns experienced by caregivers in Korean society may manifest differently than those observed in other cultural contexts. To enable accurate assessment and meaningful interpretation of HF-specific caregiver burden in this population, a culturally adapted and psychometrically validated Korean version of the HF-CQ is warranted. Therefore, this study aims to translate the HF-CQ into Korean, culturally adapt it, and evaluate its psychometric properties, thereby establishing a standardized tool to inform caregiver support strategies and policy development in HF care.

Methods

Study design

This methodological study involved translating the HF-CQ into Korean and undertaking cultural adaptation, followed by verification of the reliability and validity (content, construct, criterion, and known-group validity) of the Korean version of HF-CQ.

Setting and samples

This study included caregivers of patients with HF who attended the cardiology departments of two university hospitals in Seoul. Inclusion criteria included (1) adults aged ≥ 18 years, (2) family or informal caregivers (including relatives, partners or friends) who provided unpaid caregiving to a patient with HF, and (3) individuals who had provided regular caregiving for at least 6 months. Exclusion criteria included: (1) patients or caregivers diagnosed with a serious illness, such as terminal cancer or kidney disease requiring dialysis; (2) those with cognitive impairments, such as dementia, or a diagnosed mental illness, such as schizophrenia; and (3) those unable to complete the questionnaire. The diagnosed conditions of patients with HF were identified through both electronic medical records and self-reports, whereas those of caregivers were confirmed based on self-reports.

The required sample size was estimated based on previous studies indicating that a minimum of 200 participants

are needed for confirmatory factor analysis (CFA) [18]. To account for potential incomplete data, 330 participants were recruited. The caregivers who participated in the pilot test were excluded from main study to prevent potential response bias arising from prior exposure to the instrument.

Measures

Sociodemographic characteristics

Caregiver sociodemographic characteristics included age, sex, relationship to the patient, education level, marital status, working status, monthly income, body mass index (BMI), number of chronic conditions, living with patients, duration of caregiving, and time spent in caregiving. Patient characteristics encompassed age, sex, New York Heart Association (NYHA) functional classification, and time since the diagnosis of HF.

Heart failure-caregiver questionnaire (HF-CQ)

Caregiver burden was measured using the HF-CQ developed by Stromberg et al. [16]. The instrument comprises 21 items across three domains: physical wellbeing, emotional wellbeing, and lifestyle (5, 11, and 5 items, respectively). Each item is rated on a 5-point Likert scale (0 = “Not at all”, 1 = “Almost not”, 2 = “Moderate”, 3 = “Quite a lot”, 4 = “Very”). Domain scores are standardized to a 100-point scale, and the total score—ranging from 0–100—is calculated as the mean of the three domain scores, with higher scores indicating a greater degree of caregiving burden. In the original validation, the HF-CQ demonstrated excellent internal consistency, with a Cronbach’s alpha of 0.94 [16].

Patient Health Questionnaire-9 (PHQ-9) for depressive symptoms

Depressive symptoms were assessed using the validated Korean adaptation of the Patient Health Questionnaire-9 (PHQ-9) [19], which was originally developed by Spitzer and colleagues [20]. This scale consists of nine items, each rated on a 4-point Likert scale ranging from 0 (“not at all”) to 3 (“nearly every day”), with higher total scores reflecting more severe depressive symptoms. In the Korean validation study, the PHQ-9 showed satisfactory internal consistency (Cronbach’s alpha = 0.81) [19], and a similar level of reliability was observed in the current study (Cronbach’s alpha = 0.83).

EuroQol 5-Dimension 5-level (EQ-5D-5 L) for health-related quality of life

Health-related quality of life in caregivers of patients with HF was evaluated using the Korean version of EuroQol 5-dimension 5-level (EQ-5D-5 L) instrument [21]. This tool assesses five domains—mobility, self-care, usual activities, pain/discomfort, and anxiety/

depression—each scored on a five-point scale reflecting the degree of impairment. Higher total scores are considered to be a better quality of life. In its Korean validation, the EQ-5D-5 L demonstrates acceptable reliability (Intra-class correlation coefficient, ICC = 0.75) [21]. In the current study, reliability was of a similar magnitude (ICC = 0.77), with Cronbach’s alpha of 0.93.

Translation process

The translation process followed the World Health Organization’s recommended methodology [22], comprising forward translation, expert panel evaluation, back-translation, and pre-testing. Initially, two bilingual co-authors independently translated the original English instrument into Korean. Their translations were compared, and any discrepancies were resolved through discussion to produce a single reconciled version. The principal investigator then reviewed this draft for conceptual accuracy, linguistic appropriateness, and potential item revisions. Subsequently, an independent bilingual translator with expertise in health-related measurement tools performed a back-translation into English. The original and back-translated forms were examined to confirm conceptual and cultural equivalence, resulting in the finalized preliminary Korean version of the instrument.

Content validity

Content validity of the preliminary instrument was evaluated by a panel of ten experts.

The expert panel consisted of five nursing professors with experience in instrument development and psychometric analysis, and five clinical nurses with over ten years of experience in patient care. Individuals without direct experience in nursing HF patients or without sufficient knowledge of instrument development were excluded. Each expert evaluated the relevance and clarity of every item using a 4-point Likert scale (1 = not relevant to 4 = highly relevant). The content validity index (CVI) was then computed at both the item (I-CVI) and scale levels (S-CVI/Ave). The I-CVI was determined as the proportion of experts assigning a rating of 3 (*relevant*) or 4 (*highly relevant*), whereas the S-CVI/Ave was calculated as the means of the I-CVI values across all items. For expert panels of three or more members, I-CVI ≥ 0.78 and an S-CVI/Ave ≥ 0.90 are considered acceptable [23].

