# RESEARCH

# **Open Access**



# Psychometric validation and cultural adaptation of the Chinese version of the CarerQol-7D instrument

Richard Huan Xu<sup>1\*</sup>, Yuanshuo Xu<sup>1</sup>, Mengping Zhao<sup>2</sup>, Nan Liu<sup>2</sup>, Pei Wang<sup>3</sup>, Xiao Liang<sup>1</sup> and Lingming Zhou<sup>4\*</sup>

# Abstract

**Objective** This study aims to translate and validate the CarerQol in a sample of informal caregivers in China.

**Method** The CarerQol was translated following a standard process. Validation data was collected from multiple sources, including hospitals, patient associations, and community health centers between September 2023 and December 2024 in China. Caregivers of individuals with disability or long-term care needs were recruited. The psychometric properties of the CarerQol-7D were evaluated through ceiling and floor effects, factorial structure, convergent and divergent validity, known-groups validity, and test-retest reliability.

**Results** A total of 324 participants completed the survey and provided validated responses. The mean CarerQol level sum score was 12.8 (SD = 2.9). The CarerQol-7D showed no ceiling or floor effects at either dimension or scale levels. Test-retest reliability of the CarerQol-7D was satisfactory, as supported by ICC and Gwet's AC1 measures. Correlation analysis confirmed all hypothesized pairs of dimensions and level sum scores/utility scores between CarerQol-7D, SF-6Dv2, and ICECAP-A. The CarerQol-7D successfully distinguished between risk groups through ANOVA testing. Confirmatory factor analysis revealed that a one-factor model of the CarerQol-7D fit the data.

**Conclusion** This study provides the Chinese version of the CarerQol and confirms its sound psychometric properties for informal caregivers in China. These findings demonstrate the CarerQol's value in cost-effectiveness analyses of caregiving interventions and strategies.

Keywords CarerQol, Psychometric properties, Translation, Informal caregiver

Richard Huan Xu richard.xu@polyu.edu.hk Lingming Zhou zhoulingming@gzhmu.edu.cn <sup>1</sup>Department of Rehabilitation Sciences, Faculty of Health and Social Sciences, Hong Kong Polytechnic University, Kowloon, Hong Kong <sup>2</sup>School of Public Health, Lanzhou University, Lanzhou, China <sup>3</sup>School of Public Health, Fudan University, Shanghai, China <sup>4</sup>The Second Affiliated Hospital, Guangzhou Medical University, Guangzhou, China

# Introduction

Informal caregivers, who is the individual, typically a family member or friend, who provides unpaid care and support to a person with a chronic illness, disability, or other dependency needs, often managing physical, emotional, and daily living tasks [1], are essential to care recipients and are crucial components of the healthcare system [2]. Socioeconomic changes and rising healthcare costs have created a huge global need for informal caregivers. The World Health Organization reports that approximately 1.3 billion people need care due to aging or disability. Today, the role of informal cancer caregivers



\*Correspondence:

© 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://creative.commons.org/licenses/by-nc-nd/4.0/.

has evolved beyond basic support. They are now the primary source of care, managing medications and providing emotional support, tasks once handled by healthcare professionals [3, 4]. Research has shown that patients with engaged caregivers have better treatment adherence, improved quality of life, and better emotional wellbeing [5, 6]. However, intensive caregiving has adverse consequences: caregivers often face reduced employment opportunities, high poverty rates, and increased mental health problems [7–9]. In response, many countries in the Organization for Economic Co-operation and Development have put in place policies to support family caregivers and reduce these negative effects.

While informal caregiving is often associated with challenges such as emotional strain and physical demands, it is equally important to recognize its potential positive impacts on caregivers. Providing informal care may foster a sense of purpose, fulfillment, and strengthened familial bonds, contributing to caregivers' psychological wellbeing. For instance, many caregivers report feelings of satisfaction and pride in supporting loved ones, which can enhance their self-esteem and emotional resilience [10]. Additionally, caregiving often deepens interpersonal relationships, offering opportunities for reciprocity and mutual support that enrich caregivers' social lives [11, 12]. These positive dimensions may counterbalance the burdens and highlight the multifaceted nature of the caregiving experience, indicating the need to adopt a more holistic perspective in assessing its effects on informal caregivers.

Considering informal care in economic evaluations is important because it highlights their significant contributions to both the healthcare system and society as a whole. By quantifying the value of caregiving, policymakers can allocate resources more efficiently, ensuring that caregivers receive the support they need through financial compensation, tax benefits, and access to essential services [13–15]. This evaluation also underscores the role of caregivers in reducing healthcare costs by preventing or delaying the need for more expensive institutional care. Furthermore, recognizing the economic impact of caregiving can lead to more supportive workplace policies, helping caregivers balance their responsibilities with their employment [16]. Ultimately, understanding the economic and social value of caregiving leads to better policy formulation and resource allocation, thereby enhancing the well-being of both caregivers and care recipients [17].

Traditional methods of health economic evaluation tend to focus on a limited interpretation of health, with the aim of quantifying outcomes in terms of qualityadjusted life years (QALYs) [18]. However, many interventions, particularly in public health, end-of-life care, and social care, have effects that go beyond this narrow perspective of health. Research has highlighted the need for assessment beyond what is possible with basic conventional instruments, such as the EQ-5D, and to evaluate the impacts of interventions that include outcomes outside the QALY framework [19, 20]. This is crucial when conducting an economic evaluation that informs decision-making on resource allocation in care-related interventions.

