Zheng *et al. Health and Quality of Life Outcomes* (2025) 23:123
https://doi.org/10.1186/s12955-025-02467-8

Health and Quality of Life Outcomes

RESEARCH **Open Access**

Evaluating quality of care for dying patients from the perspective of bereaved relatives: validation of the Chinese version of the international care of the dying evaluation

Xuejiao Zheng[1,2*], Liying Zhou[3], Baiyan Li[4], Lei Dong[1], Yuna Li[5], Catriona R. Mayland[6,7], Maho Aoyama[1,8] and Mitsunori Miyashita[1]

Abstract

Background Few recognized, valid and reliable tools are used to assess the current quality of care at the very end-of-life from the bereaved relative's perspective in Mainland China. The purpose of this study was to validate the Chinese version of the international Care Of the Dying Evaluation (i-CODE) questionnaire by assessing its reliability and validity.

Methods From June 2023 to January 2024, participants were 216 bereaved relatives who were the primary caregivers of the deceased patients. Confirmatory Factor Analyses (CFA) were conducted, concurrent and discriminant validity was examined by correlating scores from the Chinese version of the i-CODE with Good Death Inventory (GDI), Care Evaluation Scale (CES), and Patient Health Questionnaire (PHQ-9). Internal reliability was assessed with Cronbach alpha (α) and test-retest reliability was assessed using Intraclass Correlation Coefficients (ICC).

Results The Chinese version of the i-CODE had four dimensions and 27 items were confirmed based on confirmatory factor analysis of the factor structure proposed by the authors of the original version. The fit indices were acceptable: Root Mean Square Error of Approximation (RMSEA) was 0.069, Standardised Root Mean Square Residual (SRMR) was 0.063, Comparative Fit Index (CFI) was 0.879, and Tucker-Lewis Index (TLI) was 0.864. The i-CODE was moderately correlated with the GDI ($r = 0.50$, $p < 0.001$) and CES ($r = 0.31$, $p < 0.001$) and not correlated with PHQ-9 ($r = 0.02$, $p = 0.765$). The internal consistency was excellent (Cronbach's $\alpha = 0.91$) and the test-retest reliability was good (ICC = 0.73).

Conclusions The Chinese version of the i-CODE demonstrated acceptable preliminary psychometric properties and broadly supported the original four-factor structure, indicating suitability for assessing the quality of care in the last days of life from the perspective of bereaved relatives in Mainland China. Further validation in larger and more diverse samples is needed, and the scale may be useful for clinical evaluation, quality improvement and cross-cultural research in end-of-life care.

Keywords Palliative care, Psychometrics, Family members, Quality of health care

---

*Correspondence:
Xuejiao Zheng
zheng.xuejiao.r8@dc.tohoku.ac.jp

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

Background

Approaching the very end of life is a unique and significant period for an individual and those important to them [1]. In China, where annual deaths already exceed ten million and are projected to rise further, the demand for palliative and end-of-life care is steadily increasing [2, 3]. However, recent evidence indicates that access to palliative care remains limited and uneven across regions and levels of care, and that the quality of death and dying is still lower than in many high-income countries [4, 5]. These disparities are driven by structural constraints, including fragmented policy implementation [6], concentration of services in tertiary urban hospitals [7], workforce shortages [8], and cultural barriers to disucssing prognosis and death [9, 10]. Therefore, addressing these systemic disparities requires robust, validated tools for benchmarking care quality and enabling evidence-based service improvement within China's specific cultural and health-system context.

An internationally recognized approach to evaluating end-of-life care involves assessing its quality through the perspectives of bereaved relatives, obtained via post-bereavement surveys [11]. Several instruments have been developed to capture different dimensions of end-of-life quality, including the international Care Of the Dying Evaluation (i-CODE) [12, 13], the Care Evaluation Scale (CES) [14], the Good Death Inventory (GDI) [15], and the Views of Informal Carers Evaluation of Services Short-Form (VOICE-SF) [16]. In Mainland China, Chinese versions of the CES and GDI are available [17, 18]. From a psychometric perspective, these instruments represent important progress: they have been translated and validated in Chinese samples of bereaved relatives of cancer patients, and either emphasise the structure and process of care (CES) or broader attributes of a "good death" across the end-of-life trajectory (GDI). However, they primarily focus on cancer populations, adopt a broader temporal perspective (not restricted to the last days), and provide less detailed assessment of acute symptom management and the peri-death period. Moreover, they were not originally designed as international benchmarking tools. As such, they are suboptimal for systematically evaluating hospital care in the final 48 h of life across both cancer and non-cancer conditions.

The international Care Of the Dying Evaluation (i-CODE) is a post-bereavement questionnaire specifically designed to assess care quality in the last 48 h of life across both cancer and non-cancer conditions. Grounded in core palliative care principles [12, 13, 19], its four-factor structure captures key domains of palliative and end-of-life care from bereaved relatives' perspective [13]. The instrument includes detailed items on symptom control (pain, dyspnoea, restlessness), communication, and emotional support. Its well-structured, moderate length makes the i-CODE sufficiently comprehensive for assessment yet feasible for routine use in clinical and quality-improvement settings. It has been implemented and validated in multiple countries and languages, enabling cross-national comparisons of quality of care for dying patients [12, 13]. For the Chinese context—where structural and cultural factors strongly influence end-of-life care—i-CODE therefore offers a conceptually coherent and practically useful framework for evaluation, particularly in hospital settings.