Pilot study

A pilot study was conducted using the content-validated preliminary tool among 10 primary caregivers of patients with HF (five women and five men; mean age = 60.3 years) who met the same inclusion criteria as the main study and were recruited from the cardiology outpatient clinics of two university hospitals in Seoul. The pilot study

examined item clarity, ease of understanding, and suitability of the questionnaire format. The average completion time of the questionnaire was 10 min. Consequently, participants reported that the intent of each item was clearly understood, the wording was free of ambiguity, and the questionnaire was easy to complete and logically structured. Therefore, no revisions were required, and the final version of the tool comprised 21 items.

Data collection

Data collection was undertaken from June 2021 to December 2023 using a self-administered structured questionnaire. Two trained research assistants collected the data. Prior to data collection, they received standardized training on the research protocol, including the survey procedures, and inter-rater reliability was checked to ensure consistency in data collection. Before participation, all respondents provided written informed consent after receiving an explanation of the study purpose, procedures, assurance of anonymity, and confirmation that the collected data would be used exclusively for research. For the test–retest reliability, a follow-up survey was conducted 4 weeks after the initial survey with participants who had consented to retest participation during the first survey.

Data analysis

Data analysis was conducted using SPSS/WIN version 27.0 and AMOS version 28.0. For general and clinical characteristics of the participants, categorical variables were summarized as frequencies and percentages, while continuous variables were reported as means and standard deviations. Item analysis included examining mean and standard deviation of each item to detect extreme values and assess overall item quality. Normality was evaluated based on skewness (± 3) and kurtosis (± 8) thresholds [24]. Items were deleted if their corrected item–total correlation (C-ITC) was less than 0.30 and if their removal resulted in an increase in the overall internal consistency of the scale [25].

Construct validity was examined through CFA using the maximum likelihood (ML) estimation with bootstrapping method (5,000 samples) to ensure robust parameter estimation under absence of multivariate normality. In addition, a Monte Carlo simulation-based CFA [26] was employed to evaluate the convergence of the results with those obtained from bootstrap ML-based CFA, thereby providing robust evidence for the reproducibility and empirical reliability of the measurement model. Multiple model fit indices were evaluated to determine the appropriateness of the factor structure. Given the sensitivity of χ^2 to sample size, the chi-square minimum discrepancy divided by degrees of freedom (CMIN/DF, χ^2/df) was also calculated. Model fit was

considered acceptable when CMIN/DF ≤ 5 , the normed fit index (NFI) ≥ 0.80 , the comparative fit index (CFI) ≥ 0.80 , the Tucker–Lewis index (TLI) ≥ 0.80 , and the standardized root mean squared residual (SRMR) ≤ 0.08 [27]. For the roots mean square error of approximation (RMSEA), values ≤ 0.08 were interpreted as indicating goodness-of-fit. Whereas values between 0.80 and 0.10 were taken to reflect an adequate fit [27]. Model selection was guided by the Akaike Information Criterion (AIC) values, with the smallest AIC value denoting the most parsimonious and best-fitting model; AIC differences of less than 0.01 were regarded as trivial.

To evaluate item-level convergent and discriminant validity, established psychometric thresholds were applied. Convergent validity was considered adequate when standardized factor loadings (λ) were ≥ 0.50 , composite reliability (CR) ≥ 0.70 , and average variance extracted (AVE) was ≥ 0.50 [27]. Discriminant validity was deemed acceptable when Heterotrait–monotrait ratio of correlations (HTMT) was below 0.85 [28]. Criterion validity was evaluated by calculating Pearson correlation coefficients between scores on the Korean version of the HF-CQ, depressive symptoms as measured by the PHQ-9, and health-related quality of life as measured by EQ-5D-5 L. Known-group validity was confirmed by conducting an independent t-test to determine whether there was a difference in the total score of the caregiver burden between the predefined groups (NYHA II vs. III/IV) [29].

Internal consistency reliability of the Korean version of HF-CQ was evaluated using Cronbach's α . Test–retest reliability was examined over a 4-week interval following the initial administration, with ICC used to assess stability. Values to Cronbach's α and ICC equal to exceeding 0.70 were interpreted as indicative of satisfactory reliability and stability [18].

Results

Sample characteristics

Table 1 shows the general and clinical characteristics of family caregivers and their care recipients. Among the total sample of 328 caregivers, the mean age of participants was 60.98 ± 13.58 years. Most caregivers were female (73.2%), spouses of patients (62.5%), had education above college level (42.1%), were married (83.5%), employed (51.2%), had a monthly income of ≥ 200 Korean won (50.6%), had a normal BMI (39.3%), had no chronic disease (57.3%), and lived with the patients (80.2%). On average, caregivers provided care for 55.06 ± 81.09 months, dedicating 25.23 ± 25.83 h/week. The patients had a mean age of 64.79 ± 13.15 years; most were male (64.9%), classified as NYHA class II (57.3%), and had an average duration since HF diagnosis of 3.70 ± 4.12 years.