Preference-based measures are essential for supporting the economic evaluation of care-related policies and interventions. These instruments, widely used in health economics, quantify the value that individuals or society place on specific health states, expressed as utility scores (0 = death, 1 = perfect health). Unlike descriptive measures, they reflect preferences facilitating the calculation of QALYs for cost-effectiveness analyses in healthcare. In the context of caregiving, the Care-related Quality of Life (CarerQol) instrument has been developed to assess the impact of health and social care interventions on informal caregivers within an economic evaluation framework [21]. The CarerQol instrument can be used to measure and assess the subjective burden of informal care, or be included in the "effect side" of multi-criteria or cost-consequence analyses of interventions for patients or older people. The CarerQol instrument can also be used in cost-utility analyses of interventions for informal carers.

The CarerQol instrument consists of two parts: a seven-item descriptive system (CarerQol-7D) and a visual analogue scale for well-being (CarerQol-VAS). The CarerQol-7D aims to capture subjective burden from the caregiver's perspective and has preference weights (numerical values assigned to specific health states to reflect their perceived desirability or utility) available for several Western countries [17, 22]. It includes five negative and two positive dimensions of informal care. The negative dimensions are relational problems, mental health problems, problems combining daily activities with care, financial problems, and physical health problems due to informal care. The two positive dimensions are fulfillment from caregiving and support with lending care. For each dimension, there are three possible answers: no, a little, and a lot. The CarerQol-VAS ranges from 0 (completely unhappy) to 10 (completely happy), allowing caregivers to indicate how happy they feel. According to an announcement from iMTA, the developer of the CarerQol, the instrument is currently available in 11 languages: English, Dutch, German, Hungarian, Italian, Japanese, Norwegian, Portuguese, Slovenian, Spanish, and Swedish. Translations can be accessed on iMTA's official website [23]. However, there is no Chinese version yet.

In China, informal caregivers face a particularly heavy burden due to traditional cultural norms, coupled with the rapidly aging population and the legacy of the one-child policy, which has intensified pressure on an increasingly small pool of potential caregivers [24]. Additionally, the healthcare system in China is still developing comprehensive support structures for informal caregivers, resulting in limited access to resources such as respite care, training, and financial assistance [25]. This lack of support compounds their burden. It is therefore crucial to develop a valid preference-based measure for policymakers to recognize and address the needs of informal caregivers to ensure effective resource allocation to improve the quality of life and well-being of informal caregivers. Therefore, this study aims to translate and validate the CarerQol involving a sample of informal caregivers in China.

## Method

#### Translation and cultural adaption of the CarerQol

The translation of the CarerQol followed the protocol by the ISPOR Task Force for Translation and Cultural Adaptation [26]. Two native Chinese speakers fluent in English independently translated the CarerQol from English to Chinese. The research team collected and synthesized these translations. A revised version was then sent to two professional translators for independent back-translation. The research team compared the back-translation with the original English version, identified discrepancies, addressed disputed items, and refined the translation with attention to linguistic accuracy and cultural relevance until a consensus was achieved. To ensure the cultural appropriateness, 10 members of the general Chinese population, representing diverse demographic backgrounds (e.g., age, education, and caregiving experience), participated in a cognitive debriefing session. During this process, participants reviewed the response options and flagged challenging words, phrases, or concepts that felt unclear or misaligned with local norms. For instance, terms related to emotional burden or caregiving support, which may carry different connotations in Chinese cultural contexts, were discussed in depth; participants interpreted their understanding of these terms in their own words, highlighting nuances such as familial duty or social expectations that differ from Western perspectives. The research team analyzed this feedback to identify areas requiring cultural adaptation. These adaptations ensured that the CarerQol remained conceptually equivalent to the original while being sensitive to Chinese cultural norms and caregiving experiences. After integrating these insights and conducting a final team review, the culturally adapted Chinese translation of the CarerQol was finalized.

#### Data source

This study collected data from multiple sources, particularly from informal caregivers of individuals with various disabilities, illnesses, and conditions. The inclusion criteria for the participants were as follows: (1) being  $\geq 18$  years old; (2) having been a primary caregiver for  $\geq 5$  years; (3) having no physical or mental health disability; and (4) being able to provide informed consent. We employed a convenience sampling approach to recruit caregivers of both adult and child recipients residing in diverse geographic regions across China, including the South, East, and West. These regions were selected to reflect varying socioeconomic contexts and caregiving dynamics.

We recruited caregivers of patients with attentiondeficit/hyperactivity disorder (ADHD) from the center for early child development at Gansu maternity and child health care hospital in Lanzhou. The data collection process is detailed in our previous paper [27]. Data were also collected from various sources in two Chinese provinces in November 2024, including the Shenzhen Shixia community service center, Shenzhen Autism Society, Guangzhou Rongai Home (individuals with intellectual disabilities), Zhejiang Prader-Willi Syndrome care & support center, and Zhejiang Xinchang people's hospital. The research team has established long-term collaborations with these institutions. For the community center and patient associations, we contacted the person in charge and provided them with the project details. The person in charge then distributed the invitation link to eligible members through their internal social network. Interested members were invited to click on the link to provide basic background information and complete the survey. For the Zhejiang Xinchang people's hospital, the questionnaire was distributed through designated nurses from the geriatric, rehabilitation, respiratory, kidney, and oncology departments. The nurses reviewed their patients' medical records and invited caregivers of eligible patients to scan a QR code to join the chat group. A research assistant double-checked all participants' background characteristics and retained only eligible participants. The participants were asked to complete the questionnaire via a separate link shared in the chat group. The questionnaire began with an informed consent form, which the participants had to sign electronically before proceeding with the survey. To assess test-retest reliability, we invited all of the participants to complete the CarerQol again one week later using the same method. The Institutional Review Board of the Hong Kong Polytechnic University approved the research protocol (Ref: HSEARS20230314005 and HSEARS20241128005). All of the participants provided written informed consent.