The introduction of the i-CODE in China holds considerable potential for advancing palliative care practices. As a crucial first step, a Chinese version must be developed and its cultural and psychometric validity established. Therefore, the aim of this study was to evaluate the psychometric properties of the Chinese version of the i-CODE. The two objectives were to assess the: 1.Validity, including construct validity using confirmatory factor analysis and concurrent validity using Pearson correlation analysis; 2. Reliability, including internal consistency and test-retest reliability.

Methods

Study design

A cross-sectional, post-bereavement survey was conducted to assess the psychometric properties of the Chinese version of the i-CODE and reported in keeping with the COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) Study Design checklist for Patient-reported outcome measurement instruments Version July 2019 [20].

Participants

Following common recommendations for psychometric validation studies, we considered a minimum of 5 to 10 participants per item to be acceptable [21, 22]. Based on the 27-item i-CODE, the required sample size was therefore at least 135 to 270 participants. In addition, an a priori power analysis for confirmatory factor analysis (CFA) indicated that approximately 160 participants would provide 80% power ($\alpha = 0.05$) to distinguish poor from acceptable model fit. On this basis, we aimed to recruit at least 200 participants.

Inclusion and exclusion criteria followed our prior publication [23]. Eligible bereaved relatives were adults (≥ 18 years) primary caregivers of patients during the last two days of life. Additional criteria included a minimum 3-month bereavement period and literacy in Chinese. Deceased patients were required to be adults (≥ 18 years) with a hospital stay of at least 72 h prior to death. We excluded cases where death occurred in intensive care units or resulted from unexpected or sudden causes (e.g., surgical complications or acute allergic reactions). Eligibility criteria targeted adult populations to align with original validation of i-CODE [12, 13] and comparable Chinese tools (CES, GDI). A minimum 72-hour hospital stay ensured meaningful exposure to inpatient care [24], while the 3-month post-bereavement survey balanced recall accuracy with emotional readiness [25].

Instruments

The instruments included the evaluation of demographic characteristics and the Chinese version of the following tools: the i-CODE, the CES, the GDI and the PHQ-9.

The international Care Of the Dying Evaluation questionnaire (i-CODE)

The Chinese version of the i-CODE was developed in a previous study following the EORTC Quality of Life Group translation procedure, involving forward-backward translation, expert review, and cognitive interview [23], resulting in a 32-item Chinese version consistent with the original structure. Of these, 27 items contribute to four composite factors and to the overall score: Factor 1 “Overall care” (8 items), Factor 2 “Communication and support” (8 items), Factor 3 “Trust, respect and dignity” (8 items) and Factor 4 “Symptom management” (3 items) [13]. The response options comprise both nominal and ordinal response options (5-point (0–4), 4-point (0–3) or 3-point (0–2) Likert scale), and the total score is calculated by summing the 27 factor items (range 0–66), with higher scores indicating better quality of care. The i-CODE was originally validated as a post-bereavement measure of care in the last 48 h of life and has shown good internal consistency and acceptable construct validity across seven countries [12, 13, 19], and was acceptable to complete in a single sitting.

Care evaluation scale (CES)

The CES was developed to measure end-of-life care for cancer patients from the bereaved family member’s perspective especially focusing on structure and process of care [14]. The short version of the CES consisted of 10 items with a 6-point Likert scale (1 = highly disagree to 6 = highly agree; 7 = N/A, if no other options were applicable) was used to assess concurrent validity. The score was transformed to a 0-100 point scale, with a high score indicating excellent care. The Chinese version of the CES has demonstrated good reliability and validity [17].

Good death inventory (GDI)

The GDI was developed to evaluate end-of-life care for cancer patients from the perspective of the bereaved family members [15]. The short version of the GDI consisted of 18 items with a 7-point Likert scale (1 = absolutely disagree to 7 = absolutely agree) was used to assess concurrent validity. Total score is calculated by summing all items. The higher score indicates higher quality of life. The Chinese version of the GDI has demonstrated good reliability and validity [18].

Patient health Questionnaire-9 (PHQ-9)

The PHQ-9 is a self-administered instrument developed to assess depression severity [26]. Each item is rated on a 4-point Likert scale (0 = not at all to 3 = nearly every day). The Chinese version of the PHQ-9 has demonstrated good reliability and validity [27].

Data collection

A convenience sampling method was used to recruit participants between June 25, 2023 and January 31, 2024. An online survey was used as the primary mode of data collection, supplemented by a small number of paper questionnaires in cases where online completion was not possible. During the data collection period, bereaved relatives from five medical institutes (three public tertiary hospitals, one private tertiary hospital, and one public secondary hospital) were recruited. Ward nurses at these hospitals contacted the primary bereaved relative via telephone or WeChat, explained the study, and provided a QR code for the online questionnaire. The first page of the questionnaire provided information about the study purpose and procedures, emphasised voluntary participation, and obtained informed consent. In the survey, respondents were prompted if any items were left unanswered, and the questionnaire could only be submitted once all items had been completed. This procedure reduced item-level missing data but also meant that some partially completed questionnaires could not be submitted and were therefore not available for analysis. Participants who returned a valid initial questionnaire were asked to complete it again online two weeks later to assess test-retest reliability.

