Article:
Mayland, C.R. orcid.org/0000-0002-1440-9953, Keetharuth, A.D., Mukuria, C. et al. (1 more author) (2022) Validation of 'Care Of the Dying Evaluation' (CODETM) within an international study exploring bereaved relatives' perceptions about quality of care in the last days of life. Journal of Pain and Symptom Management, 64 (1). E23-E33. ISSN 0885-3924

https://doi.org/10.1016/j.jpainsymman.2022.02.340

Article:
Mayland, C orcid.org/0000-0002-1440-9953, Keetharuth, AD, Mukuria, C et al. (1 more author) (Accepted: 2022) Validation of 'Care Of the Dying Evaluation' (CODETM) within an international study exploring bereaved relatives' perceptions about quality of care in the last days of life. Journal of Pain and Symptom Management. ISSN 1873-6513 (In Press)

Journal of Pain and Symptom Management

International validation of CODETM (running title)

--Manuscript Draft--

Manuscript Number:	JPSM-D-21-01320R1
Article Type:	Brief Methodological Report
Section/Category:	Research
Keywords:	palliative care; Terminal Care; Psychometrics; Factor Analysis; quality of care; quality of death
Corresponding Author:	Catriona Rachel Mayland, MBChB, FRCP, MD The University of Sheffield Sheffield, UNITED KINGDOM
First Author:	Catriona Rachel Mayland, MBChB, FRCP, MD
Order of Authors:	Catriona Rachel Mayland, MBChB, FRCP, MD
	Anju D Keetharuth, PhD
	Clara Mukuria, PhD
	Dagny Faksvåg Haugen, MD, PhD
Abstract:	Context: Assessing quality of care provided during the dying phase using validated tools aids quality assurance and recognises unmet need. Objective: To assess construct validity and internal consistency of 'Care Of the Dying Evaluation' (CODE TM) within an international context. Methods: Post-bereavement survey (August 2017-September 2018) using CODE TM . Respondents were next-of-kin to adult patients (≥ 18 years old) with cancer who had an 'expected' death within 22 study site hospitals in 7 countries: Argentina, Brazil, Germany, Norway, Poland, United Kingdom, Uruguay. Exploratory and Confirmatory Factor Analysis (EFA and CFA) were conducted, and internal reliability was assessed using Cronbach alpha (α). Known group validity was assessed by ability to discriminate quality of care based in place (Palliative Care Units (PCUs)) and country (Poland, where most deaths were in PCUs) of care. Differences were quantified using effect sizes (ES). Results: 914 CODE TM questionnaires completed (54% response rate). 527 (58%) male deceased patients; 610 (67%) next-of-kin female who were most commonly the 'spouse/partner' (411, 45%). EFA identified 4 factors: 'Overall care', 'Communication and support', 'Trust, respect and dignity', and 'Symptom management' with good reliability scores (α= 0.628 – 0.862). CFA confirmed the 4-factor model; these were highly correlated and a bifactor model showed acceptable fit. The ES for quality of care in PCU's was 0.727; ES for Poland was 0.657, supporting the sensitivity of CODE TM to detect differences. Conclusion: Within an international context, good evidence supports the validity and reliability of CODE TM for assessing the quality of care provided in the last days of life.

Title page

Validation of 'Care Of the Dying Evaluation' (CODE^{TM}) within an international study exploring bereaved relatives' perceptions about quality of care in the last days of life

Catriona R Mayland, MBChB, MD, Yorkshire Cancer Research Senior Clinical Research Fellow and Honorary Consultant in Palliative Medicine, University of Sheffield; Honorary Clinical Fellow, University of Liverpool; ORCID ID 0000-0002-1440-9953

Anju D Keetharuth, PhD, School of Health and Related Research, University of Sheffield, Sheffield, UK

Clara Mukuria, PhD, School of Health and Related Research, University of Sheffield, Sheffield, UK

Dagny Faksvåg Haugen, MD, PhD, Regional Centre of Excellence for Palliative Care, Western Norway, Haukeland University Hospital, and Department of Clinical Medicine K1, University of Bergen, Bergen, Norway

Corresponding author:

Dr Catriona Mayland, Department of Oncology and Metabolism, 2nd Floor Broomcross Building, Weston Park Hospital, Whitham Road, Sheffield S10 2SJ

E-mail: c.r.mayland@sheffield.ac.uk

Twitter: @CattyRM

Number tables: 5

Number figures: 0

Number of references: 35

Word count: 2517

Supplementary files: 5

Abstract (word count 250)

Context: Assessing quality of care provided during the dying phase using validated tools aids quality assurance and recognises unmet need.

Objective: To assess construct validity and internal consistency of 'Care Of the Dying Evaluation' (CODE^{TM}) within an international context.

Methods: Post-bereavement survey (August 2017-September 2018) using CODE^{TM}. Respondents were next-of-kin to adult patients (≥ 18 years old) with cancer who had an 'expected' death within 22 study site hospitals in 7 countries: Argentina, Brazil, Germany, Norway, Poland, United Kingdom, Uruguay. Exploratory and Confirmatory Factor Analysis (EFA and CFA) were conducted, and internal reliability was assessed using Cronbach alpha (α). Known group validity was assessed by ability to discriminate quality of care based in place (Palliative Care Units (PCUs)) and country (Poland, where most deaths were in PCUs) of care. Differences were quantified using effect sizes (ES).

Results: 914 CODE^{TM} questionnaires completed (54% response rate). 527 (58%) male deceased patients; 610 (67%) next-of-kin female who were most commonly the 'spouse/partner' (411, 45%). EFA identified 4 factors: 'Overall care', 'Communication and support', 'Trust, respect and dignity', and 'Symptom management' with good reliability scores (α = 0.628 – 0.862). CFA confirmed the 4-factor model; these were highly correlated and a bifactor model showed acceptable fit. The ES for quality of care in PCU's was 0.727; ES for Poland was 0.657, supporting the sensitivity of CODE^{TM} to detect differences.

Conclusion: Within an international context, good evidence supports the validity and reliability of CODE^{TM} for assessing the quality of care provided in the last days of life.

Key message

This article describes the validation of the 'Care Of the Dying Evaluation' (CODE™) questionnaire within an international study. The results indicate that CODE™ represents a valid and reliable tool for assessing the quality of care provided to dying patients and their families. (42 words)

Key words: palliative care; terminal care; psychometrics; factor analysis; quality of care; quality of death

Running title: International validation of CODE™

Introduction

Ensuring dying patients and their families receive high quality care and support is fundamentally important. How an individual dies has a profound impact on those bereaved. Timely, informative end-of-life discussions, (~~for example~~), are associated with (~~less depression and~~) less complicated grief.[1] Conversely, overly aggressive medical interventions and dissatisfaction with communication are linked with poorer bereavement outcomes.[2,3]

Globally, the provision of care for dying patients varies, as indicated by the Quality of Death Index.[4] From a clinical and research perspective, it is important to have valid and reliable measures to assess the current quality of care. One method of evaluation is to use validated tools and ask family members, or those deemed important to the deceased, about their experiences. An example of such a tool is the 'Care Of the Dying Evaluation' (CODE™) questionnaire. CODE™ is a shortened version of the original, validated instrument, 'Evaluating Care and Health Outcomes – for the Dying (ECHO-D).[5,6,7] Both tools are unique as their conceptual basis relates to the key components of best practice for 'care for the dying' in the last days of life.[8] CODE™ assesses the quality of patient care and the level of family support, through 32 main questions, reflecting core palliative care principles.[9] ~~The tool comprises 31 main questions about symptom control, medical and nursing care, emotional and spiritual support, communication, the provision of fluids, what to expect when an individual is dying and care at the time of death. It also has two questions related to patient dignity and respect and the level of family support.~~ Ten additional questions focus on demographic details.

CODE™ was initially validated in the United Kingdom (UK) within the community setting[10] and has subsequently been used across UK hospital, [11,12] hospice and home settings.[13] Within a systematic review of 67 tools used after death, CODE™ was one of the four recommended for use, based on initial psychometric properties.[14] Assessment of CODE™ in an international setting is therefore necessary to evaluate its robustness in the wider context of care for the dying.