# Sample size

A sample size of 324 participants is sufficient to assess the psychometric properties of the CarerQoL-7D. Methodological guidelines for psychometric evaluations, such as the COSMIN checklist, recommend a minimum of 100 participants for internal consistency and structural validity analyses, with larger samples ( $\geq$  300) preferred for stable factor analysis and subgroup comparisons [28]. Our sample exceeds these thresholds and aligns with precedent studies validating caregiver-specific instruments (e.g., n = 200-400) [21, 29].

#### Instruments

#### CarerQol

The Chinese version of the CarerQol was used in this study. Because no value set for the CarerQol is available for the Chinese population, we calculated and reported the level sum scores (LSSs) of the CarerQol-7D. Additionally, UK value set was used to calculate the utility score fo CarerQoL to support convergent validity test [22].

#### SF-6Dv2

The SF-6Dv2 is a generic preference-weighted measure of health-related quality of life consisting of six dimensions (physical functioning, role limitation, social functioning, pain, mental health, and vitality) [30]. The participants were asked to indicate their health status over the past four weeks on all six dimensions. Each dimension was rated using five response levels, except for the pain dimension, which had six response levels. The psychometric properties of the SF-6Dv2 in Chinese population is confirmed [31]. To assess the convergent validity of the CarerQol-7D utility score, we employed the UK value set for the SF-6Dv2, which ranges from -0.574 to 1 [32], we calculate both the LSSs and index values in this study.

#### ICEpop capability measure for adults (ICECAP-A)

The ICECAP-A is a measure that assesses an adult's current well-being based on Sen's capability theory [33]. It includes five attributes, namely stability, attachment, autonomy, achievement, and enjoyment, each with four response levels. The ICECAP-A has been previously translated into Chinese, cross-culturally adapted, and tested in the Chinese population [34]. However, given the absence of a Chinese preference weights, we calculated both LSSs and index values using UK value set [35].

#### **Overall health status**

A single-item question was used to assess caregivers' selfreported overall health status. Participants were asked to rate their health over the past week using a 5-point Likert-type scale, with response options ranging from "very poor" to "very good."

#### Data analysis

Descriptive statistics were used to describe the participants' background characteristics and measurement profiles. Continuous variables are presented as means with standard deviations (SD), while categorical variables are presented as frequencies (n) and proportions (%). All analyses were performed using R (R foundation, Vienna, Austria).

# **Ceiling and floor effects**

Ceiling and floor effects were examined at both the item level and the scale level. At the item level, the proportions of participants reporting the highest (ceiling) and lowest (floor) response levels were determined. At the scale level, we computed the relative frequencies of participants with the best (ceiling) and worst (floor) profiles (e.g., for the CarerQol-7D, "1111111" and "3333333," respectively). A threshold of 70% was used at the item level [36, 37], while a 15% threshold was applied at the scale level [38].

#### Convergent and divergent validity

Convergent and discriminant validity were tested to assess the strength of the relationships between items aiming to measure similar constructs, as well as between different instruments. Divergent validity was used to determine whether items in the CarerQol-7D captured aspects not covered in other measures. We used Spearman's rank-order correlations between items (rho  $\leq 0.29 =$  weak; rho  $\leq 0.49 =$  moderate; rho  $\geq 0.5 =$  strong) [39]. At the instrument level, Pearson correlations ( $r \leq 0.29 =$  weak;  $r \leq 0.49 =$  moderate;  $r \geq 0.5 =$  strong) were used for index values and LSSs [39].

We selected the SF-6Dv2 (HRQoL) and ICECAP-A (capability well-being) to validate the CarerQoL-7D's convergent validity, as they collectively reflect both health-specific and holistic well-being dimensions relevant to caregivers. Their complementary frameworks allow us to test whether the CarerQoL-7D captures both caregiving-specific burdens (aligned with SF-6Dv2) and broader life impacts (aligned with ICECAP-A). We hypothesized at least moderate correlations between conceptually overlapping items, such as "I have problems with my own physical health" in the CarerQol-7D and "physical functioning" in the SF-6DV2; "I have problems with my own mental health" in the CarerQol-7D and "mental health" in the SF-6Dv2; and "I have support to carry out my care tasks, when I need it" in the CarerQol-7D and "love, friendship, and support" in the ICECAP-A. Weak or no correlations were expected between nonoverlapping items. At the instrument level, we expected moderate correlations between CarerQol-7D, SF-6Dv2, and ICECAP-A scores or index values.

#### Known-groups validity

Known-groups validity tests assessed the ability of the CarerQol-7D to distinguish between participant groups.

The known groups were defined by caregiver's age (<35 years, 35–44 years, and  $\geq$ 45 years), caregiver's health status (poor, average, and good), educational level (primary or below, secondary, and tertiary or above), family registry (urban and rural), and perceived socioeconomic status (lower than local average, equal to local average, higher than local average). ANOVA was used to compare the mean differences in CarerQol-7D LSSs between the known groups.

#### Table 1 Participants' demographics (n = 324)

	n	%
Caregiver		
Sex		
Male	59	18.2
Female	265	81.8
Educational level		
Primary or below	11	3.4
Low secondary	42	13.0
High secondary	85	26.2
Tertiary or above	186	57.4
Employed status		
Fully employed	232	71.6
Unemployed	60	18.5
Farming	17	5.2
Retired	15	4.6
Marital status		
Single	22	6.7
Married	283	87.3
Divorced/widow(er)	19	6.0
Family registry		
Urban	213	65.7
Rural	111	34.3
Chronic disease		
Yes	36	11.1
No	288	88.9
Overall health status		
Very poor	27	8.3
Poor	9	2.8
Average	165	50.9
Good	102	31.5
Very good	21	6.5
Relationship with care receiver		
Parents	272	83.9
Grandparents	7	2.2
Children	34	10.5
Others	11	3.4
Perceived socioeconomic status		
Lower than local average	51	15.7
Equal to local average	222	68.6
Higher than local average	51	15.7
Age (mean, SD)	36.2 (9.8)	

#### Page 5 of 11

#### Factor structure

Confirmatory factor analysis (CFA) was used to examine the one-factor model of the CarerQol-7D. Model fit was evaluated by checking the comparative fit index (CFI, > 0.9), Tucker–Lewis index (TLI, > 0.9), root mean square error of approximation (RMSEA, < 0.08), and standardized root mean square residual (SRMR, < 0.08). A factor loading > 0.3 was considered acceptable [40].