Table 1 General and clinical characteristics of patients with HF and their caregivers

Variables	Caregivers	Patients
	Mean $\pm$ SD or <i>n</i> (%)	
Age (years)	60.98 $\pm$ 13.58	64.79 $\pm$ 13.15
$\leq$ 39	29 (8.8)	14 (4.3)
40–49	54 (16.5)	31 (9.5)
50–59	78 (23.8)	61 (18.6)
60–69	87 (26.5)	85 (25.9)
$\geq$ 70	80 (24.4)	137 (41.8)
Sex		
Male	88 (26.8)	213 (64.9)
Female	240 (73.2)	115 (35.1)
Relationship to the patients		
Spouse	205 (62.5)	-
Child	95 (29.0)	-
Relatives, friends etc.	28 (8.5)	-
Education level		
Below high school	190 (57.9)	-
Above college	138 (42.1)	-
Marital status		
Married	274 (83.5)	-
Unmarried	41 (12.5)	-
Separated/divorced/widowed	13 (3.9)	-
Working status		
Working	168 (51.2)	-
Not working	160 (48.8)	-
Monthly income (10,000 Korean won)		
< 200	162 (49.4)	-
$\geq$ 200	166 (50.6)	-
BMI (kg/m ²)	23.66 $\pm$ 3.50	-
Underweight (< 18.5)	15 (4.6)	-
Normal (18.5–22.9)	129 (39.3)	-
Overweight (23.0–24.9)	81 (24.7)	-
Obesity ($\geq$ 25.0)	103 (31.4)	-
Number of chronic conditions		
0	188 (57.3)	-
1	83 (25.3)	-
$\geq$ 2	57 (17.4)	-
Living with patients		
Yes	263 (80.2)	-
No	65 (19.8)	-
Duration of caregiving (month)	55.06 $\pm$ 81.09	-
Time spent in caregiving (hours/week)	25.23 $\pm$ 25.83	-
NYHA functional classification		
II	-	188 (57.3)
III or IV	-	140 (42.7)
Time since the diagnosis of HF (years)	-	3.70 $\pm$ 4.12

SD, standard deviation; HF, heart failure; BMI, body mass index; NYHA, New York Heart Association

Content validity

The item-level I-CVI was 1.00 for 18 items and 0.90 for three items (16, 18, and 19). The S-CVI/Ave was 0.99. Accordingly, no items were deleted; however, expert feedback was incorporated to enhance clarity and

comprehension, resulting in the revision and refinement of items 16, 18, and 19. Item 16 was modified from “because of doing limited things for others” to “due to the limitations in your ability to do things for others.” Item 18 was revised from “avoiding making plans” to “refrain from making plans altogether.” Item 19 was revised from “ever felt that you could not be far away” to “could not live far away.”

Item analysis

In Table 2, item analysis was conducted on 20 items, excluding item 20 (“Are you currently retired or fully disabled?”), owing to its dichotomous “Yes” or “No” responses. The means and standard deviations ranged from 0.52 to 2.32 and 0.97–1.42, respectively, indicating good item quality without extreme values. All skewness and kurtosis values met the normality assumption, falling within ± 3 and ± 8 , respectively [24]. All C-ITC values for the 20 items were above 0.30 [25], except for item 15. However, no items were deleted during item analysis because removing item 15 did not increase the overall Cronbach's α of the total scale.

Construct validity

The CFA results using the Monte Carlo simulation were consistent with those obtained through the bootstrap ML estimation, confirming the methodological rigor and empirical robustness of factor structures. Model 1 indices were as follows: χ^2 (167, $N=328$) = 696.56, $p < 0.05$; SRMR = 0.06; NFI = 0.84; TLI = 0.86; CFI = 0.87; RMSEA = 0.098. However, item 15 had a standardized factor loading below the acceptable threshold of 0.30, indicating poor factor representation; therefore, item 15 was deleted (Table 2). Following its removal, another CFA Model 2 was tested, showing improved fit indices: χ^2 (149, $N=328$) = 649.57, $p < 0.05$; SRMR = 0.06; NFI = 0.85; TLI = 0.86; CFI = 0.88; RMSEA = 0.101.

Findings of model comparison indicate that Model 2 had a lower AIC value (769.57) than that of Model 1 (822.56). Additionally, the χ^2 difference test confirmed that Model 2 provided a desired fit to the data, indicating that Model 2 provided a significantly better and more parsimonious fit than Model 1. Consequently, Model 2—with a three-factor structure comprising 19 items—was determined to be the most appropriate (Table 3). Finally, the Korean version of HF-CQ retained this three-factor structure, including 20 items in total, with the categorical item 20, consistent with the original HF-CQ conceptual framework (Model 2).

For item convergent and discriminant validity, standardized factor loadings ranged from 0.57 to 0.90, exceeding the 0.50 threshold [27] (Fig. 1). The CR values ranged from 0.82 to 0.92, surpassing the recommended 0.70 level, while AVE values ranged from 0.53 to 0.62,