The aim of this study was to evaluate the psychometric properties of CODE^{TM} in an international context. The two objectives were to assess the:

1. Construct validity of CODE^{TM} using factor analysis techniques and undertaking international country comparisons to evaluate any differences in perceptions.
2. Internal consistency of the tool.

Methods

~~Full~~ Details of study design and participants have been comprehensively described,[15] hence a summary is provided below. For clarity, ~~and simplicity, the~~ 'next-of-kin' is a collective term ~~referring to~~ for family members, friends and neighbours.

Participants

Respondents were the next-of-kin to adult cancer patients (≥ 18 years old) (~~with cancer)~~ who had died an 'expected' death within the study site hospitals (n=22) in seven South American and European countries: Argentina, Brazil, Germany, Norway, Poland, UK, and Uruguay. The patient must have been admitted to the hospital for at least three calendar days ~~(with the next-of-kin present at least some of the time during this period). If there was uncertainty about whether or not the death was 'expected', the attending physician was consulted; if this was not possible, any death that did not involve cardiopulmonary resuscitation was included.~~ The next-of-kin were eligible to complete the survey if aged ≥ 18 years, sufficiently fluent in the language, and able to provide informed consent. This was pragmatically assessed ~~in a multi-faceted way~~ by ward staff at the time of death and by research staff directly contacting potential participants to invite them to participate.

Instrument and development

Work was conducted to develop CODE^{TM} into an international tool (i-CODE) involving forward-and-back translation within each of the five different languages.[16] Pre-testing survey methods, involving

patient and public representatives and bereaved relatives, and including cognitive interviews, helped ensure good face and content validity.[16] (~~Minor language changes reflected different cultural context.~~) Consensus about the tool's content was reached using a modified nominal group technique.[17] This established a core, collective international version of CODE™ (i-CODE, Supplementary File 1) to use within the seven countries. Response options include Likert scale verbal anchors and ordinal responses where higher values represent better quality of care. As i-CODE contains all the questions from the original CODE™ questionnaire (with additional questions about advance care planning and the NHS Friends and Family test[18] being added by some countries), we use the terminology 'CODE™' within this paper.

Procedure

Data was collected between August 15th 2017 and September 15th 2018. A postal survey was planned but different approaches ~~(for data collection)~~ were adopted to reflect country-specific factors such as unreliable postal services and literacy issues. Poland, Argentina, Brazil and Uruguay undertook face-to-face or telephone interviews.

Screening for eligible cases was undertaken by the research team and information on the deceased patient's gender, age, cancer, (~~and details about their cancer,~~) length of hospital stay and place of death (type of hospital ward) as well as next-of-kin's gender and age group were collected.

Following screening for eligibility, the CODE™ questionnaire was sent or administered to next-of-kin 6-8 weeks after the patient's death. Responses were entered into a ~~Corporater Surveyor~~ database (~~**www.corporater.com**~~) and data were stored on a protected research server.

~~ClinicalTrials.gov Identifier: NCT03566732~~

Statistical Analysis

Correlation

Polychoric correlation was used to assess relationships of items prior to factor analysis; items with correlation coefficient ≥ 0.8 indicated ~~were reviewed to check for~~ local dependency. Item wording was also considered to identify any potential overlap. This informed which items to combine or drop in the analysis with supplementary analysis of the excluded questions.

Exploratory and Confirmatory factor analyses

Previous work on the dimensionality of CODE^{TM} has revealed different structures and included different items.[10,12] Therefore, we ~~first~~ carried out an Exploratory Factor Analysis (EFA) followed by Confirmatory Factor Analysis (CFA) ~~including all the questions~~, using Geomin rotation to identify potential factors present in the data. ~~Informed by the EFA, three-factor, four-factor, five-factor and six-factor Confirmatory Factor Analysis (CFA) models were estimated to allow the comparison of competing models.~~ EFA and CFA were performed treating the items as ordinal categorical, using the robust weighted least squares means and variance adjusted (WLSMV) estimator in Mplus 8.2.[19,20] Model fit was assessed by the Root Mean Square Error of Approximation (RMSEA)[21] and the Comparative Fit Index (CFI)[22] where a value of ≤.08 and >.95 was assumed to provide a good fit respectively. To obtain finer factor solutions, residual correlations and modification indices (MI) were inspected to identify potentially redundant items.[23,24] In the final models, local dependence between items ~~(where the latent variables are not sufficient to explain the association between items)~~ was introduced guided by:

i) the highest MI (>100)[24] or

ii) whether the pairs of items had been identified as conceptually similar, by constraining the pair of items as free parameter estimates in model revisions, one at a time.

Internal Reliability of the scale

Internal reliability was assessed using Cronbach alpha to quantify the extent to which the items were inter-related. Coefficients above 0.7 are acceptable, above 0.8 are good, and above 0.9 are excellent but above 0.94 suggests potential redundancy.[25]

Known-group validity

Known group validity was examined in terms of whether CODE^{TM} was able to discriminate between the different quality of care expected *a priori*. First, we hypothesised that the quality of care would be higher in hospital Palliative Care Units (PCUs) compared with other hospital wards. Second, we hypothesised that as most deaths in Poland occurred in hospital PCUs, ~~within hospitals,~~ we expected perceptions about quality of care to be higher. Differences were quantified using effect sizes (ES) across categories identified, calculated as the difference in mean scores between groups divided by the standard deviation of the lower quality of the two sub-groups. ES expressed as Cohen's *d* of 0.2 are normally considered small, 0.5 moderate, and 0.8 large.[26]

Testing for item invariance

Differential item functioning (DIF) is ~~said to be~~ present when participants with the same score level (quality of care, in this instance) endorse items differently by virtue of some characteristics other than the variation due to their scale score. The simple sum of the items in question was used as a proxy for the latent trait. Uniform DIF exists when the statistical relationship between item response and group is constant for all levels of a matching variable. Non-uniform DIF exists when there are differences among the groups for specific item responses.[27] DIF with regard to age and gender of patient and relative, ward and country was evaluated through ordinal logistic regression models.[28] Significant DIF was assessed through a dual criterion of statistical significance and a difference in explained variance (Nagelkerke pseudo R2) larger than 2%.[29] To assess the impact of the different

countries on the factor structure, we regressed the factors on country in the best performing CFA model.

Missing data

Missing data for the items was less than 5% and no data imputation was conducted.

Results

Response rate

From ~~In total,~~ 1683 potential cases ~~were~~ screened, ~~and~~ 914 CODE^{TM} questionnaires were completed (54% response rate) with at least 100 responses per country. ~~Each of the seven countries had at least 100 responses.~~

Demographics (Table 1)

The deceased patients tended to be male (527, 58%) and the most common primary cancer diagnoses were from the gastrointestinal tract (321, 35%) or respiratory system (196, 21%). Next-of-kin tended to be female (610, 67%) and the 'spouse/partner' (411, 45%) to the deceased individual. Further demographics ~~(and individual question responses)~~ have been detailed previously.[15]

Correlation results

High correlations (0.8-0.99) were observed for a number of questions and further examination of wording revealed considerable overlap (Table 2).

~~Very High correlations (0.8-0.99) were observed for the following questions: q1 and q2 (*personal/nursing care needs*); q17 and q18 (*discussion about fluids/whether discussion would have been helpful*); q21 and q22 (*spiritual needs of patient/relative*); q24 and q25 (*what to expect when dying/whether discussion would have been helpful*). Upon examining the wording of these items, we confirmed that this was an indication of local dependence. We dropped item 22 as the correlation~~

~~with q21 was 0.99. We kept the question pertaining to the spiritual needs of the patient rather than the family carer as the foremost responsibility of the healthcare organisation is to the patient. Replication of analyses with q22 did not change the results or conclusions, hence we do not report these analyses. We combined questions 17 and 18 and questions 24 and 25 as these were related questions. Further examination of the wording of the questionnaire suggests overlap in the following pairs of questions: 10 and 11 (*presence of pain/whether enough was done to control pain*), 12 and 13 *(presence of restlessness/whether enough was done to control restlessness*), 14 and 15 (*presence of 'noisy rattle'/whether enough was done to control 'noisy rattle'*). Therefore, we also carried out the analyses excluding questions 10, 12 and 14.~~

Exploratory Factor Analysis ~~(EFA)~~ results

Eigenvalue analysis identified a strong first factor and four or five weaker factors with values >1 (Supplementary Figure S1, Supplementary Table S1). ~~In the Geomin rotated five factor EFA, the fifth factor consisted of only two items pertaining to the religious and spiritual needs of the patient (q21) and relative (q22) being met and given the high correlation (0.99) between the two items, we excluded q22 resulting in a four-factor model (Supplementary Table S1).~~ EFA resulted in a four-factor model, namely: 'Overall care', 'Communication and support', 'Trust, respect and dignity', and 'Symptom management'.