#### Test-retest reliability

Only caregivers of children with ADHD who completed both assessments (e.g., baseline and follow-up) were included in the test-retest reliability analysis. For the test-retest reliability of the CarerQol-7D, we used Gwet's AC1 to evaluate item-level agreement (0.21–0.4 = fair agreement, 0.41-0.6 = moderate agreement,  $\ge 0.61$  = strong agreement) [41]. For test-retest reliability at the scale level, we used the internal consistency coefficient (ICC, two-way random effects, absolute agreement, with  $\ge 0.7$  considered acceptable) [42].

## Results

#### Participant demographics

The survey received valid responses from 324 eligible caregivers (Table 1). Among them, 81.8% were women with a mean age of 36.2 years, and 57.4% had completed tertiary education or higher. The background characteristics of care receivers are presented in Appendix.

## **Ceiling and floor effects**

None of the dimensions showed significant ceiling or floor effects. Specifically, 41.4% of the participants reported no problems with daily activities and financial matters, while 20.4% of the participants reported significant mental health issues. Among the participants, only 0.9% (n=3) reported the highest possible care status on the CarerQol-7D. The mean LSS on the CarerQol-7D was 12.8 (SD = 2.9), ranging from 7 to 21 (Table 2).

#### **Test-retest reliability**

The data from 155 (48%) participants who indicated "unchanged" health status in the second survey were used to assess test-retest reliability (Table 2). Gwet's AC1 showed that with the exception of "relational problems" and "support when needed" (moderate agreement), all other dimensions of the CarerQol-7D showed strong agreement between the two surveys. The ICC (value = 0.71) confirmed good test-retest reliability for the CareQol-7D at the scale level.

## Convergent and divergent validity

Tables 3 and 4 report the convergent and divergent validity of the CarerQol-7D, respectively. The CarerQol-7D dimensions were significantly correlated with most of the

CarerQol-7D	Proportion (%)			Mean	SD	Median	Range	Gwet's AC1/ ICC
	No problems	Some problems	A lot of problems	-				( <i>n</i> = 155)
Fulfilment*	22.8	57.1	20.1	-	-	-	-	0.64
Relational problems	37.3	48.5	14.2	-	-	-	-	0.55
Mental health	28.4	51.2	20.4	-	-	-	-	0.62
Daily activities	41.4	46.0	12.7	-	-	-	-	0.62
Financial problems	41.4	39.8	18.8	-	-	-	-	0.73
Support when need*	26.9	54.6	18.5	-	-	-	-	0.58
Physical health	37.3	44.8	17.9	-	-	-	-	0.68
Best caring status (1111111)	0.9	-	-	-	-	-	-	-
Level sum score	-	-	-	12.8	2.9	13	7–21	0.71
CarerQoI-VAS	-	-	-	6.2	2.1	7	0-10	-

#### Table 2 Profile of CarerQol-7D and test-retest reliability

\*Reversed scoring

Table 3 Correlations of dimensions between CarerQol-7D, SF-6Dv2, and ICECAP-A

CarerQol-7D							
SF-6Dv2	Fulfilment	Relational problems	Mental health	Daily activities	Financial problems	Support when need	Physical health
Physical functioning	-0.05	0.19	0.15	0.30	0.38	0.06	0.36
Role limitation	0.03	0.17	0.25	0.30	0.34	0.08	0.38
Social functioning	0.09	0.23	0.28	0.38	0.41	0.17	0.39
Pain	0.03	0.25	0.36	0.40	0.39	0.05	0.45
Mental health	0.17	0.24	0.50	0.40	0.38	0.13	0.45
Vitality	0.14	0.22	0.36	0.40	0.37	0.06	0.40
ICECAP-A							
Stability	0.27	0.21	0.39	0.40	0.42	0.16	0.40
Attachment	0.19	0.20	0.38	0.34	0.33	0.37	0.38
Autonomy	0.20	0.15	0.36	0.33	0.32	0.25	0.33
Achievement	0.20	0.13	0.32	0.33	0.31	0.20	0.28
Enjoyment	0.14	0.15	0.24	0.31	0.31	0.14	0.22
CarerQoI-VAS	-0.42	-0.20	-0.40	-0.37	-0.34	-0.27	-0.36

Table 4 Correlations of utility score/level sum score between CarerQol-7D, SF-6Dv2, and ICECAP-A

Instrument		Instrument	Correlation coefficient (95% Confidence Interval)
CarerQol-7D UK utility score	~	SF-6Dv2 UK utility score	0.50 (0.41, 0.58)
CarerQol-7D UK utility score	~	ICECAP-A utility score	0.40 (0.30, 0.48)
CarerQol-7D UK utility score	~	CarerQoI-VAS	0.27(0.16, 0.36)
CarerQol-7D level sum score	~	SF-6Dv2 level sum score	0.58 (0.50, 0.65)
CarerQol-7D level sum score	~	ICECAP-A level sum score	0.59 (0.51, 0.65)
CarerQol-7D level sum score	~	CarerQoI-VAS	-0.56 (-0.63, -0.48)

hypothesized dimensions of the SF-6Dv2 and ICECAP-A. The "mental health" dimension of the CarerQol-7D showed a strong association with the "mental health" dimension of the SF-6Dv2, while the "physical health" dimension of the CarerQol-7D showed a moderate association with the "physical functioning" dimension of the SF-6Dv2. The "support when needed" dimension of the CarerQol-7D showed a moderate association with the "attachment" dimension of the ICECAP-A. Finally, the LSS of the CarerQol-7D demonstrated strong associations with both the utility and LSSs of the SF-6Dv2 and ICECAP-A.