Table 2 Item analysis and standardized factor loadings of model 1

Item	Mean $\pm$ SD or n (%)	Skewness	Kurtosis	C-ITC	Cron- bach's α if item deleted	Stan- dardized factor loadings	Item Status (Korean version of the HF-CQ)
Factor 1: Physical well-being							
1. How much has caregiving been physically hard work?	0.83 $\pm$ 1.08	1.11	0.30	0.72	0.93	0.87	Retained
2. How much has caregiving made you feel physically tired?	0.81 $\pm$ 1.04	1.07	0.34	0.74	0.93	0.90	Retained
3. How much has caregiving caused you aches and pains?	0.61 $\pm$ 0.90	1.30	0.69	0.60	0.94	0.72	Retained
4. How much has caregiving made it difficult to sleep?	0.59 $\pm$ 0.98	1.62	1.78	0.62	0.93	0.68	Retained
5. How much has caregiving made you neglect your own health?	0.60 $\pm$ 0.98	1.56	1.59	0.66	0.93	0.73	Retained
Factor 2: Emotional well-being							
6. How much have you felt the person you care for asks too much of you?	0.54 $\pm$ 1.00	1.98	3.20	0.65	0.93	0.67	Retained
7. How much have you felt like you need to do more for the person you care for?	1.28 $\pm$ 1.40	0.58	-1.10	0.66	0.93	0.64	Retained
8. How much has caregiving made you feel stressed?	1.11 $\pm$ 1.20	0.79	-0.34	0.76	0.93	0.83	Retained
9. How much has caregiving made you feel frustrated?	0.66 $\pm$ 1.03	1.58	1.72	0.66	0.93	0.77	Retained
10. How much has caregiving made it difficult to focus or concentrate on other things?	0.66 $\pm$ 1.02	1.57	1.78	0.76	0.93	0.80	Retained
11. How much has caregiving made you feel mentally tired?	1.08 $\pm$ 1.19	0.89	-0.11	0.72	0.93	0.81	Retained
12. How much has caregiving made you worry about the person you care for?	2.32 $\pm$ 1.24	-0.52	-0.74	0.54	0.94	0.57	Retained
13. How much has caregiving made you feel lonely?	0.77 $\pm$ 1.06	1.20	0.49	0.66	0.93	0.70	Retained
14. How much has caregiving made you feel emotionally drained?	0.84 $\pm$ 1.13	1.23	0.56	0.75	0.93	0.83	Retained
15. How much support have you had from family or friends?	2.17 $\pm$ 1.32	-0.28	-0.97	0.18	0.93	0.16	Deleted
16. How guilty have you felt due to the limitations in your ability to do things for others?	0.59 $\pm$ 0.97	1.63	1.92	0.65	0.93	0.63	Retained
Factor 3: Lifestyle							
17. How much have you felt like you have no time for yourself?	0.78 $\pm$ 1.12	1.25	0.49	0.70	0.93	0.78	Retained
18. How much has caregiving caused you to change your plans or refrain from making plans altogether?	0.95 $\pm$ 1.23	1.02	-0.20	0.74	0.93	0.83	Retained
19. How much have you felt that you could not live far away?	1.29 $\pm$ 1.42	0.64	-0.99	0.59	0.94	0.63	Retained
*20. Are you currently retired or fully disabled?	Yes: 248 (75.6) No: 80 (24.4)						Retained
21. How much has caregiving made it difficult to do paid work?	0.52 $\pm$ 0.98	1.93	2.91	0.59	0.94	0.65	Retained

C-ITC, corrected item-total correlation; SD, standard deviation; HF-CQ: Heart failure-caregiver burden questionnaire

*Dichotomous response item (Yes or No)

Table 3 Model fit indices from confirmatory factor analysis ($N = 328$)

	χ^2	DF	CMIN/DF (χ^2/df)	SRMR	NFI	TLI	CFI	RMSEA	AIC
Model 1	696.56 ($p < 0.001$)	167	4.17	0.06	0.84	0.86	0.87	0.098	822.56
Model 2	649.57 ($p < 0.001$)	149	4.36	0.06	0.85	0.86	0.88	0.101	769.57

AIC, akaike information criterion; CFI, comparative fit index; DF, degrees of freedom; NFI, normed fit index; SRMR, standardized root mean square Residual; RMSEA, root mean square error of approximation; TLI, tucker-lewis index