Data analysis

All data were analysed using SPSS 26.0 and Amos 24.0. A two-tailed P-value ≤ 0.05 was considered statistically significant. Descriptive statistics summarized the demographic characteristics of deceased patients and bereaved relatives. Confirmatory Factor Analysis (CFA) was conducted to evaluate the fit of the four-factor i-CODE model. Model fit was considered acceptable if the the Root Mean Square Error of Approximation (RMSEA) and Standardised Root Mean Square residual (SRMR) < 0.08 and Comparative Fit Index (CFI) and Tucker-Lewis Index (TLI) ≥ 0.90 [28]. Model modification was performed based on modification indices (MIs) to improve model fit. Residual correlations and MIs were first examined to identify potentially redundant items. Only item pairs with high MI values (MI > 50) [29] were considered as candidates for releasing correlated residuals, and each modification was introduced sequentially, freeing one pair of items at a time. In addition to MI magnitude, standardised expected parameter change and theoretical justification [30] were also taken into account when deciding which residual covariances to free. This stepwise, conceptually guided procedure was adopted to enhance model fit indices (CFI and RMSEA) while minimizing the risk of overfitting and maintaining the interpretability of the measurement model. Factor loadings were classified as low (< 0.30), moderate (0.30–0.59), and high (≥ 0.60) [31]. Pearson correlation analysis was performed between the i-CODE total score and CES [17], GDI [18], and PHQ-9 [27] to assess concurrent validity and discriminant validity. Correlation coefficient ≤ 0.30, > 0.30, > 0.70 indicated low, moderate and high correlations, respectively [32]. Internal consistency was assessed using Cronbach's α for i-CODE total score and four factors. A Cronbach's α ≥ 0.70 indicates acceptable internal consistency and ≥ 0.80 indicates good internal consistency [33, 34]. Test-retest reliability was evaluated using intraclass correlation coefficient (ICC), with < 0.50, ≥ 0.50, and > 0.70 indicating poor, moderate, good reliability, respectively [35, 36].

Results

In total, 242 participants completed the survey (222 online and 20 paper). After excluding 26 questionnaires due to response patterns or obvious internal inconsistencies, 216 valid responses were included in the psychometric analyses (effective response rate, 89.3%). Of the 216 participants invited for the retest survey, 107 provided valid responses (effective response rate, 49.5%). The overall response rate could not be estimated because the sampling frame did not allow calculation of an accurate denominator.

Participant characteristics

Deceased patients were predominantly male (68.5%; mean age 58.70 ± 20.08 years), with cancer, cerebrovascular disease and cardiovascular disease as the leading causes of death. Bereaved relatives were also more often male (53.7%; mean age 33.16 ± 11.75 years), most commonly the patients' children (38.4%); additional characteristics are presented in Table 1.

Construct validity

Confirmatory Factor Analysis (CFA) was conducted to examin the construct validity of the original four-factor i-CODE model (Fig. 1). The model showed suboptimal fit: RMSEA = 0.087, CFI = 0.808, TLI = 0.786, and SRMR = 0.068 (Table 2). After freeing two theoretically justified residual covariances—between Q6 (confidence and trust in the nurses) and Q7 (confidence and trust in the doctors) and between Q3 ((environment was comfortable) and Q4 (environment had adequate privacy)—model fit improved to an overall acceptable, though not optimal: RMSEA = 0.069, CFI = 0.879, TLI = 0.864, and SRMR = 0.063 (Table 2). RMSEA and SRMR met commonly cited criteria (< 0.08), whereas CFI and TLI remained slightly below the conventional 0.90 threshold.

Standardised factor loadings for all 27 items are presented in Table 3. With the exception of one item, all loadings were ≥ 0.30, ranging from 0.309 to 0.890 across the four factors. Item 23 ("Before they died, were you told they were likely to die soon?") showed a very low loading (0.135) on the Factor 3 "Trust, respect and dignity" but was retained due to its conceptual importance and to maintain comparability with the original i-CODE. Interfactor correlations ranged from 0.605 to 0.945 (Table 4), indicating moderate to high associations among the latent constructs, with particularly strong correlations between Factor 1 and 3, and between Facor 2 and 3.

Concurrent and discriminant validity

Pearson correlation analysis was employed to explore the correlation between the i-CODE total score and the other validated scales. The i-CODE total score was moderately correlated with the GDI ($r = 0.502$, $p < 0.001$) and low-to-moderately correlated with the CES ($r = 0.311$, $p < 0.001$). In contrast, its correlation with the PHQ-9 was negligible and not statistically significant ($r = 0.015$, $p = 0.765$), indicating that higher perceived quality of care in the last days of life was largely independent of relatives' depressive symptom severity (Table 5).

Internal consistency

The overall i-CODE and four factors showed moderate to excellent reliability scores (Factor 1 'Overall care' α = 0.832; Factor 2 'Communication and support' α = 0.778; Factor 3 'Trust, respect and dignity' α = 0.698; Factor 4 'Symptom management' α = 0.769; Overall α = 0.913).

Test-retest reliability

Moderate to good test-retest reliability was evaluated with ICC for overall i-CODE and four factors. Overall ICC = 0.734 ($p < 0.001$); Factor 1 'Overall care' ICC = 0.697 ($p < 0.001$); Factor 2 'Trust, Respect & Dignity' ICC = 0.616 ($p < 0.001$); Factor 3 'Communication and Support' ICC = 0.645 ($p < 0.001$), Factor 4'Symptom Management' ICC = 0.676 ($p < 0.001$).

Discussion

The i-CODE has been increasingly used internationally to evaluate care in the last days of life, but evidence on its measurement properties in Mainland China has been lacking. This study provides the first evaluation of the Chinese i-CODE in a hospital-bereaved sample, showing acceptable preliminary reliability and validity and broadly supporting the original four-factor framework.