Estimating four-factor and bi-factor CFA models

The four-factor model returned an acceptable fit (RMSEA = 0.069, CFI = 0.939). The standardized coefficients of the four factors ranged from 0.366 to 0.915 and were all statistically significant ($p < 0.001$). The correlations between factors ranged between 0.656 and 0.858 (Table 3). We extended the four-factor model to account for areas of strain within factor solutions through local correlations. Two pairs of items were allowed to correlate with each other (q1: ~~*There was enough help available to meet his/her)*~~ *personal care needs, (~~such as washing, personal hygiene and toileting needs)~~* and q2: (~~*There was enough help)*~~ *with nursing care needs, (~~such as giving medicines and helping~~*); q3: (~~*The bed*~~

~~*area and surrounding)*~~ *environment was comfortable* ~~*(for him/her)*~~ and q4: (~~*The bed area and surrounding)*~~ *environment had adequate privacy* ~~*(for him/her*~~). The model fit improved (RMSEA = 0.062; CFI = 0.951) with the factors still highly correlated (Table 3) leading us to estimate a bi-factor model. The model achieved acceptable fit (RMSEA = 0.057; CFI = 0.962) and the loadings for the global factor are higher than for the individual factors for the majority of items. Explained common variance (ECV) values were 77.9%, showing a strong global factor and suggesting a unidimensional model.[30]

Reliability results

The four-factor scale showed moderate to excellent reliability scores (Factor 1 'Overall care' α = 0.862; Factor 2 'Communication and support' α = 0.824; Factor 3 'Trust, respect and dignity' α = 0.618; Factor 4 'Symptom management' α = 0.796; Overall α = 0.922).

Known-group validity ~~***differences***~~

The ES for the quality of care in PCUs and for Poland were 0.727 and 0.657 respectively. In both cases, the results are as hypothesised, suggesting a higher quality of care in PCUs and in Poland compared with other countries.

Testing for item invariance

In tests of DIF with regard to age, gender, ward and country, several statistically significant instances of DIF were found (Table 4, Supplementary Tables S3a-S6). Uniform DIF was found for q4 for patient age ~~(of the patient)~~ and q27 and q29 for relative age ~~(of the relative).~~ There was uniform DIF for q5 for relative gender ~~(of the relative).~~ The largest number of instances of DIF was seen with regard to country. (~~In particular)~~, Perceptions about many aspects of care were often higher from Polish respondents, e.g., ~~including (These included) the questions asking about aspects of nursing care~~, symptom control and ~~being provided with sensitive~~ support after ~~the~~ death. Additionally, perceptions

about aspects (~~the questions relating to~~) of nursing care in Argentina had lower mean scores (Table 5).

Discussion

Within an international context, CODE™ was found to be valid and reliable in assessing the quality of care provided in the last days of life for those dying from cancer in hospital. Assessment of construct validity identified that a bi-factor model with four distinct factors – 'overall care', 'communication and support', 'trust, respect and dignity', and 'symptom management' - provided the best model fit. For 27 of the 32 questions, the factor loadings were substantially higher on the general factor than on the group factors. Each of the four identified factors had good construct validity and internal consistency. Additionally, our *a priori* hypotheses were supported by our findings: CODE was sufficiently sensitive to detect differences in perceptions of care between countries; and the quality of care provided within PCUs was perceived to be greater compared with other ward settings. This study further builds on the quality of psychometric evidence[14] for CODE™, namely for: internal consistency, hypothesis testing and cross-cultural properties.

The original structure of CODE™ (~~the measure~~) is around three composite scales 'Environment', 'Care' and 'Communication'. These scales did not include all (~~the CODE™~~) question items but were based upon theoretical assumptions developed and validated from the 'ECHO-D' questionnaire.[10] (~~Environment was not a distinct factor in our findings~~). Vogt et al 2020 found a 7-factor model based on principal component analysis of a selected 28 'core items' with some items loading on more than one factor.[12] These differences (~~with the reported findings~~) may be due to methods or samples used and the items included in the final models. Our four identified factors represent meaningful concepts reflecting principles of holistic palliative care. Previous studies have concluded that scales with similar ECV values (77.9%) are sufficiently unidimensional.[31,32] Within a recognised multi-faceted concept, ~~the general factor reflected here is the overall quality of care for the dying.~~

unidimensionality implies that an overall score could be calculated for 'quality of care for the dying' ~~CODE™~~ should this be desired.

One of the key objectives was to undertake international country comparisons to evaluate for differences in perceptions, which was observed for nine question items. In this context, significant DIF suggests that respondents are answering the questions differently by virtue of being in different countries. These differences may be a tool artefact (~~of the tool~~) or may be picking up real differences in perceptions about the quality of care in ~~(the)~~ different countries. For example, within the Argentinian ~~(hospital)~~ study sites, issues relating to inadequate nursing numbers ~~(of nurses)~~ may have influenced perceptions about these aspects of care.[33] This compares with findings from Poland where most deaths occurred within hospital PCU's and perceptions about nursing care were higher. ~~As most of the deaths in Poland occurred within hospital PCUs, the level of skills and experience of nursing staff are likely to be higher.~~ In these instances, we can conclude that DIF may not actually be problematic but reflects true differences in perceptions.

Almost 70 tools have been identified to assess quality of death, dying and care (~~care at the end of life, and quality of dying and death)~~ with variable levels of use and psychometric assessment.[14] ~~The~~ Choice of tool is ~~will be~~ influenced by many factors including study purpose and setting, the measure's validity and reliability (~~of the measure), study setting~~ and ease of use. CODE™ as a tool has certain strengths. The content and format are acceptable and relevant across several countries reflecting different cultures. CODE™ has the potential to assess interventions used within the last days of life, as well as ~~potentially being one of the outcomes assessing interventions~~ those used earlier in the disease trajectory. Additionally, CODE™ can ~~be used to~~ facilitate quality improvement work, with direct user-feedback helping inform clinical practice (~~findings coming straight from users and having direct clinical relevance and impact~~).[34]

For statistical purposes, we have dropped some items. While this does not necessarily mean that these items are not helpful, there may be scope to create a shortened version of CODE™. Shortened

tools can be a more feasible way to assess quality of care for larger populations[35], may reduce completion burden, and encourage higher response rates.

~~In this paper, we combined and dropped some question items based on local correlations. While this was necessary for statistical purposes, it does not necessarily mean that these items are not helpful. (Indeed, some items have important clinical context (e.g., reflecting whether spiritual needs are met for both the patient and family) or help the logical flow of CODE™.) The high correlations, however, suggest that there may be scope to create a shortened version of CODE™. Shortened tools can be a more feasible way to assess quality of care (in the last days of life) for larger populations[33], and may help reduce completion the burden, and encourage (of completion encouraging) higher response rates.~~

Our study has several limitations. It ~~The study~~ was only conducted with the next-of-kin who had experienced a hospital death related to cancer (~~although CODE™ has been used previously in different care settings and evaluating deaths from non-malignant disease~~). We could not conduct a multilevel EFA and CFA because there were only seven 'clusters' (countries) and a minimum of 30-50 is required for such a technique. Additional psychometric work using an independent dataset, reflecting a different population is needed to undertake further CFA of the constructs identified within this study. Future work will include testing (~~other psychometric properties such as~~) inter-rater reliability and the concurrent and predictive validity of CODE™. Additionally, using CODE™ with other instrument(s) is needed to assess criterion validity, although a pragmatic choice is required as no current 'gold standard' tool exists.