Model 2 = Model 1 with item 15 deleted

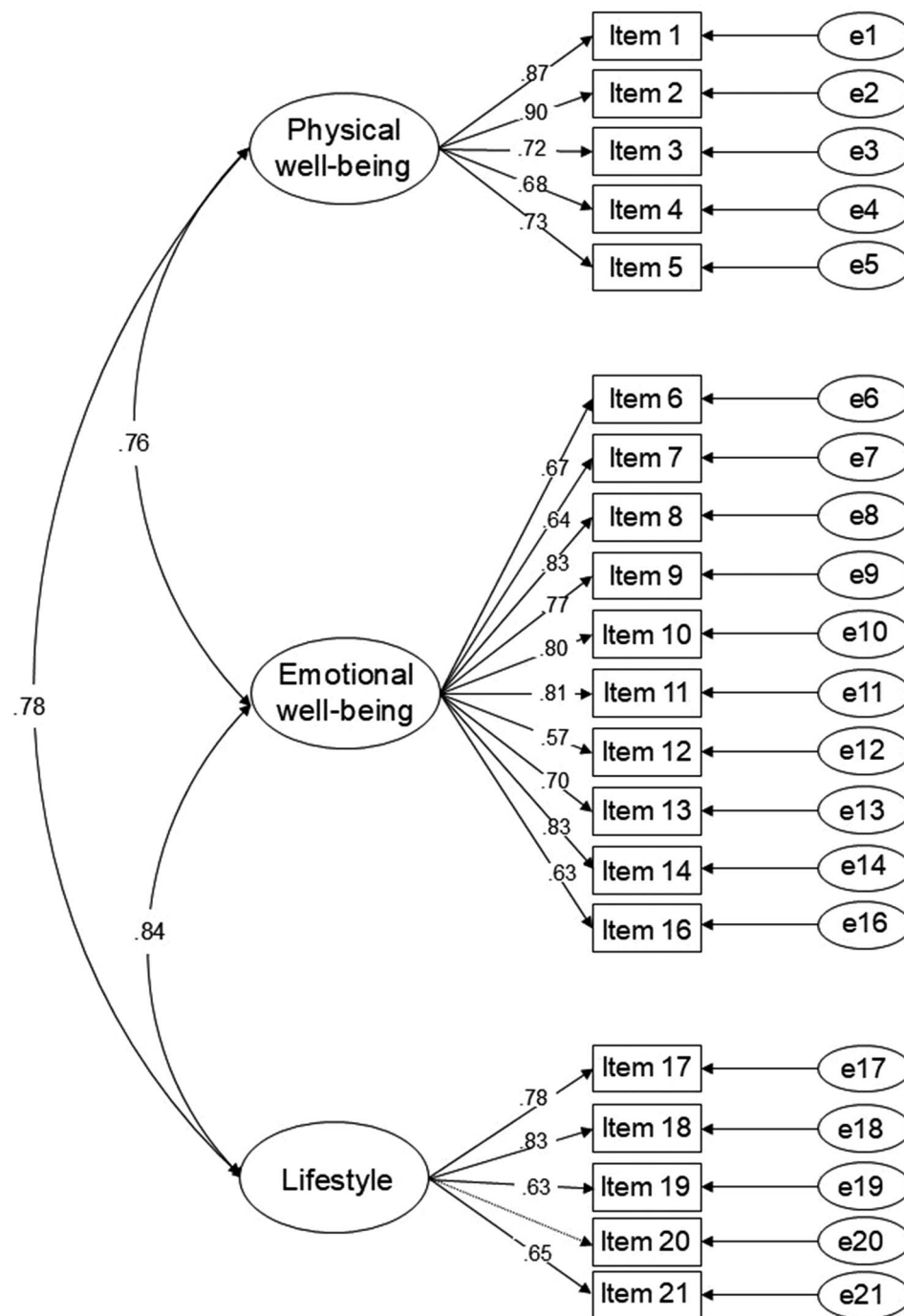


Fig. 1 Final measurement model from confirmatory factor analysis

above the 0.50 criteria [27]. Discriminant validity was confirmed with HTMT values between 0.69 and 0.76, all < 0.85 cutoff. Accordingly, convergent and discriminant validity were established (Table 4).

Internal consistency and test-retest reliability

The total internal consistency of the Korean version of the HF-CQ was high, with a Cronbach's α of 0.95. Cronbach's α values for the subscales were 0.89, 0.91, and 0.80 for physical well-being, emotional well-being, and lifestyle, respectively (Table 4). Among the 328 participants, a subset of 223 completed an additional survey after a 4-week interval. The test-retest reliability of the Korean

Table 4 Construct validity, internal consistency, and test-retest reliability

Factors	Mean ± SD	CR	AVE	HTMT			Internal consistency (N = 328) Cronbach's α	Test-retest reliability (N = 223)	
				1	2	3		ICC	95% CI
Factor 1 (5 items) Physical well-being	17.36 ± 20.87	0.89	0.62	1			0.89	0.71	[0.62, 0.78]
Factor 2 (10 items) Emotional well-being	24.62 ± 21.20	0.92	0.53	0.72	1		0.91	0.70	[0.60, 0.77]
Factor 3 (5 items) Lifestyle	22.14 ± 23.77	0.82	0.53	0.69	0.76	1	0.80	0.76	[0.68, 0.81]
Total Score (20 items)	22.19 ± 19.74						0.95	0.73	[0.65, 0.80]

AVE, average variance extracted; CI, confidence interval; CR, composite reliability; HTMT, heterotrait-monotrait ratio of the correlations; ICC, intraclass correlation coefficient; SD, standard deviation

Table 5 Criterion validity among HF-CQ, PHQ-9, and EQ-5D-5L

Measures (variables)	Mean ± SD	HF-CQ	PHQ-9	EQ-5D-5L
HF-CQ (Caregiver burden)	22.19 ± 19.74	1		
PHQ-9 (Depression)	3.74 ± 3.62	0.44 ($p < 0.001$)	1	
EQ-5D-5 L (Health-related quality of life)	0.90 ± 0.10	-0.23 ($p < 0.001$)	-0.41 ($p < 0.001$)	1

EQ-5D-5 L, EuroQol 5-Dimension, 5 level version; HF-CQ, Heart Failure Caregiver Questionnaire; PHQ-9, Patient Health Questionnaire-9, SD, standard deviation

version of the HF-CQ was ICC = 0.73 (95% CI = 0.65–0.80), with subscale ICC values ranging from 0.70 to 0.76 (Table 4).

Criterion validity

For criterion validity, correlations were examined among caregiver burden (Korean version of HF-CQ), depressive symptoms, and health-related quality of life. The Korean version of HF-CQ showed a significant positive correlation with depressive symptoms measured using the PHQ-9 ($r = 0.44$, $p < 0.001$), and a significant negative correlation with health-related quality of life measured using the EQ-5D-5 L ($r = -0.23$, $p < 0.001$; Table 5).

Known-group validity

For known-group validity, the differences between caregivers of the NYHA class II and class III/IV were evaluated. The mean difference in caregiver burden was not statistically significant between caregivers of patients in NYHA Class II (20.61 ± 18.43) and those of patients in NYHA classes III/IV (24.30 ± 21.26) ($t(326) = -1.680$, $p = 0.094$).

Discussion

This study demonstrates that the Korean version of the HF-CQ is a reliable instrument for assessing the multi-dimensional burden of family caregiving in HF. Adapted from a Western-developed measure, it addresses the gap

in tools that capture the physical, emotional, and lifestyle influence of caregiving within the Korean sociocultural context. The final version includes 20 items across three dimensions—physical wellbeing, emotional wellbeing, and lifestyle—offering a comprehensive and culturally relevant assessment of caregiver burden.

In this present study, content validity indices exceeded recommended thresholds, indicating strong semantic and conceptual alignment with the original construct of “caregiver burden.” Several lifestyle items (e.g., Items 17, 18, and 19) were rephrased to reflect caregiving-related lifestyle restrictions within the Korean context. For instance, Item 18, originally phrased as “avoid making plans,” was revised to “refrain from making plans altogether?” to better accurately capture the passive nature of social withdrawal commonly reported by Korean caregivers. Furthermore, during the cultural adaptation process, the tool was comprehensively evaluated for linguistic clarity, ease of understanding, and cultural appropriateness by 10 primary family caregivers of patients with HF, confirming that face validity was secured.