Table 1 General characteristic of deceased patients and bereaved relatives (*n* = 216)

Characteristic	*n*	%	Characteristic	*n*	%
< Deceased patients>			**< Bereaved relatives>**		
Age, years (mean + SD, range)			**Age, years (mean + SD, range)**		
58.7 ± 20.08 (18 ~ 96)			33.16 ± 11.75 (18 ~ 76)		
18 ~ 44	57	26.4	18 ~ 30	109	50.5
45 ~ 59	33	15.3	31 ~ 40	65	30.1
60 ~ 74	72	33.3	40 ~ 50	25	11.6
75~	54	25	50~	17	7.9
Gender			**Gender**		
Male	148	68.5	Male	116	53.7
Female	68	31.5	Female	100	46.3
Marital status			**Marital status**		
Married	174	80.6	Married	116	53.7
Unmarried	24	11.1	Unmarried	92	42.6
Widow	15	6.9	Widow	4	1.9
Divorce	3	1.4	Divorce	4	1.9
Religious belief			**Religious belief**		
No specific religious beliefs	160	74.1	No specific religious beliefs	177	81.9
Buddhism	33	15.3	Buddhism	19	8.8
Islam	5	2.3	Islam	8	3.7
Catholicism · Christianity	6	2.8	Catholicism · Christianity	5	2.3
Other	12	5.6	Other	7	3.2
Hospitalisation days			**Educational level**		
3 ~ 6days	36	16.7	Primary school · junior high school	11	5.1
7 ~ 10days	27	12.5	High school	43	19.9
11 ~ 15days	56	25.9	College	64	29.6
16 ~ 30days	41	19.0	Undergraduate · graduate	94	43.5
31days or more	56	25.9	Other	4	1.9
Cause of death			**Relationship to patient**		
Cancer	73	33.8	Spouse	37	17.1
Cerebrovascular disease	26	12.0	Children	83	38.4
Cardiovascular disease	34	15.7	Son-in-law · daughter-in-law	15	6.9
Chronic obstructive pulmonary disease (COPD)	11	5.1	Parents	5	2.3
Diabetes	6	2.8	Siblings	42	19.4
Cirrhosis	10	4.6	others	34	15.7
Hypertension	10	4.6	**Per capita monthly family income(yuan)**		
End stage renal disease	5	2.3	< 1000	8	3.7
Dementia	2	0.9	1000 ~ 1999	10	4.6
Motor neuron disease	3	1.4	2000 ~ 2999	25	11.6
Senility	12	5.6	3000 ~ 3999	23	10.6
Don't know	9	4.2	4000 ~ 4999	33	15.3
Other	15	6.9	5000 or more	117	54.2

Construct validity and factor structure

Our findings provide partial support for the four-factor structure originally proposed for the i-CODE [12, 13]. These domains are grounded in palliative care quality frameworks that emphasise the interplay between care processes and family-reported outcomes [19]. In our CFA, RMSEA and SRMR met conventional criteria (< 0.08), whereas CFI and TLI fell slightly below 0.90. This pattern is not uncommon in complex, multifactorial models evaluated in relatively modest samples, where absolute and residual-based indices may indicate acceptable fit while incremental indices remain below "excellent" thresholds; therefore, fit indices should be interpreted collectively and in context rather than as rigid cut-offs [37–39].

Model modifications in this study were deliberately conservative and theory-driven. Residual covariances were freed only for item pairs with high modification indices and clear conceptual overlap, which improved model fit while preserving the original factor framework and limiting data-driven re-specification, thereby supporting the interpretability and reducing the risk of overfitting [40].

Compared with the original i-CODE validation, which also applied theory-guided modifications and reported stronger global fit (RMSEA = 0.057, CFI = 0.962) [13], our Chinese validation yieded acceptable but comparatively lower indices. To our knowledge, the German version of i-CODE (CODE-GER) has been published to date, but its validation focused on concurrent validity and did not report a separate CFA for the four-factor structure, limiting direct comparisons of model-fit indices with our findings [41]. In the Chinese context, related family-reported end-of-life care instruments such as the CES (RMSEA = 0.047, CFI = 0.952, TLI = 0.946) [17] and GDI (RMSEA = 0.044, CFI = 0.900, TLI = 0.892) [18] have shown acceptable factor structures with slightly stronger fit than observed here. The i-CODE, however, targets the last 48 h of life and includes more heterogeneous items on acute symptom management and circumstances around death, which may partly explain the comparatively less optimal global fit in our sample.

High inter-factor correlations were also observed. Although this pattern may indicate partial construct overlap statistically, it is conceptually understandable in the Chinese context, where overall care, respectful treatment and supportive communication are often experienced as closely intertwined facets of a single care experience [42, 43]. Similar high correlations were reported in the original validation, where bifactor model was explored as alternative structure [13]. Given the multidimensional nature of palliative care and the practical value of domain-level feedback for quality improvement, we retained four-domain model. Future studies