Conclusion

This study provides good evidence for the validity and reliability of CODE™ within several different countries representing different cultural contexts. Using CODE™ provides part of the means to ensure ~~areas of~~ unmet needs are recognised and that efforts are made to ensure quality of care provided to dying patients and their families is at the level of the best.

Disclosures

We have no conflicts of interest to declare.

Acknowledgements

We wish to thank all the bereaved family members and friends for taking their time to complete the CODE™ questionnaire and sharing their views. Additionally, we wish to thank our wider ERANet-LAC CODE Project Group:

Core scientific group: Katrin Ruth Sigurdardottir, Marit Irene Tuen Hansen, Karl Ove Hufthammer, Wojciech Leppert, Katarzyna Wolszczak, Eduardo Garcia Yanneo, Vilma Tripodoro, Gabriel Goldraij, Martin Weber, Christina Gerlach, Lair Zambon, Juliana Nalin Passarini, Ivete Bredda Saad, John Ellershaw, and Grace Ting.

Local study coordinators: Anne Kvikstad, Eva Gravdahl, Julia Bratke, Janet Bakken, Kristin Vassbotn Guldhav, Christina Demmerle.

We also thank our colleagues within the International Collaborative for Best Care for the Dying Person and the White Rose collaboration for valuable discussions and support.

Funding

Funding for this study was provided by the White Rose Collaborative Fund https://whiterose.ac.uk/collaborationfunds/measuring-evaluating-and-improving-the-quality-of-care-for-dying-people-and-their-families/

And through the ERANet-LAC CODE project: “International Care Of the Dying Evaluation (CODE): quality of care for cancer patients as perceived by bereaved relatives” (reference ELAC2015/T07–0545, January 2017–January 2020) as one of the successful proposals of the second Joint Call for Transnational Research and/or Innovation Projects within the ERANet-LAC framework, cofunded by

the European Commission's seventh Framework Programme, with the overall aims to improve the quality of care and quality of life of patients dying from cancer.

The following national funding agencies are acknowledged: Ministry for Science, Technology and Productive Innovation, Argentina; Research Support Foundation of the State of Sao Paulo, Brazil; Federal Ministry of Education and Research, Germany; The Research Council of Norway (grant number 271051); National Centre for Research and Development, Poland; and Marie Curie and the Northwest Coast Clinical Research Network Funding, U.K. Dr. Catriona Mayland is funded by Yorkshire Cancer Research. The funders of the study played no role in the study design, data collection, data analysis, data interpretation, or writing of the report.

Ethics and research governance

Study approval was given by the Regional Committee for Medical and Health Research Ethics West (2017/640/REK vest), Norway; Ethics Committee of the Medical Council of the province of Rhineland-Palatine, Germany (approval nos. 837.331.13(901 6F) and 837.292.17 (111261); Guía de Buenas Prácticas de Investigación Clínica en Seres Humanos, Ministerio de Salud de la Nación Argentina (Resolución 1480/2011); Bioethics Committee, Poland - KB507/2017 (13.06.17); Health Research Authority and East of England – Cambridge East Research Ethics Committee, U.K. (IRAS project ID 225922; REC 17/EE/0302); National Research Ethics Commission, Brazil (ref. 2308.216); and Committee on Bioethics from the “Mutualista Asociacion Hospital Evangelico,” Uruguay (ref. 29/5/2017/01).

The study was registered on ClinicalTrials.gov -Identifier: NCT03566732

References

1. Yamaguchi T, Maeda I, Hatano Y, et al. Effects of End-of-Life Discussions on the Mental Health of Bereaved Family Members and Quality of Patient Death and Care. J Pain Symptom Manage. 2017 Jul;54(1):17-26.e1. doi: 10.1016/j.jpainsymman.2017.03.008. Epub 2017 Apr 24. PMID: 28450216.
2. Wright AA, Zhang B, Ray A, et al. Associations between end-of-life discussions, patient mental health, medical care near death, and caregiver bereavement adjustment. JAMA. 2008 Oct 8;300(14):1665-73. doi: 10.1001/jama.300.14.1665. PMID: 18840840; PMCID: PMC2853806.
3. Miyajima K, Fujisawa D, Yoshimura K, et al. Association between quality of end-of-life care and possible complicated grief among bereaved family members. J Palliat Med. 2014 Sep;17(9):1025-31. doi: 10.1089/jpm.2013.0552. Epub 2014 Jul 22. PMID: 25050607.
4. The Economist Intelligence Unit . The 2015 quality of death index. Lien Foundation. https://www.eiuperspectives.economist.com/sites/default/files/2015%20EIU%20Quality%20of%20Death%20Index%20Oct%2029%20FINAL.pdf (accessed 16.11.2021).
5. Mayland CR, Williams EM, Ellershaw JE. Assessing quality of care for the dying: the development and initial validation of a postal self-completion questionnaire for bereaved relatives. Palliat Med. 2012 Oct;26(7):897-907. doi: 10.1177/0269216311424953. Epub 2011 Nov 1. PMID: 22045726.
6. Mayland CR, Williams EM, Addington-Hall J, et al. Does the 'Liverpool Care Pathway' facilitate an improvement in quality of care for dying cancer patients?. Br J Cancer. 2013;108(10):1942-1948. doi:10.1038/bjc.2013.203.
7. Mayland CR, Williams EM, Addington-Hall J, et al. Assessing the quality of care for dying patients from the bereaved relatives' perspective: further validation of "Evaluating care and health outcomes--for the dying". J Pain Symptom Manage. 2014 Apr;47(4):687-96. doi: 10.1016/j.jpainsymman.2013.05.013. Epub 2013 Nov 5. PMID: 24210451.

8. Mayland C, Williams E, Ellershaw J. How well do current instruments using bereaved relatives' views evaluate care for dying patients? Palliat Med. 2008 Mar;22(2):133-44. doi: 10.1177/0269216307085742. PMID: 18372378.
9. World Health Organization. Palliative care. 2020. Available at https://www.who.int/news-room/fact-sheets/detail/palliative-care (accessed 21.08.2021).
10. 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 Jun;4(2):167-174. doi: 10.1136/bmjspcare-2013-000596. Epub 2014 Mar 28. PMID: 24681559.
11. Mayland CR, Mulholland H, Gambles M, et al How well do we currently care for our dying patients in acute hospitals: the views of the bereaved relatives? BMJ Supportive & Palliative Care 2017;**7**:316-325.
12. Vogt A, Stiel S, Heckel M, 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 Sep 22;18(1):311. doi: 10.1186/s12955-020-01473-2. PMID: 32962706; PMCID: PMC7507719.
13. Mayland C, McGlinchey T, Gambles M, et al. Quality assurance for care of the dying: engaging with clinical services to facilitate a regional cross-sectional survey of bereaved relatives' views. BMC Health Serv Res. 2018 Oct 10;18(1):761. doi: 10.1186/s12913-018-3558-z. PMID: 30305082; PMCID: PMC6180653.
14. Kupeli N, Candy B, Tamura-Rose G, et al. Tools Measuring Quality of Death, Dying, and Care, Completed after Death: Systematic Review of Psychometric Properties. Patient. 2019 Apr;12(2):183-197. doi: 10.1007/s40271-018-0328-2. PMID: 30141020; PMCID: PMC6397142.
15. Haugen DF, Hufthammer KO, Gerlach C, et al; ERANet-LAC CODE Project Group. Good Quality Care for Cancer Patients Dying in Hospitals, but Information Needs Unmet: Bereaved

Relatives' Survey within Seven Countries. Oncologist. 2021 Jul;26(7):e1273-e1284. doi: 10.1002/onco.13837. Epub 2021 Jun 17. PMID: 34060705; PMCID: PMC8265351.