#### **Known-groups validity**

Table 5 reports the sensitivity of the CarerQol-7D in known-groups comparisons and its ability to distinguish between risk groups. The analysis revealed statistically significant differences across all known groups. Specifically, the participants who had good health, higher education levels, and lived in urban areas reported better care-related quality of life than their counterparts.

# **Factorial structure**

The CFA results supported the one-factor model of the CarerQol-7D, with CFI and TLI values greater than 0.9

#### Table 5 Known-group validity of CarerQol-7D

	Mean (SD)	P-value
Caregiver's age		
< 35	12.5(2.8)	0.002
35–44	13.0(2.9)	
≥45	13.8(2.8)	
Caregiver's health status		
Poor	15.9(2.6)	< 0.01
Average	13.4(2.6)	
Good	11.3(2.3)	
Educational level		
Primary or below	15.5(3.4)	< 0.001
Secondary	13.5(2.8)	
Tertiary or above	12.3(2.7)	
Family registry		
Urban	12.4(2.7)	< 0.001
Rural	13.8(2.9)	
Perceived socioeconomic status		
Lower than local average	15.4(2.7)	< 0.001
Equal to local average	12.5(2.7)	
Higher than local average	11.9(2.6)	

and RMSEA and SRMR values less than 0.05 (Fig. 1).

# Model fit indicators:

RMSEA = 0.07 CFI = 0.951 TLI =0.927 SRMR =0.05 However, the standardized factor loading of the "support" item was statistically significant and below the acceptable threshold (>0.3). After removing this item and rerunning the CFA, the model's performance showed only minimal improvement rather than substantial enhancement.

# Discussion

The findings of this study demonstrate that the Chinese version of the CarerQol-7D is a valid and reliable instrument for measuring the impact of caregiving for people with long-term disabilities on informal caregivers in China. To the best of our knowledge, this study is the first to evaluate a preference-based measure that is culturally adapted to the Chinese population and exhibits satisfactory psychometric properties to support cost–utility analysis of care-related interventions, policies, and strategies.

We observed that the correlations between the CarerQol-7D and the CarerQol-VAS, SF-6Dv2, and ICE-CAP-A were statistically significant and in the expected direction, aligning with prior studies [29, 43–45]. These findings highlight shared constructs of wellbeing and autonomy in caregiving, supporting the convergent



validity of CarerQol-7D in this sample. However, some dimensional associations between measures showed less strength than expected. For example, the "relational problems" dimension of the CarerQol-7D showed a weak association with the "attachment" dimension of the ICE-CAP-A. The reason might be that the "relational problems" dimension specifically addresses how caregiving affects relationships, which may include stress or strain due to caregiving responsibilities. In contrast, "attachment" is broader and focuses on a person's general ability to form and maintain close relationships, which may not be directly affected by caregiving in the same way [46]. The convergent validity of the CarerQol-7D at the dimension level requires further assessment.

As the factor structure of the CarerQol-7D has never been tested in an Asian population, we conducted this analysis. Our CFA results supported the one-factor model of the CarerQol-7D. However, two items requiring reverse scoring, namely "fulfilment" and "support," showed low factor loadings. We revised the model by removing items with low factor loadings individually and together, but these changes did not significantly improve model performance. While exploratory factor analysis suggested a two-factor model structure, CFA of this two-factor structure ("fulfilment" and "support" as factor 2) revealed consistently low and nonsignificant factor loadings, indicating that this modification was not meaningful. It is difficult to determine whether this is a language or technical issue, as few previous studies have reported on the factor structure of the CarerQol-7D, but most of them supported the it is a unidimensional tool by several studies for utility purposes [13, 29, 45]. However, previous studies examining the psychometric properties of this instrument in Chinese populations have shown similar problems with negatively worded items [47]. Therefore, we recommend revising the Chinese version of these two items to align their logic with the other five items and to test their performance in future studies.

In our study, we found that two CarerQol-7D items exhibited ceiling effects slightly exceeding 40%. These results align with prior research. For instance, in a Dutch study, 64% and 71% of participants reported 'no problems' with daily activities and financial problems, respectively [13]. Similarly, a study across eight European countries reported 63-71% and 68-75% with 'no problems' on these dimensions [29]. Another study spanning Australia, the USA, Germany, and the UK showed ceiling effects for these items ranging from 50 to 75% [22]. To determine whether these findings indicate a problematic ceiling effect, we considered the context of preferencebased instruments and drew on the EQ-HWB studies as a comparative benchmark. Unlike generic HRQoL measures like the EQ-5D, which often aim to discriminate across a broad range of health states in general populations (where ceiling effects above 15-30% may signal limited sensitivity), the CarerQol targets a specific group, caregivers, whose experiences may naturally skew toward extreme ends of the scale (e.g., high burden or high satisfaction). The EQ-HWB studies' use of a 70% threshold suggests that, for instruments tailored to such populations, higher proportions of extreme responses may not necessarily undermine the tool's validity or utility, provided the items remain meaningful and preference weights (e.g., utility scores) can still be derived effectively.