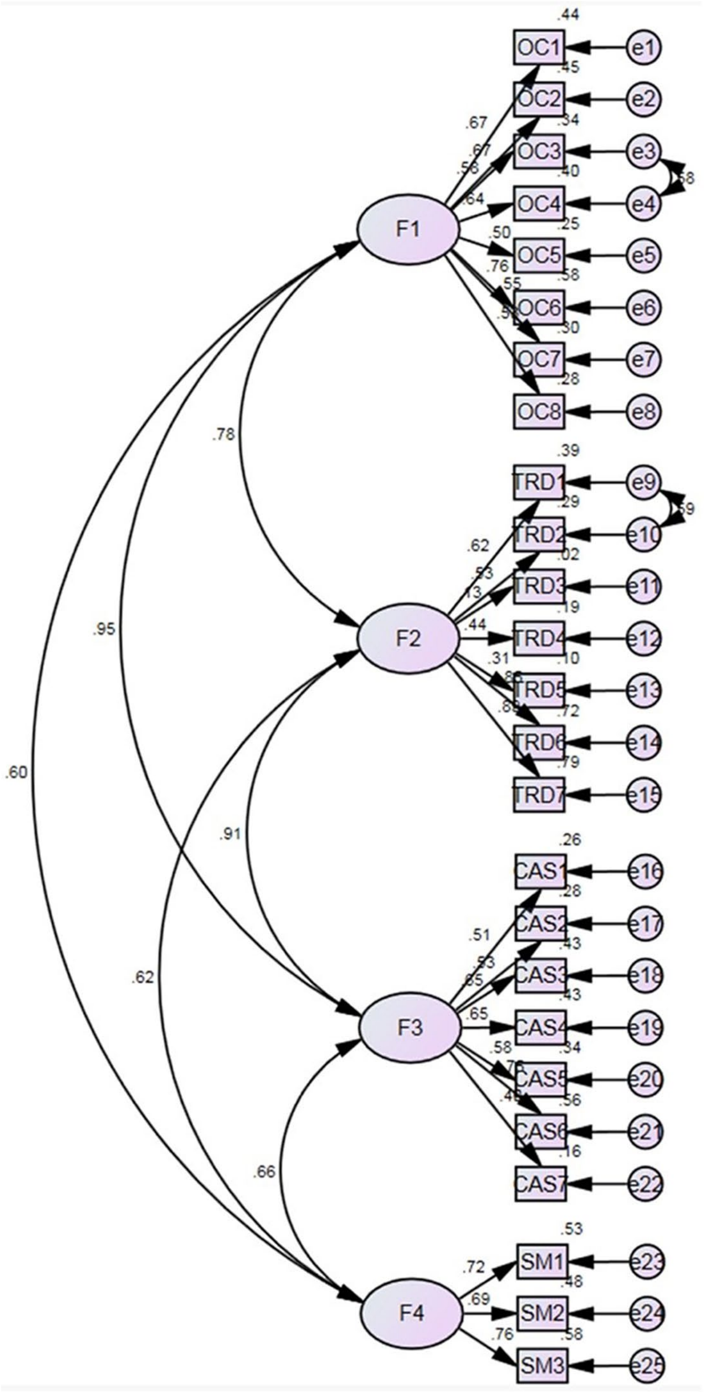

Fig. 1 Confirmatory Factor Analysis Model of Chinese Version of the i-CODE. Note: F1: Overall care; F2: Trust, respect and dignity; F3: Communication and support; F4: Symptom management

with larger samples should formally compare correlated four-factor, bifactor and higher-order models to assess whether adding a general "overall quality of care" factor provides a better representation of the construct in Chinese settings.

The low loading item 23 may have contributed to the less-than-optimal global fit. Prognostic communication and timely recognition of dying are widely recognised as core elements of high-quality end-of-life care and are central to the construct that the i-CODE seeks to capture [44–46]. Removing this item would also reduce comparability with international datasets using the full i-CODE item set. The weak loading may therefore reflect socio-cultural influences on prognostic disclosure and recall

Table 2 Model fit indices for Chinese version of the i-CODE

Fit index	Recommended Cutoff	Initial Model	Modification Model	Final Model
χ^2/df	< 3.00	2.611	2.303	2.027
RMSEA	< 0.08	0.087	0.078	0.069
CFI	≥ 0.90	0.808	0.845	0.879
TLI	≥ 0.90	0.786	0.827	0.864
IFI	≥ 0.90	0.811	0.847	0.880
SRMR	< 0.08	0.068	0.068	0.063

Note: χ^2/df = Chi-Square to Degrees of Freedom; RMSEA = Root Mean Square Error of Approximation; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; IFI = Incremental Fit Index; SRMR = Standardised Root Mean Square Residual

Table 3 The standardised factor loadings of each item and the assigned factor

Item No.	Item Content (Short Description)	Assigned Factor	Standardised Loading
Q1	*Personal care needs*	Factor 1: Overall Care	0.667
Q2	*Nursing care needs*	Factor 1: Overall Care	0.670
Q3	*Environment was Comfortable*	Factor 1: Overall Care	0.583
Q4	*Environment had Adequate privacy*	Factor 1: Overall Care	0.635
Q5	*Cleanliness of the ward area*	Factor 1: Overall Care	0.502
Q8	*Nurses had time to listen and discuss*	Factor 1: Overall Care	0.761
Q27	*Died in the right place*	Factor 1: Overall Care	0.548
Q32	*Willingness to recommend the organization*	Factor 1: Overall Care	0.529
Q6	*Confidence and trust in nurses*	Factor 2: Trust, Respect and Dignity	0.623
Q7	*Confidence and trust in doctors*	Factor 2: Trust, Respect and Dignity	0.535
Q23	*Informed about impending death*	Factor 2: Trust, Respect and Dignity	0.135
Q24/25	*Discussion about dying process helpful*	Factor 2: Trust, Respect and Dignity	0.436
Q29	*Sensitivity from team after death*	Factor 2: Trust, Respect and Dignity	0.309
Q30a	*Treated with respect and dignity (doctors)*	Factor 2: Trust, Respect and Dignity	0.848
Q30b	*Treated with respect and dignity (nurses)*	Factor 2: Trust, Respect and Dignity	0.890
Q16	*Involvement in care decisions*	Factor 3: Communication and Support	0.509
Q17/18	*Discussion about fluids via drip*	Factor 3: Communication and Support	0.527
Q19	*Clarity of explanation about condition*	Factor 3: Communication and Support	0.654
Q20	*Emotional support from team*	Factor 3: Communication and Support	0.653
Q21	*Religious/spiritual needs met*	Factor 3: Communication and Support	0.584
Q28	*Support at time of death*	Factor 3: Communication and Support	0.746
Q31	*Supported during last two days*	Factor 3: Communication and Support	0.396
Q11	*Relief of Pain*	Factor 4: Symptom Management	0.725
Q13	*Relief of restlessness*	Factor 4: Symptom Management	0.693
Q15	*Relief of noisy rattle (breathing)*	Factor 4: Symptom Management	0.760