16. 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 Mar;33(3):357-368. doi: 10.1177/0269216318818299. Epub 2019 Jan 10. PMID: 30628867; PMCID: PMC6376606.
17. Jones, J, Hunter, D. Consensus methods for medical and health services research. BMJ 1995; 311: 376–380.
18. National Health Service. Friends and family test. https://www.england.nhs.uk/fft/ (accessed 25.09.2021).
19. Flora DB, Curran PJ. An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. Psychol Methods. 2004 Dec;9(4):466-91. doi: 10.1037/1082-989X.9.4.466. PMID: 15598100; PMCID: PMC3153362.
20. Muthen LK, Muthen BO. Mplus User's Guide (Sixth ed.) Los Angeles, CA: Muthen & Muthen. 1998-2011.
21. MacCallum RC, Browne MW, Sugawara HM. Power analysis and determination of sample size for covariance structure modeling. Psychological Methods 1996; 1(2): 130–149. https://doi.org/10.1037/1082-989X.1.2.130.
22. Hu L-t, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Structural Equation Modeling. 1999: 6(1): 1–55. https://doi.org/10.1080/10705519909540118.
23. Schreiber JB, Stage FK, King J, et al. Reporting Structural Equation Modeling and Confirmatory Factor Analysis Results: A Review. The Journal of Educational Research. 2006; *99*(6): 323–337. https://doi.org/10.3200/JOER.99.6.323-338.

24. Watt T, Groenvold M, Deng N, et al. Confirmatory factor analysis of the thyroid-related quality of life questionnaire ThyPRO. Health Qual Life Outcomes. 2014 Sep 10;12:126. doi: 10.1186/s12955-014-0126-z. PMID: 25213857; PMCID: PMC4172819.
25. Fayers PM, Machin D. Quality of life: the assessment, analysis and interpretation of patient-reported outcomes. John Wiley & Sons, 2013.
26. Cohen J. Statistical power analysis for the behavior science. Lawrance Eribaum Association. 1988.
27. Hanson BA. Uniform DIF and DIF Defined by Differences in Item Response Functions. Journal of Educational and Behavioral Statistics. 1998;23 (3):244-253. doi:10.3102/10769986023003244
28. Zumbo BD. A handbook on the theory and methods of differential item functioning (DIF). Ottawa: National Defense Headquarters: 1999.
29. Nagelkerke NJ. A note on a general definition of the coefficient of determination. Biometrika 1991; 78(3): 691-2.
30. Reise, S. P., Moore, T. M., & Haviland, M. G. Bifactor models and rotations: Exploring the extent to which multidimensional data yield univocal scale scores. Journal of Personality Assessment. 2010; 92(6): 544–559.
31. Reise SP, Bonifay WE, Haviland MG. Scoring and modeling psychological measures in the presence of multidimensionality. J Pers Assess. 2013;95(2):129-40. doi: 10.1080/00223891.2012.725437. Epub 2012 Oct 2. PMID: 23030794.
32. Keetharuth AD, Bjorner JB, Barkham M, et al. Exploring the item sets of the Recovering Quality of Life (ReQoL) measures using factor analysis. Qual Life Res. 2019 Apr;28(4):1005-1015. doi: 10.1007/s11136-018-2091-1. Epub 2018 Dec 21. PMID: 30578454; PMCID: PMC6439178.

33. Rubinstein A, Zerbino MC, Cejas C, et al. Making Universal Health Care Effective in Argentina: A Blueprint for Reform. Health Syst Reform 2018;4(3):203-13. doi: 10.1080/23288604.2018.1477537 [published Online First: 2018/08/02]
34. Goldraij G, Tripodoro VA, Aloisio M The ERANet-LAC CODE Project Group, et al. One chance to get it right: improving clinical handovers for better symptom control at the end of life. BMJ Open Quality 2021;10:e001436. doi: 10.1136/bmjoq-2021-001436.
35. Pérez-Cruz PE, Padilla Pérez O, Bonati P, et al. Validation of the Spanish Version of the Quality of Dying and Death Questionnaire (QODD-ESP) in a Home-Based Cancer Palliative Care Program and Development of the QODD-ESP-12. J Pain Symptom Manage. 2017 Jun;53(6):1042-1049.e3. doi: 10.1016/j.jpainsymman.2017.02.005. Epub 2017 Mar 16. PMID: 28323080.

Title page

Validation of 'Care Of the Dying Evaluation' (CODE^{TM}) within an international study exploring bereaved relatives' perceptions about quality of care in the last days of life

Catriona R Mayland, MBChB, MD, Yorkshire Cancer Research Senior Clinical Research Fellow and Honorary Consultant in Palliative Medicine, University of Sheffield; Honorary Clinical Fellow, University of Liverpool; ORCID ID 0000-0002-1440-9953

Anju D Keetharuth, PhD, School of Health and Related Research, University of Sheffield, Sheffield, UK

Clara Mukuria, PhD, School of Health and Related Research, University of Sheffield, Sheffield, UK

Dagny Faksvåg Haugen, MD, PhD, Regional Centre of Excellence for Palliative Care, Western Norway, Haukeland University Hospital, and Department of Clinical Medicine K1, University of Bergen, Bergen, Norway

Corresponding author:

Dr Catriona Mayland, Department of Oncology and Metabolism, 2nd Floor Broomcross Building, Weston Park Hospital, Whitham Road, Sheffield S10 2SJ

E-mail: c.r.mayland@sheffield.ac.uk

Twitter: @CattyRM

Number tables: 5

Number figures: 0

Number of references: 35

Word count: 2517

Supplementary files: 5

Abstract (word count 250)

Context: Assessing quality of care provided during the dying phase using validated tools aids quality assurance and recognises unmet need.

Objective: To assess construct validity and internal consistency of 'Care Of the Dying Evaluation' (CODE^{TM}) within an international context.

Methods: Post-bereavement survey (August 2017-September 2018) using CODE^{TM}. Respondents were next-of-kin to adult patients (≥ 18 years old) with cancer who had an 'expected' death within 22 study site hospitals in 7 countries: Argentina, Brazil, Germany, Norway, Poland, United Kingdom, Uruguay. Exploratory and Confirmatory Factor Analysis (EFA and CFA) were conducted, and internal reliability was assessed using Cronbach alpha (α). Known group validity was assessed by ability to discriminate quality of care based in place (Palliative Care Units (PCUs)) and country (Poland, where most deaths were in PCUs) of care. Differences were quantified using effect sizes (ES).

Results: 914 CODE^{TM} questionnaires completed (54% response rate). 527 (58%) male deceased patients; 610 (67%) next-of-kin female who were most commonly the 'spouse/partner' (411, 45%). EFA identified 4 factors: 'Overall care', 'Communication and support', 'Trust, respect and dignity', and 'Symptom management' with good reliability scores (α = 0.628 – 0.862). CFA confirmed the 4-factor model; these were highly correlated and a bifactor model showed acceptable fit. The ES for quality of care in PCU's was 0.727; ES for Poland was 0.657, supporting the sensitivity of CODE^{TM} to detect differences.

Conclusion: Within an international context, good evidence supports the validity and reliability of CODE^{TM} for assessing the quality of care provided in the last days of life.

Key message

This article describes the validation of the 'Care Of the Dying Evaluation' (CODE™) questionnaire within an international study. The results indicate that CODE™ represents a valid and reliable tool for assessing the quality of care provided to dying patients and their families. (42 words)

Key words: palliative care; terminal care; psychometrics; factor analysis; quality of care; quality of death

Running title: International validation of CODE™

Introduction

Ensuring dying patients and their families receive high quality care and support is fundamentally important. How an individual dies has a profound impact on those bereaved. Timely, informative end-of-life discussions are associated with less complicated grief.[1] Conversely, overly aggressive medical interventions and dissatisfaction with communication are linked with poorer bereavement outcomes.[2,3]

Globally, the provision of care for dying patients varies, as indicated by the Quality of Death Index.[4] From a clinical and research perspective, it is important to have valid and reliable measures to assess the current quality of care. One method of evaluation is to use validated tools and ask family members, or those deemed important to the deceased, about their experiences. An example of such a tool is the 'Care Of the Dying Evaluation' (CODE™) questionnaire. CODE™ is a shortened version of the original, validated instrument, 'Evaluating Care and Health Outcomes – for the Dying (ECHO-D).[5,6,7] Both tools are unique as their conceptual basis relates to the key components of best practice for 'care for the dying' in the last days of life.[8] CODE™ assesses the quality of patient care and the level of family support, through 32 main questions, reflecting core palliative care principles.[9] Ten additional questions focus on demographic details.