The CarerQol-7D successfully distinguished between caregiver subgroups based on several relevant characteristics. In line with previous findings, older caregivers and those with poorer health status tended to report higher LSSs on the CarerQoL-7D [44, 48]. Additionally, caregivers with higher education levels were more likely to report higher LSSs on the CarerQol-7D, which is consistent with an Iranian study [49] but differs from results from European countries [29]. One advantage of our known-groups validity is that our care recipients represented a more diverse sample than that used in previous studies [29, 49], with different types of disabilities across various age groups. This diversity enhances the generalizability of our findings. This generally supports the known-groups validity of the CarerQol-7D.

The CarerQol-7D exhibited strong test-retest reliability in our study, with an ICC of 0.71 over a 1-week interval, aligning with findings from a newly published systematic review, which reported an average ICC of 0.62 across three studies [1]. Our result closely matches Hoefman et al. (ICC = 0.75, 2 weeks) [50] and exceeds McCaffrey et al. (ICC=0.67, assumed 2-4 weeks) [45], supporting the instrument's suitability for longitudinal research to track changes over time. For the 'financial problems' dimension, we observed excellent test-retest stability, likely due to the short 1-week interval and a high ceiling effect (41.4% reported 'no problems'), consistent with Hoefman et al., where stable financial circumstances contributed to reliability [50]. In contrast, the 'relational problems' dimension showed lower agreement between tests, a pattern also noted before [50], possibly reflecting the dynamic nature of caregiver-care recipient relationships even within a week. A longer interval, such as the 24 weeks in Vluggen et al. (ICC=0.41) [51], might amplify such fluctuations, suggesting a need for further investigation. A strength of our study is the evaluation of test-retest reliability at the scale level, with an ICC of 0.71 surpassing the meta-analytic average (0.62), affirming the CarerQol-7D's consistency over time and its reliability for both clinical assessments and research purposes in Chinese context.

The psychometric properties of the CarerQol-7D were robust across two subsamples, caregivers of children and caregivers of adults (Appendix), demonstrating its adaptability to diverse caregiving contexts. For children's caregivers, convergent validity was supported by moderate to strong correlations between LSS of the CarerQol-7D, SF-6Dv2, and ICECAP-A, though correlations with utility scores were slightly lower. Similar pattern of correlation was found for the adult's caregivers, possibly reflecting differences in perceived caregiving burden or cultural interpretations of wellbeing. Known-groups validity was confirmed in both subsamples, with significant LSS differences across age, health status, education, and socioeconomic status, aligning with the overall sample findings. However, the CarerQol-VAS showed a weaker correlation with LSS in adult's caregivers compared to children's caregivers, suggesting potential subgroup-specific variations in how subjective wellbeing aligns with caregiving experiences. No CFA was conducted for the adult's caregiver subsample due to its limited sample size (n=81), which was insufficient for reliable factorial analysis, which demonstrate the future evaluations in this population are needed, particularly for positive dimensions like fulfillment and support, which showed lower factor loadings in the overall and children's caregiver CFA.

Several limitations should be addressed. First, our data came from different sources, and although our data collection method was consistent, there could be problems in ensuring high data quality and reducing potential biases introduced by variations in how the data were recorded or reported across these sources. Second, our re-test data came only from the ADHD hospital sample. We lacked community samples to assess the test-retest reliability of the CarerQol-7D. Although this did not pose any methodological problems, it may have introduced selection bias into our analysis, affecting the interpretation of our results. Third, our survey only included caregivers who reported providing care for recipients for more than five years. While this criterion effectively captures the long-term effects of informal caregiving, it may overlook how caregiving impacts the quality of life and well-being of individuals with shorter caregiving durations. Consequently, the generalizability of our findings could be limited, as the experiences of caregivers newer to their roles remain unrepresented. Fourth, we recognize that the generalizability of findings may be limited to populations with similar characteristics to our sample. However, as this study focused on establishing psychometric properties (e.g., factor structure, internal consistency) rather than population-level inferences, representativeness was not the primary concern. Future research with probability sampling will be critical to validate these findings in broader populations. Finally, due to privacy concerns, we did not collect detailed information about care recipients, such as the duration of caregiving, the daily time spent providing informal care, or whether the caregiver lives with the recipient. These details are crucial for understanding the quality of life and wellbeing of caregivers. Consequently, this lack of data prevents us from assessing the known-groups validity of the CarerQol-7D comprehensively. This limitation may impact the instrument's applicability and warrants further investigation.

#### Conclusion

This study used a sample of informal caregivers of people with diverse disabilities to validate the Chinese version of the CarerQol-7D, confirming its satisfactory construct, convergent, and known-groups validity, and test–retest reliability. The findings indicate that the CarerQol instrument can be valuable for economic evaluation studies and help clinicians and policymakers make informed decisions about care-related interventions and resource allocation, ensuring that caregivers' needs are accurately identified and addressed.

#### Supplementary Information

The online version contains supplementary material available at https://doi.or g/10.1186/s12955-025-02379-7.

Supplementary Material 1

#### Author contributions

RHX: Conceptualization, Methodology, Software, Validation, Formal analysis, Data Curation, Supervision, Visualization, Writing - Original Draft, Writing -Review & Editing, YSX: Project administration, Data Curation, Writing - Review & Editing. MPZ: Project administration, Writing - Review & Editing. NL: Project administration, Writing - Review & Editing. PW: Conceptualization, Software, Writing - Review & Editing. XL: Conceptualization, Project administration, Supervision, Writing - Review & Editing. LMZ: Conceptualization, Formal analysis, Funding acquisition, Writing - Review & Editing.

#### Funding

This project is supported by the Guangdong Province Medical Research Fund (Grant No. B2022205) and the 2023 Plan on Enhancing Scientific Research in GMU.

#### Data availability

No datasets were generated or analysed during the current study.

#### Declarations

Conflict of interest None.

#### **Ethical approval**

This study was conducted in accordance with the ethical principles of the Declaration of Helsinki. The Institutional Review Board of Hong Kong Polytechnic University approved the research protocol (Ref: HSEARS20230314005 and HSEARS20241128005). All the participants provided written informed consent.