Our CFA results showed that item 15, “How much support have you had from family or friends?”, failed to meet the criteria for C-ITC and factor loadings in the present and original studies [16]. Removing Item 15 improved model fit indices. All items in the emotional well-being factor were found to consistently assess the emotional aspect; however, Item 15 exhibited conceptual heterogeneity, as it measured the positive construct of “support,” rather than the negative construct of “burden.” Convergent validity was supported via strong correlations with related constructs, while discriminant validity was confirmed via low correlations with unrelated measures and significant differences across known groups.

As no gold standard exists for assessing caregiver burden in HF, criterion validity was examined using related constructs: depressive symptoms measured with the PHQ-9 [30, 31] and health-related quality of life assessed with the EQ-5D-5 L [32, 33]. The Korean version of HF-CQ demonstrated a moderate positive correlation with depressive symptoms and a negative correlation

with health-related quality of life, consistent with the theoretical expectation that higher caregiver burden is associated with poorer psychological well-being and reduced quality of life [29, 34]. These findings strengthen the evidence for the validity of the instrument and highlight its utility in identifying caregivers at heightened risk for adverse psychological and health outcomes.

In this study, caregiver burden for patients with NYHA Class II was slightly lower than that for patients with NYHA Class III/IV. However, the difference was not statistically significant, a finding similar to the original scale development study [16]. These findings may be attributable to the limited sample size, or to other patients' disease related factors contributing to an increased caregiving burden for family members. A recent systematic review identified key patient factors influencing caregiver burden, including the frequency of hospital readmissions, medication use, and the number and type of comorbidities [2, 29]. Accordingly, future research should take these factors into account and further validate the HF-CQ in a representative sample through known-group comparisons.

The instrument demonstrated excellent internal consistency and strong test–retest reliability, surpassing established psychometric thresholds [18] and closely approximating the reliability indices reported in the original validation study [16]. Therefore, the Korean version of HF-CQ is confirmed to possess high reliability and measurement stability.

The mean caregiver burden score in this study is comparable to that of the original U.S. sample [16]. However, the higher lifestyle burden in Korean HF caregivers may reflect the enduring influence of Confucian and patriarchal values. In Korean society, caregiving—especially by women—is still predominantly regarded as a familial duty rather than a potential source of personal strain [35]. In this cultural context, Korean women often view family caregiving as a natural obligation. As a result, they may perceive physical and emotional strain as less burdensome, leading to lower scores in these domains despite substantial caregiving demands [35]. These cultural influences highlight the need for disease-specific instruments that can capture such nuanced patterns.

The Korean version of the HF-CQ offers a more comprehensive assessment of the multifaceted burden of caregiving than that of generic tools such as the Zarit Burden Interview or the Caregiver Burden Inventory [12, 14]. Nevertheless, the Korean version differs slightly from the original HF-CQ, as one item from the emotional well-being domain was removed, yielding a final 20-item instrument. This structural variation warrants caution when comparing total scores or emotional well-being outcomes with the original 21-item version. Despite this difference, the validated Korean HF-CQ remains a robust

measure for accurately assessing caregiver burden in HF, enabling cross-cultural comparisons, and guiding the development of targeted interventions to reduce burden. Future research should evaluate its responsiveness to change and explore potential mediating pathways linking caregiver burden, caregiver characteristics, and patient outcomes in HF care.

Limitation

This study has some limitations. First, participants were recruited via convenience sampling from a single medical institution in Korea, which may restrict the generalizability of the findings to all family caregivers of patients with HF. Second, cut-off points of the Korean version of HF-CQ were not examined in this study. Therefore, it is crucial to further examine the optimal threshold for identifying high-risk groups with large, representative Korean caregiver populations across multiple settings and confirm its broader applicability. Third, the emotional well-being domain was reduced to 20 items following the removal of one item, which restricts direct comparability of emotional well-being domain scores and total scores with those obtained from the original 21-item HF-CQ. Caution is therefore warranted when making such comparisons. Lastly, longitudinal stability was not assessed, leaving it uncertain whether the factor structure of the tool remains stable over time or across different caregiver subgroups.

Conclusions

The present study demonstrates a comprehensive psychometric validation of the Korean adaptation of the HF-CQ, originally developed in a Western context. The findings indicate strong content, construct, and criterion validity, as well as high internal consistency and test–retest reliability.

The validated Korean version of HF-CQ is a 20-item scale encompassing three domains—physical wellbeing, emotional wellbeing, and lifestyle—designed to capture the multidimensional nature of caregiving burden in HF. When implemented in clinical and community settings, the HF-CQ offers a standardized, validated tool for screening high-risk caregivers of patients with HF, enabling timely recognition of elevated burden and guiding the development of targeted interventions to reduce stress and improve outcomes.

Abbreviations

HF	Heart Failure
HF-CQ	Heart Failure Caregiver Burden questionnaire
CFA	Confirmatory Factor Analysis
BMI	Body Mass Index
PHQ-9	Patient Health Questionnaire-9
NYHA	New York Heart Association
ICC	Intraclass Correlation Coefficient
CVI	Content Validity Index

I-CVI	Item-level Content Validity Index
S-CVI/Ave	Scale-level Content Validity Index/ Average method
C-ITC	Corrected Item–Total Correlation
CMIN/DF	Chi-square Minimum discrepancy/Degrees of Freedom
NFI	Normed Fit Index
CFI	Comparative Fit Index
SRMR	Standardized Root Mean squared Residual
TLI	Tucker–Lewis Index
RMSEA	Root Mean Square Error of Approximation
AIC	Akaike Information Criterion
CR	Composite Reliability
AVE	Average Variance Extracted
HTMT	Heterotrait–Monotrait ratio of correlations

Acknowledgements

The authors acknowledge and thank all the research participants.