CODE™ was initially validated in the United Kingdom (UK) within the community setting[10] and has subsequently been used across UK hospital,[11,12] hospice and home settings.[13] Within a systematic review of 67 tools used after death, CODE™ was one of the four recommended for use, based on initial psychometric properties.[14] Assessment of CODE™ in an international setting is therefore necessary to evaluate its robustness in the wider context of care for the dying.

The aim of this study was to evaluate the psychometric properties of CODE™ in an international context. The two objectives were to assess the:

1. Construct validity of CODE^{TM} using factor analysis techniques and undertaking international country comparisons to evaluate any differences in perceptions.
2. Internal consistency of the tool.

Methods

Details of study design and participants have been comprehensively described,[15] hence a summary is provided below. For clarity, 'next-of-kin' is a collective term for family members, friends and neighbours.

Participants

Respondents were the next-of-kin to adult cancer patients (≥ 18 years old) who had died an 'expected' death within the study site hospitals (n=22) in seven South American and European countries: Argentina, Brazil, Germany, Norway, Poland, UK, and Uruguay. The patient must have been admitted to the hospital for at least three calendar days. The next-of-kin were eligible to complete the survey if aged ≥ 18 years, sufficiently fluent in the language, and able to provide informed consent. This was pragmatically assessed by ward staff at the time of death and by research staff directly contacting potential participants to invite them to participate.

Instrument and development

Work was conducted to develop CODE^{TM} into an international tool (i-CODE) involving forward-and-back translation within each of the five different languages.[16] Pre-testing survey methods, involving patient and public representatives and bereaved relatives, and including cognitive interviews, helped ensure good face and content validity.[16] Consensus about the tool's content was reached using a modified nominal group technique.[17] This established a core, collective international version of CODE^{TM} (i-CODE, Supplementary File 1) to use within the seven countries. Response options include Likert scale verbal anchors and ordinal responses where higher values represent better quality of

care. As i-CODE contains all the questions from the original CODE™ questionnaire (with additional questions about advance care planning and the NHS Friends and Family test[18] being added by some countries), we use the terminology 'CODE™' within this paper.

Procedure

Data was collected between August 15th 2017 and September 15th 2018. A postal survey was planned but different approaches were adopted to reflect country-specific factors such as unreliable postal services and literacy issues. Poland, Argentina, Brazil and Uruguay undertook face-to-face or telephone interviews.

Screening for eligible cases was undertaken by the research team and information on the deceased patient's gender, age, cancer, length of hospital stay and place of death (type of hospital ward) as well as next-of-kin's gender and age group were collected.

Following screening for eligibility, the CODE™ questionnaire was sent or administered to next-of-kin 6-8 weeks after the patient's death. Responses were entered into a database and data were stored on a protected research server.

Statistical Analysis

Correlation

Polychoric correlation was used to assess relationships of items prior to factor analysis; items with correlation coefficient ≥ 0.8 indicated local dependency. Item wording was also considered to identify any potential overlap. This informed which items to combine or drop in the analysis with supplementary analysis of the excluded questions.

Exploratory and Confirmatory factor analyses

Previous work on the dimensionality of CODE™ has revealed different structures and included different items.[10,12] Therefore, we carried out an Exploratory Factor Analysis (EFA) followed by

Confirmatory Factor Analysis (CFA), using Geomin rotation to identify potential factors present in the data. EFA and CFA were performed treating the items as ordinal categorical, using the robust weighted least squares means and variance adjusted (WLSMV) estimator in Mplus 8.2.[19,20] Model fit was assessed by the Root Mean Square Error of Approximation (RMSEA)[21] and the Comparative Fit Index (CFI)[22] where a value of ≤.08 and >.95 was assumed to provide a good fit respectively. To obtain finer factor solutions, residual correlations and modification indices (MI) were inspected to identify potentially redundant items.[23,24] In the final models, local dependence between items was introduced guided by:

i) the highest MI (>100)[24] or
ii) whether the pairs of items had been identified as conceptually similar, by constraining the pair of items as free parameter estimates in model revisions, one at a time.

Internal Reliability of the scale

Internal reliability was assessed using Cronbach alpha to quantify the extent to which the items were inter-related. Coefficients above 0.7 are acceptable, above 0.8 are good, and above 0.9 are excellent but above 0.94 suggests potential redundancy.[25]

Known-group validity

Known group validity was examined in terms of whether CODE™ was able to discriminate between the different quality of care expected *a priori*. First, we hypothesised that the quality of care would be higher in hospital Palliative Care Units (PCUs) compared with other hospital wards. Second, we hypothesised that as most deaths in Poland occurred in hospital PCUs, we expected perceptions about quality of care to be higher. Differences were quantified using effect sizes (ES) across categories identified, calculated as the difference in mean scores between groups divided by the standard deviation of the lower quality of the two sub-groups. ES expressed as Cohen's *d* of 0.2 are normally considered small, 0.5 moderate, and 0.8 large.[26]

Testing for item invariance

Differential item functioning (DIF) is present when participants with the same score level (quality of care, in this instance) endorse items differently by virtue of some characteristics other than the variation due to their scale score. The simple sum of the items in question was used as a proxy for the latent trait. Uniform DIF exists when the statistical relationship between item response and group is constant for all levels of a matching variable. Non-uniform DIF exists when there are differences among the groups for specific item responses.[27] DIF with regard to age and gender of patient and relative, ward and country was evaluated through ordinal logistic regression models.[28] Significant DIF was assessed through a dual criterion of statistical significance and a difference in explained variance (Nagelkerke pseudo R2) larger than 2%.[29] To assess the impact of the different countries on the factor structure, we regressed the factors on country in the best performing CFA model.

Missing data

Missing data for the items was less than 5% and no data imputation was conducted.

Results

Response rate

From 1683 potential cases screened, 914 CODE^{TM} questionnaires were completed (54% response rate) with at least 100 responses per country.

Demographics (Table 1)

The deceased patients tended to be male (527, 58%) and the most common primary cancer diagnoses were from the gastrointestinal tract (321, 35%) or respiratory system (196, 21%). Next-of-

kin tended to be female (610, 67%) and the 'spouse/partner' (411, 45%) to the deceased individual. Further demographics have been detailed previously.[15]

Correlation results

High correlations (0.8-0.99) were observed for a number of questions and further examination of wording revealed considerable overlap (Table 2).

Exploratory Factor Analysis results

Eigenvalue analysis identified a strong first factor and four or five weaker factors with values >1 (Supplementary Figure S1, Supplementary Table S1/S2). EFA resulted in a four-factor model, namely: 'Overall care', 'Communication and support', 'Trust, respect and dignity', and 'Symptom management'.

Estimating four-factor and bi-factor CFA models

The four-factor model returned an acceptable fit (RMSEA = 0.069, CFI = 0.939). The standardized coefficients of the four factors ranged from 0.366 to 0.915 and were all statistically significant ($p < 0.001$). The correlations between factors ranged between 0.656 and 0.858 (Table 3). We extended the four-factor model to account for areas of strain within factor solutions through local correlations. Two pairs of items were allowed to correlate with each other: q1 (*personal care needs)* and q2 (*nursing care needs*); q3 (*environment was comfortable)* and q4: (*environment had adequate privacy*). The model fit improved (RMSEA = 0.062; CFI = 0.951) with the factors still highly correlated (Table 3) leading us to estimate a bi-factor model. The model achieved acceptable fit (RMSEA = 0.057; CFI = 0.962) and the loadings for the global factor are higher than for the individual factors for the majority of items. Explained common variance (ECV) values were 77.9%, showing a strong global factor and suggesting a unidimensional model.[30]

Reliability results

The four-factor scale showed moderate to excellent reliability scores (Factor 1 'Overall care' α = 0.862; Factor 2 'Communication and support' α = 0.824; Factor 3 'Trust, respect and dignity' α = 0.618; Factor 4 'Symptom management' α = 0.796; Overall α = 0.922).