#### **Competing interests**

The authors declare no competing interests.

Received: 12 December 2024 / Accepted: 29 April 2025 Published online: 09 May 2025

- Cejalvo E, Martí-Vilar M, Gisbert-Pérez J, et al. The CarerQol instrument: A systematic review, validity analysis, and generalization reliability study. J Clin Med. 2025;14. https://doi.org/10.3390/jcm14061916.
- Chong E, Crowe L, Mentor K et al. Systematic review of caregiver burden, unmet needs and quality-of-life among informal caregivers of patients with pancreatic cancer. Support Care Cancer. 2023;31.
- Romito F, Goldzweig G, Cormio C, et al. Informal caregiving for cancer patients. Cancer. 2013;119. https://doi.org/10.1002/cncr.28057.
   Maleria H, Waa M, Haderberg G, and Angel G, ange
- Molassiotis A, Wang M. Understanding and supporting informal Cancer caregivers. Curr Treat Options Oncol. 2022;23. https://doi.org/10.1007/s1186 4-022-00955-3.
- Geerlings AD, Kapelle WM, Sederel CJ, et al. Caregiver burden in Parkinson's disease: a mixed-methods study. BMC Med. 2023;21. https://doi.org/10.1186/ s12916-023-02933-4.
- Plöthner M, Schmidt K, De Jong L et al. Needs and preferences of informal caregivers regarding outpatient care for the elderly: A systematic literature review. BMC Geriatr. 2019;19.
- Chiao CY, Wu HS, Hsiao CY. Caregiver burden for informal caregivers of patients with dementia: A systematic review. Int Nurs Rev. 2015;62.
- van Beusekom I, Bakhshi-Raiez F, de Keizer NF, et al. Reported burden on informal caregivers of ICU survivors: A literature review. Crit Care. 2016;20. htt ps://doi.org/10.1186/s13054-016-1185-9.
- Kaggwa MM, Najjuka SM, Mamun MA, et al. Involvement and burden of informal caregivers of patients with mental illness: the mediating role of affiliated stigma. BMC Psychiatry. 2023;23. https://doi.org/10.1186/s12888-023-0455 3-x.
- Metzelthin SF, Verbakel E, Veenstra MY, et al. Positive and negative outcomes of informal caregiving at home and in institutionalised long-term care: A cross-sectional study. BMC Geriatr. 2017;17. https://doi.org/10.1186/s12877-0 17-0620-3.
- Yuan Q, Zhang Y, Samari E, et al. Positive aspects of caregiving among informal caregivers of persons with dementia in the Asian context: a qualitative study. BMC Geriatr. 2023;23. https://doi.org/10.1186/s12877-023-03767-8.
- Roth DL, Fredman L, Haley WE. Informal caregiving and its impact on health: A reappraisal from population-based studies. Gerontologist. 2015;55. https:// doi.org/10.1093/geront/gnu177.
- Hoefman RJ, van Exel J, Brouwer WBF. Measuring the impact of caregiving on informal carers: A construct validation study of the CarerQol instrument. Health Qual Life Outcomes. 2013;11. https://doi.org/10.1186/1477-7525-11-1 73.
- Krol M, Papenburg J, van Exel J. Does including informal care in economic evaluations matter?? A systematic review of inclusion and impact of informal care in Cost-Effectiveness studies. PharmacoEconomics. 2015;33.
- 15. Urwin S, Lau YS, Grande G et al. The challenges of measuring informal care time: A review of the literature. PharmacoEconomics. 2021;39.
- Bobinac A, van Exel NJA, Rutten FFH, et al. Caring for and caring about: disentangling the caregiver effect and the family effect. J Health Econ. 2010;29. htt ps://doi.org/10.1016/j.jhealeco.2010.05.003.
- Baji P, Farkas M, Golicki D, et al. Development of population tariffs for the CarerQol instrument for Hungary, Poland and Slovenia: A discrete choice experiment study to measure the burden of informal caregiving. Pharmaco-Economics. 2020;38. https://doi.org/10.1007/s40273-020-00899-2.
- Drummond M. Methods for the economic evaluation of health care programmes. Fourth Edi. Oxford: Oxford University Press; 2015.
- Al-Janabi H, Flynn TN, Coast J. QALYs and carers. PharmacoEconomics. 2011;29.
- Greco G, Lorgelly P, Yamabhai I. Outcomes in economic evaluations of public health interventions in Low- and Middle-Income countries: health, capabilities and subjective wellbeing. Health Econ (United Kingdom). 2016;25. https:/ /doi.org/10.1002/hec.3302.
- Brouwer WBF, Van Exel NJA, Van Gorp B, et al. The CarerQol instrument: A new instrument to measure care-related quality of life of informal caregivers for use in economic evaluations. Qual Life Res. 2006;15. https://doi.org/10.1007/s 11136-005-5994-6.
- Hoefman RJ, van Exel J, Brouwer WBF. Measuring Care-Related quality of life of caregivers for use in economic evaluations: CarerQol tariffs for Australia, Germany, Sweden, UK, and US. PharmacoEconomics. 2017;35. https://doi.org /10.1007/s40273-016-0477-x.
- 23. Institute for Medical Technology Assessment -CarerQoL. https://www.imta.nl/ questionnaires/carerqol/instruments/