Table 4 Correlations between factors

Factor	Factor 1	Factor 2	Factor 3	Factor 4
Factor 1	1.000	0.783	0.945	0.605
Factor 2		1.000	0.913	0.619
Factor 3			1.000	0.657
Factor 4				1.000

Note: F1: Overall care; F2: Trust, respect and dignity; F3: Communication and support; F4: Symptom management

Table 5 Concurrent validity and discriminant validity of Chinese version of i-CODE

Factors	GDI	CES	PHQ-9
Overall Care	0.416^{***}	0.272^{***}	0.035
Trust, Respect & Dignity	0.446^{***}	0.283^{***}	-0.058
Communication and Support	0.483^{***}	0.280^{***}	0.010
Symptom Management	0.357^{***}	0.213^{***}	0.092
i-CODE	0.502^{***}	0.311^{***}	0.015

Note: Table are Pearson's correlation coefficients. CES = Care Evaluation Scale; GDI = Good Death Inventory; PHQ-9 = Patient Health Questionnaire – 9 items

$^{***}P < 0.001$

rather than irrelevance of the construct. Further qualitative and quantitative research is warranted to explore culturally specific interpretations of this item and to determine whether refinement is needed in future Chinese applications.

Concurrent and discriminant validity

The pattern of correlations observed in our study was similar to that reported in the CODE-GER, which found a moderate correlations ($r = -0.41$, $p < 0.001$) with established palliative care measures [41]. As hypothesized, the Chinese i-CODE showed a low-to-moderate correlation with the CES, which emphasises the structure and process of care [17], and a moderate correlation with the GDI, which focuses on broader attributes of a "good death" across the end-of-life trajectory [18]. Notably, The relatively low correlation between the i-CODE Symptom Management subscale and the GDI further underscores the i-CODE's specific focus on symptom management in the last 48 h of life. Conversely, the negligible and non-significant correlation with the PHQ-9 (r = 0.015, p = 0.765) [27] supports discriminant validity, indicating that relatives' evaluations of end-of-life quality are independent of depressive symptoms severity. Therefore, these findings provide evidence of concurrent and discriminant validity and support the structure validity of the Chinese i-CODE.

Internal consistency and test–retest reliability

The Chinese i-CODE showed good internal consistency (α = 0.91) for the total score, comparable the original international validation (α = 0.92) [13] and the CODE-GER study (α = 0.86) [41]. Across the four domains, internal consistency ranged from satisfactory to good, consistent with COSMIN guidance that reliability should be assessed at the subscale level in multidimensional instruments [20]. Test–retest reliability (ICC) was moderate to good but slightlylower than that reported in the German study (ICC = 0.85) [41]. This difference may reflect population-specific response patterns as well as measurement characteristics of the source instrument that can influence temporal stability across cultural contexts [47]. In addition, subtle linguistic or conceptual shifts introduced during translation may affect the stability of respondents' interpretations over time, a recognized challenge in cross-cultural validation research [48]. Nevertheless, the ICC values indicate that the Chinese i-CODE yields reasonably stable ratings among bereaved relatives, supporting its use in both clinical and research settings.

Implications, limitations and future directions

These findings suggest that the Chinese i-CODE provides a practical means of capturing bereaved relatives' views on care in the last two days of life, helping healthcare teams identify specific strengths and targets for improvement (e.g. symptom relief, communication and family support, and perceived dignity). More broadly, a validated i-CODE can facilitate cross-study and cross-country comparisons and support more patient- and family-centred monitoring of end-of-life care quality as palliative care services continue to develop in China.

This study has several limitations. First, data were collected primarily via a voluntary online survey, which may over-represent younger and more digitally literate relatives and limit generalisability. Second, although the sample size was sufficient for testing the four-factor CFA model, it may be underpowered for more complex structures, warranting replication in larger and more diverse samples. Third, the very low loading of Q23 highlights the need to further explore culturally specific understandings of prognostic communication. Finally, as the sample was restricted to relatives of adult in-hospital decedents, future studies should evaluate the Chinese i-CODE in home and community settings. Overall, these findings provide preliminary support for the Chinese i-CODE as a measure of care quality in the last days of life.

Conclusion

This study provides preliminary evidence that the Chinese version of the i-CODE is a culturally appropriate instrument with acceptable psychometric properties for assessing the quality of care in the last days of life from the perspective of bereaved family members in Mainland China. The original four-factor structure was broadly supported, with satisfactory internal consistency, moderate test–retest reliability, and adequate concurrent and discriminant validity. However, a modest sample size, slightly suboptimal global fit indices, and one low-loading item indicate that further refinement and replication in larger and more diverse samples are warranted. Despite these limitations, the Chinese i-CODE has potential to support clinical evaluation, quality improvement, and cross-cultural research in end-of-life care.

Acknowledgements

We gratefully thank all the members of the research team for making the study progress smoothly. Additionally, we would like to thank all the bereaved relatives for sharing their views.