Known-group validity

The ES for the quality of care in PCUs and for Poland were 0.727 and 0.657 respectively. In both cases, the results are as hypothesised, suggesting a higher quality of care in PCUs and in Poland compared with other countries.

Testing for item invariance

In tests of DIF with regard to age, gender, ward and country, several statistically significant instances of DIF were found (Table 4, Supplementary Tables S3a-S6). Uniform DIF was found for q4 for patient age and q27 and q29 for relative age. There was uniform DIF for q5 for relative gender. The largest number of instances of DIF was seen with regard to country. Perceptions about many aspects of care were often higher from Polish respondents e.g. symptom control, support after death. Additionally, perceptions about aspects of nursing care in Argentina had lower mean scores (Table 5).

Discussion

Within an international context, CODE^{TM} was found to be valid and reliable in assessing the quality of care provided in the last days of life for those dying from cancer in hospital. Assessment of construct validity identified that a bi-factor model with four distinct factors – 'overall care', 'communication and support', 'trust, respect and dignity', and 'symptom management' - provided the best model fit. For 27 of the 32 questions, the factor loadings were substantially higher on the general factor than on the group factors. Each of the four identified factors had good construct validity and internal consistency. Additionally, our *a priori* hypotheses were supported by our findings: CODE was

sufficiently sensitive to detect differences in perceptions of care between countries; and the quality of care provided within PCUs was perceived to be greater compared with other ward settings. This study further builds on the quality of psychometric evidence[14] for CODE^{TM}, namely for: internal consistency, hypothesis testing and cross-cultural properties.

The original structure of CODE^{TM} is around three composite scales 'Environment', 'Care' and 'Communication'. These scales did not include all question items but were based upon theoretical assumptions developed and validated from the 'ECHO-D' questionnaire.[10] Vogt et al 2020 found a 7-factor model based on principal component analysis of a selected 28 'core items' with some items loading on more than one factor.[12] These differences may be due to methods or samples used and the items included in the final models. Our four identified factors represent meaningful concepts reflecting principles of holistic palliative care. Previous studies have concluded that scales with similar ECV values (77.9%) are sufficiently unidimensional.[31,32] Within a recognised multi-faceted concept, unidimensionality implies that an overall score for 'quality of care for the dying' could be calculated for CODE^{TM} should this be desired.

One of the key objectives was to undertake international country comparisons to evaluate for differences in perceptions, which was observed for nine question items. In this context, significant DIF suggests that respondents are answering the questions differently by virtue of being in different countries. These differences may be a tool artefact or may be picking up real differences in perceptions about the quality of care in different countries. For example, within the Argentinian study sites, issues relating to inadequate nursing numbers may have influenced perceptions about these aspects of care.[33] This compares with findings from Poland where most deaths occurred within hospital PCUs and perceptions about nursing care were higher. In these instances, we can conclude that DIF may not actually be problematic but reflects true differences in perceptions.

Almost 70 tools have been identified to assess quality of death, dying and care with variable levels of use and psychometric assessment.[14] Choice of tool is influenced by many factors including study

purpose and setting, the measure's validity, reliability, and ease of use. CODE™ as a tool has certain strengths. The content and format are acceptable and relevant across several countries reflecting different cultures. CODE™ has the potential to assess interventions used within the last days of life, as well as those used earlier in the disease trajectory. Additionally, CODE™ can facilitate quality improvement work, with direct user-feedback helping inform clinical practice.[34]

For statistical purposes, we have dropped some items. While this does not necessarily mean that these items are not helpful, there may be scope to create a shortened version of CODE™. Shortened tools can be a more feasible way to assess quality of care for larger populations[35], may reduce completion burden, and encourage higher response rates.

Our study has several limitations. It was only conducted with the next-of-kin who had experienced a hospital death related to cancer. We could not conduct a multilevel EFA and CFA because there were only seven 'clusters' (countries) and a minimum of 30-50 is required for such a technique. Additional psychometric work using an independent dataset, reflecting a different population is needed to undertake further CFA of the constructs identified within this study. Future work will include testing inter-rater reliability and the concurrent and predictive validity of CODE™. Additionally, using CODE™ with other instrument(s) is needed to assess criterion validity, although a pragmatic choice is required as no current 'gold standard' tool exists.

Conclusion

This study provides good evidence for the validity and reliability of CODE™ within several different countries representing different cultural contexts. Using CODE™ provides part of the means to ensure unmet needs are recognised and that efforts are made to ensure quality of care provided to dying patients and their families is at the level of the best.

Disclosures

We have no conflicts of interest to declare.

Acknowledgements

We wish to thank all the bereaved family members and friends for taking their time to complete the CODE^{TM} questionnaire and sharing their views. Additionally, we wish to thank our wider ERANet-LAC CODE Project Group:

Core scientific group: Katrin Ruth Sigurdardottir, Marit Irene Tuen Hansen, Karl Ove Hufthammer, Wojciech Leppert, Katarzyna Wolszczak, Eduardo Garcia Yanneo, Vilma Tripodoro, Gabriel Goldraij, Martin Weber, Christina Gerlach, Lair Zambon, Juliana Nalin Passarini, Ivete Bredda Saad, John Ellershaw, and Grace Ting.

Local study coordinators: Anne Kvikstad, Eva Gravdahl, Julia Bratke, Janet Bakken, Kristin Vassbotn Guldhav, Christina Demmerle.

We also thank our colleagues within the International Collaborative for Best Care for the Dying Person and the White Rose collaboration for valuable discussions and support.

Funding

Funding for this study was provided by the White Rose Collaborative Fund https://whiterose.ac.uk/collaborationfunds/measuring-evaluating-and-improving-the-quality-of-care-for-dying-people-and-their-families/

And through the ERANet-LAC CODE project: “International Care Of the Dying Evaluation (CODE): quality of care for cancer patients as perceived by bereaved relatives” (reference ELAC2015/T07–0545, January 2017–January 2020) as one of the successful proposals of the second Joint Call for Transnational Research and/or Innovation Projects within the ERANet-LAC framework, cofunded by

the European Commission's seventh Framework Programme, with the overall aims to improve the quality of care and quality of life of patients dying from cancer.

The following national funding agencies are acknowledged: Ministry for Science, Technology and Productive Innovation, Argentina; Research Support Foundation of the State of Sao Paulo, Brazil; Federal Ministry of Education and Research, Germany; The Research Council of Norway (grant number 271051); National Centre for Research and Development, Poland; and Marie Curie and the Northwest Coast Clinical Research Network Funding, U.K. Dr. Catriona Mayland is funded by Yorkshire Cancer Research. The funders of the study played no role in the study design, data collection, data analysis, data interpretation, or writing of the report.

Ethics and research governance

Study approval was given by the Regional Committee for Medical and Health Research Ethics West (2017/640/REK vest), Norway; Ethics Committee of the Medical Council of the province of Rhineland-Palatine, Germany (approval nos. 837.331.13(901 6F) and 837.292.17 (111261); Guía de Buenas Prácticas de Investigación Clínica en Seres Humanos, Ministerio de Salud de la Nación Argentina (Resolución 1480/2011); Bioethics Committee, Poland - KB507/2017 (13.06.17); Health Research Authority and East of England – Cambridge East Research Ethics Committee, U.K. (IRAS project ID 225922; REC 17/EE/0302); National Research Ethics Commission, Brazil (ref. 2308.216); and Committee on Bioethics from the “Mutualista Asociacion Hospital Evangelico,” Uruguay (ref. 29/5/2017/01).