- Tang B, Harary E, Kurzman R, et al. Clinical characterization and the caregiver burden of dementia in China. Value Health Reg Issues. 2013;2. https://doi.org/ 10.1016/j.vhri.2013.02.010.
- 25. Li X, Lu J, Hu S et al. The primary health-care system in China. Lancet. 2017;390.
- Wild D, Grove A, Martin M, et al. Principles of good practice for the translation and cultural adaptation process for Patient-Reported outcomes (PRO) measures: report of the ISPOR task force for translation and cultural adaptation. Value Health. 2005;8:94–104. https://doi.org/10.1111/j.1524-4733.2005.04054.
- Liang X, Zhao M, Su L, et al. Sleep problems in children with ADHD: associations with internalizing symptoms and physical activity. J Autism Dev Disord Published Online First. 2024. https://doi.org/10.1007/s10803-024-06623-9.
- Prinsen CAC, Mokkink LB, Bouter LM, et al. COSMIN guideline for systematic reviews of patient-reported outcome measures. Qual Life Res. 2018;27:1147– 57. https://doi.org/10.1007/s11136-018-1798-3.
- Voormolen DC, van Exel J, Brouwer W, et al. A validation study of the CarerQol instrument in informal caregivers of people with dementia from eight European countries. Qual Life Res. 2021;30. https://doi.org/10.1007/s11136-020-02 657-5.
- Brazier JE, Mulhern BJ, Bjorner JB et al. Developing a new version of the SF-6D health state classification system from the SF-36v2: SF-6Dv2. Med Care. 2020.
- Wu J, Xie S, He X, et al. The simplified Chinese version of SF-6Dv2: translation, cross-cultural adaptation and preliminary psychometric testing. Qual Life Res. 2020;29. https://doi.org/10.1007/s11136-020-02419-3.
- Mulhern BJ, Bansback N, Norman R et al. Valuing the SF-6Dv2 classification system in the united Kingdom using a discrete-choice experiment with duration. Med Care. 2020.
- Al-Janabi H, Flynn T, Coast J. Development of a self-report measure of capability wellbeing for adults: the ICECAP-A. Qual Life Res. 2012;21:167–76. https: //doi.org/10.1007/s11136-011-9927-2.
- Tang C, Xiong Y, Wu H, et al. Adaptation and assessments of the Chinese version of the ICECAP-A measurement. Health Qual Life Outcomes. 2018;16:11– 45. https://doi.org/10.1186/s12955-018-0865-3.
- Flynn TN, Huynh E, Peters TJ, et al. Scoring the ICECAP-A capability instrument. Estimation of a UK general population tariff. Health Econ. 2015;24:258– 69. https://doi.org/10.1002/hec.3014.
- Peasgood T, Mukuria C, Brazier J, et al. Developing a new generic health and wellbeing measure: psychometric survey results for the EQ-HWB. Value Health. 2022;25. https://doi.org/10.1016/j.jval.2021.11.1361.
- McDool E, Mukuria C, Peasgood T. Psychometric performance of the EQ health and wellbeing short in a united Kingdom population sample. Value Health. 2024;27:1215–24. https://doi.org/10.1016/j.jval.2024.05.007.
- Terwee CB, Bot SDM, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol. 2007;60:34–42. https://doi.org/10.1016/j.jclinepi.2006.03.012.
- Cohen J. Statistical power analysis for the behavioral sciences. 2nd ed. Hillsdale, N.J.: L. Erlbaum Associates; 1988.
- Mair P. Modern psychometrics with R. Cham: Springer International Publishing: Imprint: Springer; 2018.
- 41. Gwet KL, Li Gwet K, Gwet KL. Handbook of Inter-Rater Reliability: the definitive guide to measuring the extent of agreement among raters. 2010.
- 42. Nunnally J, Bernstein I. Psychometric theory (3rd ed.). 1994.
- McLoughlin C, Goranitis I, Al-Janabi H. Validity and responsiveness of Preference-Based Quality-of-Life measures in informal carers: A comparison of 5 measures across 4 conditions. Value Health. 2020;23. https://doi.org/10.1 016/j.jval.2020.01.015.
- 44. Baji P, Brouwer WBF, van Exel J, et al. Validation of the Hungarian version of the CarerQol instrument in informal caregivers: results from a cross-sectional survey among the general population in Hungary. Qual Life Res. 2021;30. htt ps://doi.org/10.1007/s11136-020-02662-8.
- McCaffrey N, Bucholc J, Rand S, et al. Head-to-Head comparison of the psychometric properties of 3 Carer-Related Preference-Based instruments. Value Health. 2020;23. https://doi.org/10.1016/j.jval.2020.07.005.
- Al-Janabi H, Peters TJ, Brazier J, et al. An investigation of the construct validity of the ICECAP-A capability measure. Qual Life Res. 2013;22:1831–40. https://d oi.org/10.1007/s11136-012-0293-5.
- Xu RH, Wong EL-Y, Lu SY-J, et al. Validation of the Toronto empathy questionnaire (TEQ) among medical students in China: analyses using three psychometric methods. Front Psychol. 2020;11:810.

- caregivers. Value Health. 2024;27. https://doi.org/10.1016/j.jval.2024.03.003.
  49. Faraji H, Akbarzadeh I, Sahraian MA, et al. Psychometric validation of the Persian version of the CarerQol-7D instrument on caregivers of patients with multiple sclerosis. Curr J Neurol Published Online First: 8 September. 2024. htt ps://doi.org/10.18502/cjn.v23i1.16433.
- Hoefman RJ, Van Exel NJA, Foets M, et al. Sustained informal care: the feasibility, construct validity and test-retest reliability of the CarerQol-instrument to measure the impact of informal care in long-term care. Aging Ment Health. 2011;15. https://doi.org/10.1080/13607863.2011.575351.
- Vluggen TPMM, van Haastregt JCM, Tan FE, et al. Effectiveness of an integrated multidisciplinary geriatric rehabilitation programme for older persons with stroke: a multicentre randomised controlled trial. BMC Geriatr. 2021;21. h ttps://doi.org/10.1186/s12877-021-02082-4.

## Publisher's note

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