Author contributions

XZ conceived and designed the study, conducted the statistical analyses, and drafted the initial manuscript. LZ, BL, LD and YL coordinated participant recruitment, data collection. CRM, MA and MM provided methodological guidance, supervised the study design and analysis, and critically revised the manuscript. All authors contributed to the interpretation of data, reviewed the manuscript, and approved the final version for submission.

Funding

This study was supported by JST SPRING, Grant Number JPMJSP2114. The authors have no relevant financial or non-financial interests to disclose. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

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 was carried out in line with the principles of the Declaration of Helsinki, including subsequent revisions, and was approved by the Institutional Review Board (IRB) of the Tohoku University (Approval Number: 2022-1-1022). Consent was secured from all individuals after they were fully informed prior to their involvement in the research.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]Department of Palliative Nursing, Tohoku University Graduate School of Medicine, 2-1 Seiryo-machi, Aoba-ku, Sendai, Japan
[2]Changchun Humanities and Sciences College, Changchun, China
[3]Department of Respiratory and Critical Care, The Second Norman Bethune Hospital of Jilin University, Changchun, China
[4]School of Nursing, Changchun Vocational College of Health, Changchun, China
[5]Department of Neurosurgery, The First Hospital of JILIN University, Changchun, China
[6]Division of Clinical Medicine, School of Medicine & Population Health, University of Sheffield, Sheffield, England
[7]Palliative Care Unit, University of Liverpool, Liverpool, England
[8]Department of Nursing, Yamagata Prefectural University of Health Sciences, Yamagata, Japan

Received: 28 September 2025 / Accepted: 10 December 2025

References

1. Radbruch L, De LL, Knaul F, et al. Redefining palliative Care-A new Consensus-Based definition. J Pain Symptom Manage. 2020;60(4):754–64.
2. Central People's Government of the People's Republic of China. Interpretation of the main economic data for 2023 by the head of the relevant department of the National Statistical Office (in Chinese). 2024. https://www.gov.cn/lianbo/bumen/202401/content_6926737.htm. Accessed 3 November 2024.
3. Department of Economic and Social Affairs Population Division of United Nation. World Population Prospects 2024. https://population.un.org/wpp/Graphs/Probabilistic/MORT/Deaths/156. Accessed 3 November 2024.
4. Li L, Gong X. Developing supportive policy environments for hospice care in china: a quantitative policy evaluation based on the PMC-Index model. BMC Palliat Care. 2025;24(1):248.
5. Tripodoro VA, Fidalgo JFL, Pons JJ, et al. First-Ever global ranking of palliative care: 2025 world map under the new WHO framework. J Pain Symptom Manage. 2025;70(5):447–58.
6. Shi Z, Du M, Zhu S, et al. Factors influencing accessibility of palliative care: a systematic review and meta-analysis. BMC Palliat Care. 2025;24(1):80.
7. Lin H, Huang Y, Wang Y. Facilitators and barriers to palliative care delivery in rural china: a qualitative study of the perceptions and experiences of rural healthcare professionals. BMC Palliat Care. 2025;24(1):211.
8. Wang J, Song S, Mai G, et al. The comparison of Spatial patterns and factors associated with healthcare provider knowledge in palliative care in various regions of China. Humanit Soc Sci Commun. 2024;11:510.
9. He L, Liao H, Pan Y, et al. Current public perception of, and barriers to, palliative care in china: a mixed-methods study. Lancet Oncol. 2022;23:S8.
10. Zuo T, Yeap SY, Lu G. Navigating the delicate balance of autonomy and harmony: a case study on the cultural adaptation of palliative care interventions in China. BMC Palliat Care. 2025;24(1):168.
11. Haugen DF, Hufthammer KO, Gerlach C, et al. Good quality care for cancer patients dying in Hospitals, but information needs unmet: bereaved relatives' survey within seven countries. Oncologist. 2021;26(7):e1273–84.
12. Mayland CR, Gerlach C, Sigurdardottir K, et al. Assessing quality of care for the dying from the bereaved relatives' perspective: using pre-testing survey methods across seven countries to develop an international outcome measure. Palliat Med. 2019;33(3):357–68.
13. Mayland CR, Keetharuth AD, Mukuria C, et al. Validation of 'Care of the dying evaluation' (CODE™) within an international study exploring bereaved relatives' perceptions about quality of care in the last days of life. J Pain Symptom Manage. 2022;64(1):e23–33.
14. Miyashita M, Aoyama M, Nakahata M, et al. Development the care evaluation scale version 2.0: a modified version of a measure for bereaved family members to evaluate the structure and process of palliative care for cancer patient. BMC Palliat Care. 2017;16(1):8.
15. Miyashita M, Morita T, Sato K, et al. Good death inventory: a measure for evaluating good death from the bereaved family member's perspective. J Pain Symptom Manage. 2008;35(5):486–98.
16. Hunt KJ, Richardson A, Darlington AE, et al. Developing the methods and questionnaire (VOICES-SF) for a National retrospective mortality follow-back survey of palliative and end-of-life care in England. BMJ Support Palliat Care. 2019;9(1):e5.
17. Zhao J, You L, Tao H, et al. Validation of the Chinese version of the care evaluation scale for measuring the quality of structure and process of end-of-life care from the perspective of bereaved family. BMC Palliat Care. 2021;20(1):85.
18. Zhao J, Wong FKY, You L, et al. Validation of the Chinese version of the good death inventory for evaluating End-of-Life care from the perspective of the bereaved family. J Pain Symptom Manage. 2019;58(3):472–80.
19. Mayland CR, Lees C, Germain A, et al. Caring for those who die at home: the use and validation of 'Care of the dying evaluation' (CODE) with bereaved relatives. BMJ Support Palliat Care. 2014;4(2):167–74.
20. COSMIN [homepage on the internet]. COSMIN Study Design checklist for Patient-reported outcome measurement instruments. Version July 2019. https://www.cosmin.nl/wp-content/uploads/COSMIN-study-designing-checklist_final.pdf. Accessed 1 April 2024.
21. Terwee CB, Bot SD, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol. 2007;60(1):34–42.
22. Streiner DL, Norman GR, Cairney J. Health measurement scales: a practical guide to their development and use. 5th ed. Oxford University Press; 2015.
23. Zheng X, Zhao J, Mayland CR, et al. Translation and cultural adaptation of the Chinese version of the international 'Care of the dying evaluation' (CODE™) tool for assessing the quality of care for dying patients from the perspective of bereaved family members. BMC Palliat Care. 2025;24(1):128.
24. Zhou P, Tang C, Wang J, et al. Assessing the relationship between the distress levels in patients with irreversible terminal delirium and the good quality of death from the perspective of bereaved family. BMC Palliat Care. 2025;24(1):14.
25. Ham L, Fransen HP, van den Borne B, et al. Bereaved relatives' quality of life before and during the COVID-19 pandemic: results of the prospective, multicenter, observational eQuiPe study. Palliat Med. 2021;35(8):1502–7.
26. 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.
27. Bian C, He X, Qian J, et al. Application of the patient health questionnaire depressive symptom cluster scale in a general hospital (in Chinese). J Tongji Univ (Medical Edition). 2009;30(05):136–40.
28. Wang J, Wang X, Jiang B. Structural equation models: methods and applications. Beijing: Higher Education Press; 2011.
29. Rodríguez-Jiménez MC, Pérez Jorge D, González-Contreras AI, et al. Development and validation of EDU-COV-NEE-NEAE questionnaire to measure educational response, ICT usage and impact of COVID-19 pandemic on students with special educational needs. Heliyon. 2023;9(9):e19231.
30. Garnier-Villarreal M, Jorgensen DT. Evaluating local model misspecification with modification indices in bayesian structural equation modeling. Struct Equation Modeling: Multidisciplinary J. 2025;32(2):304–18.
31. Kamaruzzaman S, Ploubidis GB, Fletcher A, et al. A reliable measure of frailty for a community dwelling older population. Health Qual Life Outcomes. 2010;8:123.
32. Weber JC, Lamb DR. Statistics and research in physical education. St. Louis: The C.V.Mosby Company; 1970.
33. Devellis RF. Scale development: theory and applications. 4th ed. Los Angeles: Sage; 2017.
34. Pourmomeny AA, Ghanei B, Alizadeh F. Reliability and validity of the Persian Language version of the international consultation on incontinence Questionnaire-Male lower urinary tract symptoms (ICIQ-MLUTS). Lower Urinary Tract Symptoms. 2018;10(2):190–2.
35. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med. 2016;15(2):155–63.
36. Webb KL, Ryan J, Wolfe R, et al. Test-Retest reliability and minimal detectable change of four cognitive tests in Community-Dwelling older adults. J Alzheimers Dis. 2022;87(4):1683–93.
37. Shi D, Lee T, Maydeu-Olivares A. Understanding the model size effect on SEM fit indices. Educ Psychol Meas. 2019;79(2):310–34.
38. McNeish D, An J, Hancock GR. The thorny relation between measurement quality and fit index cutoffs in latent variable models. J Pers Assess. 2018;100(1):43–52.