The study was registered on ClinicalTrials.gov -Identifier: NCT03566732

References

1. Yamaguchi T, Maeda I, Hatano Y, et al. Effects of End-of-Life Discussions on the Mental Health of Bereaved Family Members and Quality of Patient Death and Care. J Pain Symptom Manage. 2017 Jul;54(1):17-26.e1. doi: 10.1016/j.jpainsymman.2017.03.008. Epub 2017 Apr 24. PMID: 28450216.
2. Wright AA, Zhang B, Ray A, et al. Associations between end-of-life discussions, patient mental health, medical care near death, and caregiver bereavement adjustment. JAMA. 2008 Oct 8;300(14):1665-73. doi: 10.1001/jama.300.14.1665. PMID: 18840840; PMCID: PMC2853806.
3. Miyajima K, Fujisawa D, Yoshimura K, et al. Association between quality of end-of-life care and possible complicated grief among bereaved family members. J Palliat Med. 2014 Sep;17(9):1025-31. doi: 10.1089/jpm.2013.0552. Epub 2014 Jul 22. PMID: 25050607.
4. The Economist Intelligence Unit . The 2015 quality of death index. Lien Foundation. https://www.eiuperspectives.economist.com/sites/default/files/2015%20EIU%20Quality%20of%20Death%20Index%20Oct%2029%20FINAL.pdf (accessed 16.11.2021).
5. Mayland CR, Williams EM, Ellershaw JE. Assessing quality of care for the dying: the development and initial validation of a postal self-completion questionnaire for bereaved relatives. Palliat Med. 2012 Oct;26(7):897-907. doi: 10.1177/0269216311424953. Epub 2011 Nov 1. PMID: 22045726.
6. Mayland CR, Williams EM, Addington-Hall J, et al. Does the 'Liverpool Care Pathway' facilitate an improvement in quality of care for dying cancer patients?. Br J Cancer. 2013;108(10):1942-1948. doi:10.1038/bjc.2013.203.
7. Mayland CR, Williams EM, Addington-Hall J, et al. Assessing the quality of care for dying patients from the bereaved relatives' perspective: further validation of "Evaluating care and health outcomes--for the dying". J Pain Symptom Manage. 2014 Apr;47(4):687-96. doi: 10.1016/j.jpainsymman.2013.05.013. Epub 2013 Nov 5. PMID: 24210451.

8. Mayland C, Williams E, Ellershaw J. How well do current instruments using bereaved relatives' views evaluate care for dying patients? Palliat Med. 2008 Mar;22(2):133-44. doi: 10.1177/0269216307085742. PMID: 18372378.
9. World Health Organization. Palliative care. 2020. Available at https://www.who.int/news-room/fact-sheets/detail/palliative-care (accessed 21.08.2021).
10. 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 Jun;4(2):167-174. doi: 10.1136/bmjspcare-2013-000596. Epub 2014 Mar 28. PMID: 24681559.
11. Mayland CR, Mulholland H, Gambles M, et al How well do we currently care for our dying patients in acute hospitals: the views of the bereaved relatives? BMJ Supportive & Palliative Care 2017;**7**:316-325.
12. Vogt A, Stiel S, Heckel M, 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 Sep 22;18(1):311. doi: 10.1186/s12955-020-01473-2. PMID: 32962706; PMCID: PMC7507719.
13. Mayland C, McGlinchey T, Gambles M, et al. Quality assurance for care of the dying: engaging with clinical services to facilitate a regional cross-sectional survey of bereaved relatives' views. BMC Health Serv Res. 2018 Oct 10;18(1):761. doi: 10.1186/s12913-018-3558-z. PMID: 30305082; PMCID: PMC6180653.
14. Kupeli N, Candy B, Tamura-Rose G, et al. Tools Measuring Quality of Death, Dying, and Care, Completed after Death: Systematic Review of Psychometric Properties. Patient. 2019 Apr;12(2):183-197. doi: 10.1007/s40271-018-0328-2. PMID: 30141020; PMCID: PMC6397142.
15. Haugen DF, Hufthammer KO, Gerlach C, et al; ERANet-LAC CODE Project Group. Good Quality Care for Cancer Patients Dying in Hospitals, but Information Needs Unmet: Bereaved

Relatives' Survey within Seven Countries. Oncologist. 2021 Jul;26(7):e1273-e1284. doi: 10.1002/onco.13837. Epub 2021 Jun 17. PMID: 34060705; PMCID: PMC8265351.

16. 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 Mar;33(3):357-368. doi: 10.1177/0269216318818299. Epub 2019 Jan 10. PMID: 30628867; PMCID: PMC6376606.
17. Jones, J, Hunter, D. Consensus methods for medical and health services research. BMJ 1995; 311: 376–380.
18. National Health Service. Friends and family test. https://www.england.nhs.uk/fft/ (accessed 25.09.2021).
19. Flora DB, Curran PJ. An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. Psychol Methods. 2004 Dec;9(4):466-91. doi: 10.1037/1082-989X.9.4.466. PMID: 15598100; PMCID: PMC3153362.
20. Muthen LK, Muthen BO. Mplus User's Guide (Sixth ed.) Los Angeles, CA: Muthen & Muthen. 1998-2011.
21. MacCallum RC, Browne MW, Sugawara HM. Power analysis and determination of sample size for covariance structure modeling. Psychological Methods 1996; 1(2): 130–149. https://doi.org/10.1037/1082-989X.1.2.130.
22. Hu L-t, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Structural Equation Modeling. 1999: 6(1): 1–55. https://doi.org/10.1080/10705519909540118.
23. Schreiber JB, Stage FK, King J, et al. Reporting Structural Equation Modeling and Confirmatory Factor Analysis Results: A Review. *The Journal of Educational Research. 2006; 99*(6): 323–337. https://doi.org/10.3200/JOER.99.6.323-338.

24. Watt T, Groenvold M, Deng N, et al. Confirmatory factor analysis of the thyroid-related quality of life questionnaire ThyPRO. Health Qual Life Outcomes. 2014 Sep 10;12:126. doi: 10.1186/s12955-014-0126-z. PMID: 25213857; PMCID: PMC4172819.
25. Fayers PM, Machin D. Quality of life: the assessment, analysis and interpretation of patient-reported outcomes. John Wiley & Sons, 2013.
26. Cohen J. Statistical power analysis for the behavior science. Lawrance Eribaum Association. 1988.
27. Hanson BA. Uniform DIF and DIF Defined by Differences in Item Response Functions. Journal of Educational and Behavioral Statistics. 1998;23 (3):244-253. doi:10.3102/10769986023003244
28. Zumbo BD. A handbook on the theory and methods of differential item functioning (DIF). Ottawa: National Defense Headquarters: 1999.
29. Nagelkerke NJ. A note on a general definition of the coefficient of determination. Biometrika 1991; 78(3): 691-2.
30. Reise, S. P., Moore, T. M., & Haviland, M. G. Bifactor models and rotations: Exploring the extent to which multidimensional data yield univocal scale scores. Journal of Personality Assessment. 2010; 92(6): 544–559.
31. Reise SP, Bonifay WE, Haviland MG. Scoring and modeling psychological measures in the presence of multidimensionality. J Pers Assess. 2013;95(2):129-40. doi: 10.1080/00223891.2012.725437. Epub 2012 Oct 2. PMID: 23030794.
32. Keetharuth AD, Bjorner JB, Barkham M, et al. Exploring the item sets of the Recovering Quality of Life (ReQoL) measures using factor analysis. Qual Life Res. 2019 Apr;28(4):1005-1015. doi: 10.1007/s11136-018-2091-1. Epub 2018 Dec 21. PMID: 30578454; PMCID: PMC6439178.

33. Rubinstein A, Zerbino MC, Cejas C, et al. Making Universal Health Care Effective in Argentina: A Blueprint for Reform. Health Syst Reform 2018;4(3):203-13. doi: 10.1080/23288604.2018.1477537 [published Online First: 2018/08/02]
34. Goldraij G, Tripodoro VA, Aloisio M The ERANet-LAC CODE Project Group, et al. One chance to get it right: improving clinical handovers for better symptom control at the end of life. BMJ Open Quality 2021;10:e001436. doi: 10.1136/bmjoq-2021-001436.
35. Pérez-Cruz PE, Padilla Pérez O, Bonati P, et al. Validation of the Spanish Version of the Quality of Dying and Death Questionnaire (QODD-ESP) in a Home-Based Cancer Palliative Care Program and Development of the QODD-ESP-12. J Pain Symptom Manage. 2017 Jun;53(6):1042-1049.e3. doi: 10.1016/j.jpainsymman.2017.02.005. Epub 2017 Mar 16. PMID: 28323080.