Author contributions

DYK: Data curation, Visualization, Software, Writing – original draft, Writing – review & editing. HMC: Formal analysis, Data curation, Writing – original draft, Writing – review & editing. EJK: Formal analysis, Writing – original draft, Writing – review & editing. YJS: Conceptualization, Methodology, Validation, Supervision, Project administration, Writing – original draft, Writing – review & editing.

Funding

This research was supported by the National Research Foundation of Korea (NRF) grants funded by the Korea government (MSIT) (grant nos. 2021R1A2C2004406 and RS-2025-00562535).

Data availability

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate

This study received approval from the Institutional Review Board of Chung-Ang University (approval no. 1041078-20250204-HR-028) for the ethical protection of human subjects. The study was conducted in accordance with the Declaration of Helsinki. Written informed consent was obtained from all participants. Participants voluntarily consented to take part after fully understanding the study's purpose and procedures.

Consent for publication

Informed consent for the publication of participant's personal or clinical details was obtained from all individuals included in this study.

Competing interests

The authors declare no competing interests.

Author details

¹School of Nursing, Chung-Ang University, 84 Heukseok-ro, Dongjak-gu, Seoul 06974, Republic of Korea

²Red Cross College of Nursing, Chung-Ang University, 84 Heukseok-ro, Dongjak-gu, Seoul 06974, Republic of Korea

Received: 19 August 2025 / Accepted: 20 November 2025

Published online: 24 November 2025

References

- Skouri H, Girerd N, Monzo L, Petrie MC, Böhm M, Adamoet M, et al. Clinical management and therapeutic optimization of patients with heart failure with reduced ejection fraction and low blood pressure. A clinical consensus statement of the heart failure association (HFA) of the ESC. *Eur J Heart Fail*. 2025;27(4):707–22. <https://doi.org/10.1002/ehf.3618>.
- Savarese G, Becher PM, Lund LH, Seferovic P, Rosano GMC, Coats AJS. Global burden of heart failure: A comprehensive and updated review of epidemiology. *Cardiovasc Res*. 2023;118(17):3272–87. <https://doi.org/10.1093/cvr/cvac013>.
- Lippi G, Sanchis-Gomar F. Global epidemiology and future trends of heart failure. *AME Med J*. 2020;5:1–6. <https://doi.org/10.21037/amj.2020.03.03>.
- Heidenreich PA, Bozkurt B, Aguilar D, Allen LA, Byun JJ, Colvin MM, et al. 2022 AHA/ACC/HFSA guideline for the management of heart failure: A report of the American college of Cardiology/American heart association joint committee on clinical practice guidelines. *Circulation*. 2022;145(18):e895–1032. <https://doi.org/10.1161/CIR.0000000000001063>.
- Buck HG, Howland C, Stawnychy MA, Aldossary H, Cortés YI, DeBerg J, et al. Caregivers' contributions to heart failure self-care: an updated systematic review. *J Cardiovasc Nurs*. 2024;39(3):266–78. <https://doi.org/10.1097/JCN.0000000000001060>.
- Jering K, Claggett B, Redfield MM, Shah SJ, Anand IS, Martinez F, et al. Burden of heart failure signs and symptoms, prognosis, and response to therapy: the PARAGON-HF trial. *JACC Heart Fail*. 2021;9(5):386–97. <https://doi.org/10.1016/j.jchf.2021.01.011>.
- Kitko L, McIlvennan CK, Bidwell JT, Dionne-Odom JN, Dunlay SM, Lewis LM, et al. Family caregiving for individuals with heart failure: A scientific statement from the American heart association. *Circulation*. 2020;141(22):e864–78. <https://doi.org/10.1161/CIR.0000000000000768>.
- Suksatan W, Tankumpuan T, Davidson PM. Heart failure caregiver burden and outcomes: A systematic review. *J Prim Care Community Health*. 2022;13:1–19. <https://doi.org/10.1177/2150131922112584>.
- Sala-González M, Pérez-Jover V, Guilabert M, Mira JJ. Mobile apps for helping informal caregivers: A systematic review. *Int J Environ Res Public Health*. 2021;18(4):1–18. <https://doi.org/10.3390/ijerph18041702>.
- Lilleheie I, Debesay J, Bye A, Bergland A. The tension between carrying a burden and feeling like a burden: a qualitative study of informal caregivers' and care recipients' experiences after patient discharge from hospital. *Int J Qual Stud Health Well-being*. 2021;16(1):1–14. <https://doi.org/10.1080/17482631.2020.1855751>.
- Malki ST, Johansson P, Andersson G, Andréasson F, Mourad G. Caregiver burden, psychological well-being, and support needs among Swedish informal caregivers. *BMC Public Health*. 2025;25(1):1–11. <https://doi.org/10.1186/s12889-025-22074-y>.
- Zarit SH, Reever KE, Bach-Peterson J. Relatives of the impaired elderly: correlates of feelings of burden. *Gerontologist*. 1980;20(6):649–55. <https://doi.org/10.1093/geront/20.6.649>.
- Novak M, Guest C. Application of a multidimensional caregiver burden inventory. *Gerontologist*. 1989;29(6):798–803. <https://doi.org/10.1093/geront/29.6.798>.
- Elmståhl S, Malmberg B, Annerstedt L. Caregiver's burden of patients 3 years after stroke assessed by a novel caregiver burden scale. *Arch Phys Med Rehabil*. 1996;77(2):177–82. [https://doi.org/10.1016/s0003-9993\(96\)90164-1](https://doi.org/10.1016/s0003-9993(96)90164-1).
- Robinson BC. Validation of a caregiver strain index. *J Gerontol*. 1983;38(3):344–8. <https://doi.org/10.1093/geronj/38.3.344>.
- Strömberg A, Bonner N, Grant L, Bennett B, Chung ML, Jaarsma T, et al. Psychometric validation of the heart failure caregiver questionnaire (HF-CQ®). *Patient*. 2017;10(5):579–92. <https://doi.org/10.1007/s40271-017-0228-x>.
- Lee CJ, Lee H, Yoon M, Chun KH, Kong MG, Jung MH, et al. Heart failure statistics 2024 update: A report from the Korean society of heart failure. *Int J Heart Fail*. 2024;6(2):56–69. <https://doi.org/10.36628/ijhf.2024.0010>.
- Polit DF, Yang FM. Measurement and the measurement of change. Philadelphia (PA): Wolters Kluwer; 2015.
- Han C, Jo SA, Kwak JH, Pae CU, Steffens D, Jo I, et al. Validation of the patient health questionnaire-9 Korean version in the elderly population: the Ansan geriatric study. *Compr Psychiatry*. 2008;49(2):218–23. <https://doi.org/10.1016/j.comppsych.2007.08.006>.
- Spitzer RL, Kroenke K, Williams JB. Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. Primary care evaluation of mental disorders. Patient health questionnaire. *JAMA*. 1999;282(18):1737–44. <https://doi.org/10.1001/jama.282.18.1737>.
- Kim TH, Jo MW, Lee SI, Kim SH, Chung SM. Psychometric properties of the EQ-5D-5L in the general population of South Korea. *Qual Life Res*. 2013;22(8):2245–53. <https://doi.org/10.1007/s11136-012-0331-3>.
- World Health Organization. Process of translation and adaptation of instruments [Internet]. Geneva: World Health Organization. 2017 [cited 2025 Jun 28]. Available from: <https://www.scribd.com/document/533869240/WHO-Guidelines-on-Translation-and-Adaptation-of-Instruments>
- Zamanzadeh V, Ghahramanian A, Rassouli M, Abbaszadeh A, Alavi-Majd H, Nikanfar AR. Design and implementation content validity study:

- development of an instrument for measuring patient-centered communication. *J Caring Sci*. 2015;4(2):165–78. <https://doi.org/10.15171/jcs.2015.017>.
24. Mishra P, Pandey CM, Singh U, Gupta A, Sahu C, Keshri A. Descriptive statistics and normality tests for statistical data. *Ann Card Anaesth*. 2019;22(1):67–72. https://doi.org/10.4103/aca.ACA_157_18.
 25. Tabachnick BG, Fidell LS, Ullman JB. Using multivariate statistics. 6th ed. Boston, MA: Pearson; 2013. pp. 497–516.
 26. Peugh JL, Litson K, Feldon DF. Equivalence testing to judge model fit: A Monte Carlo simulation. *Psychol Methods*. 2025;30(4):888–925. <https://doi.org/10.1037/met0000591>.
 27. Hair JF Jr, Black WC, Babin BJ, Anderson RE. Multivariate data analysis. 8th ed. Boston, MA: Cengage; 2019.
 28. Henseler J, Ringle CM, Sarstedt M. A new criterion for assessing discriminant validity in variance-based structural equation modeling. *J Acad Mark Sci*. 2015;43:115–35. <https://doi.org/10.1007/s11747-014-0403-8>.
 29. Cheung A, Kuti EL, Osenenko KM, Friesen M, Donato BMK. Impact of caring for individuals with heart failure in the United States: a systematic literature review. *J Cardiovasc Nurs*. 2024;39(2):128–41. <https://doi.org/10.1097/JCN.0000000000001005>.
 30. Abdelhalim DS, Ahmed MM, Hussein HA, Khalaf OO, Sarhan MD. Burden of care, depression, and anxiety among family caregivers of people with dementia. *J Prim Care Community Health*. 2024;15:21501319241288029. <https://doi.org/10.1177/21501319241288029>.
 31. Bozkurt Zincir S, Sunbul M, Zincir S, et al. Burden and depressive symptoms associated with adult-child caregiving for individuals with heart failure. *Sci World J*. 2014;2014:641817. <https://doi.org/10.1155/2014/641817>.
 32. Yao H, Li K, Li C, Hu S, Huang Z, Chen J, et al. Caregiving burden, depression, and anxiety in informal caregivers of people with mental illness in china: A cross-sectional survey. *BMC Psychiatry*. 2024;24(1):824. <https://doi.org/10.1186/s12888-024-06239-4>.
 33. Valcárcel-Nazco C, Ramallo-Fariña Y, Linertová R, Ramos-Goñi JM, García-Pérez L, Serrano-Aguilar P. Health-related quality of life and perceived burden of informal caregivers of patients with rare diseases in selected European countries. *Int J Environ Res Public Health*. 2022;19(13):8208. <https://doi.org/10.3390/ijerph19138208>.
 34. Bidwell JT, Vellone E, Lyons KS, D'Agostino F, Riegel B, Paturzo M, et al. Caregiver determinants of patient clinical event risk in heart failure. *Eur J Cardiovasc Nurs*. 2017;16(8):707–14. <https://doi.org/10.1177/1474515117711305>.
 35. Song DY. A study on the predictors of double burden of caregiving among women carrying both childcare and care for the elderly: focusing on East-Asia country cases. *Family Cult*. 2017;29(2):95–125. <https://doi.org/10.21478/family.29.2.201706.004>.

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.