39. Lai K, Green SB. The problem with having two watches: assessment of fit when RMSEA and CFI disagree. Multivar Behav Res. 2016;51(2–3):220–39.
40. Kline RB. Principles and practice of structural equation modeling. 5th ed. New York: Guilford Press; 2023.
41. Vogt A, Stiel S, Heckel M, Goebel S, Mai SS, Seifert A, et al. Assessment of the quality of end-of-life care: translation and validation of the German version of the care of the dying evaluation (CODE-GER) - a questionnaire for bereaved relatives. Health Qual Life Outcomes. 2020;18(1):311.
42. Prince M. Measurement validity in cross-cultural comparative research. Epidemiol Psichiatr Soc. 2008;17(3):211–20.
43. Lopez-Vergara HI, Yang M, Weiss NH, et al. The cultural equivalence of measurement in substance use research. Exp Clin Psychopharmacol. 2021;29(5):456–65.
44. National Clinical Guideline Centre (UK). Care of dying adults in the last days of life. London: National Institute for Health and Care Excellence (NICE); 2015.
45. Clayton JM, Hancock KM, Butow PN, et al. Clinical practice guidelines for communicating prognosis and end-of-life issues with adults in the advanced stages of a life-limiting illness, and their caregivers. Med J Aust. 2007;186(S12):S77–105.
46. Maya S, Chittem M, Chawak S, Parker PA, Banerjee SC. Experiences of prognosis disclosure versus nondisclosure among family caregivers of persons with advanced cancer. Death Stud. 2024;48(9):905–15.
47. Boateng GO, Neilands TB, Frongillo EA, Melgar-Quiñonez HR, Young SL. Best practices for developing and validating scales for Health, Social, and behavioral research: A primer. Front Public Health. 2018;6:149.
48. Sousa VD, Rojjanasrirat W. Translation, adaptation and validation of instruments or scales for use in cross-cultural health care research: a clear and user-friendly guideline. J Eval Clin Pract. 2011;17(2):268–74.

Publisher's note