Instruments to measure skills and knowledge of physicians and medical students in palliative care: a systematic review of psychometric properties

Journal:	*Medical Teacher*
Manuscript ID	CMTE-2021-1682.R2
Manuscript Categories:	Articles
Date Submitted by the Author:	n/a
Complete List of Authors:	Lopez-Garcia, Mónica; CUDECA Foundation, ; Malaga Institute of Biomedical Research Rubio-Lamia, Leticia Olga; University of Málaga, Department of Human Anatomy, Legal Medicine and History of Science.; Malaga Institute of Biomedical Research Martín-de-las-Heras, Stella; University of Málaga, Department of Human Anatomy, Legal Medicine and History of Science.; Malaga Institute of Biomedical Research Suarez, Juan; University of Málaga, Department of Human Anatomy, Legal Medicine and History of Science.; Malaga Institute of Biomedical Research Pérez-Carceles, Maria Dolores; Universidad de Murcia Faculty of Medicine, Legal and forensic medicine; Virgen de la Arrixaca Murcia Biosanitary Research Institute Martin-Martin, Jaime; University of Málaga, Legal and forensic medicine
Keywords:	Psychometrics < Assessment, Palliative care < Discipline, Clinical skills < Learning outcomes, General < Phase of education, Medicine < Profession

Instruments to measure skills and knowledge of physicians and medical students in palliative care: a systematic review of psychometric properties

Mónica López-García[a,b], Leticia Rubio[b,c]*, Stella Martin-de-las-Heras[b,c] Juan Suárez[b,c], María D. Pérez-Cárceles[d], Jaime Martin-Martin[a,c]

[a] Cudeca Hospice (Palliative Care Foundation), Málaga, Spain;[b] Instituto de Investigación Biomédica de Málaga-IBIMA, 29010, Málaga, Spain;[c] Department of Legal Medicine. Facultad de Medicina, Universidad de Málaga, 29071, Málaga, Spain; [d] Department of Legal and Forensic Medicine, Faculty of Medicine, Biomedical Research Institute (IMIB-Arrixaca), University of Murcia, Murcia, Spain

*Corresponding:

Professor Leticia Rubio

Department of Human Anatomy, Legal Medicine and History of Science; Legal Medicine Area. Facultad de Medicina, Universidad de Málaga, 29071, Málaga, Spain.

Email: lorubio@uma.es

Instruments to measure skills and knowledge of physicians and medical students in palliative care: a systematic review of psychometric properties

Purpose: Palliative care is constantly increasing around the world. The knowledge and skills of future physicians in this area are crucial. This study evaluates the psychometric properties of knowledge and skills questionnaires used in palliative care, validated in physicians or medical students based on the COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) methodology. Methods: A systematic review was carried out in Cosmin Databases, Cochrane Library, PsycINFO, SciELO, Cinahl, and Medline up to September 2020 (updated June 2021), based on the COSMIN methodology and PRISMA recommendations. The psychometric properties of each included questionnaire were identified. Methodological quality, quality of results, and quality of evidence were evaluated. Results: The search strategy yielded 12 questionnaires assessing knowledge and skills in physicians or medical students. The Palliative Care Knowledge Questionnaire for PEACE (PEACE-Q) and Palliative Care Knowledge Test (PCKT) were the instruments with the highest scores for methodological quality, quality of results, and quality of evidence based on the COSMIN methodology. Conclusions: PEACE-Q and PCKT should be the preferred choice to assess palliative care knowledge and skills in physicians. In-depth studies following COSMIN validation criteria are recommended in order to improve the psychometric properties and cross-cultural validation of the questionnaires.

Keywords: palliative care, knowledge, questionnaires, legal medicine, education.

Practice points:

- Validated questionnaires should be used to measure knowledge and skills in palliative care in undergraduate and postgraduate training programmes.
- The Palliative Care Knowledge Questionnaire for PEACE and the Palliative Care Knowledge Test (PCKT) are the instruments of choice.
- In-depth studies following COnsensus-based Standards for the selection of health Measurement Instruments (COSMIN) criteria are recommended in order to increase the validity, reliability and cross-cultural validation of the questionnaires based on palliative care in a range of contexts.

Background

At least 40 million people are in need of palliative care. Hospice and palliative care often encompass support for family members, which would more than double this need (Centeno et al. 2013). The global challenge is to incorporate a culture of palliative care in health systems, providing a professional, scientific and human response (Ferris et al. 2007).

Palliative care requires a personal and clinical approach to the patient. In clinical practice, the knowledge and skills of health professionals who provide end-of-life care can have a significant impact on patients (Harris et al. 2014). Healthcare professionals working in palliative care teams must have the knowledge necessary to ensure quality of palliative care in this population (Weissman and Blust 2005; Miyashita et al. 2007; Hales et al. 2010; Schulman-Green et al. 2011).

Earlier studies observed that both educational training in medical schools and also healthcare workers' skills in palliative care are insufficient, hence limiting access to palliative care and diminishing the level of service available all around the world (Hawley

2017; Centeno et al. 2017; Rubio et al. 2020). Palliative care is an essential element in training and must be improved in all educational levels (Morrison et al. 2007; Centeno et al. 2007; Elsner et al. 2017; Sociedad Española de Cuidados Paliativos).

Tools that measure palliative care knowledge and skills are necessary to assess competency development. Questionnaires are one tool in the assessment armamentarium to measure palliative care competence. Questionnaires can assess health professional's palliative care skills at each care level, and be duly validated in order to obtain optimal psychometric properties (Mokkink et al. 2010; Mokkink et al. 2018). Questionnaires complement other assessment methods, such as oral or short-answer exams or Objective Structured Clinical Examinations (OSCE). Questionnaires may be used before and after a training programme in order to assess participants' knowledge acquisition and identify deficiencies. This helps create training programmes based on identified needs and reinforces aspects of patient care (Hesselink et al. 2010; DiBiasio 2016). Moreover, self-assessment of competence acquired through questionnaires develops health professionals' skills in palliative care and improves quality of clinical care (Block 2002; Desbiens et al. 2012; Ferrell et al. 2019).

Several validated questionnaires have been used to assess the skills of physicians and medical students and both outcome and population vary among the studies. Al Ansari et al. (2019) analysed physician's knowledge and attitudes toward palliative care in primary care and in general hospitals; Connoly et al. (2018) focused on attitudes, behaviours and knowledge of consultants and physicians-in-training; Brock et al. (2015) designed a questionnaire for specialist paediatricians; and Gryschek et al. (2020) assessed self-efficacy in communication, patient management and multidisciplinary teamwork in palliative care of medical students.

Some authors have used several measurement indicators to assess health professionals' skills in this area (Abu-Saad Huijer and Dimassi 2007; Lange et al. 2011; Yaakup et al. 2014; Bing-Jonsson et al. 2015; Gutiérrez Sánchez et al. 2018). More specifically, Soikkeli-Jalonen et al. analysed existing questionnaires to assess nurses' skills and knowledge in palliative care, identifying nine specific instruments for this purpose, which require more validation studies (Soikkeli-Jalonen et al. 2020). However, to our knowledge, no similar studies have been carried out in physicians or medical students, and there is no consensus on which instruments are most effective to measure acquired skills and knowledge in palliative care in these populations.

The aim of this systematic review was twofold: (1) to identify and describe the validated instruments for measuring knowledge and skills around palliative care among physicians and medical students, and (2) to assess the psychometric properties of these instruments based on COSMIN methodology.

Material and Methods

Design

A systematic review was carried out based on COSMIN methodological standards (Prinsen et al. 2018) and Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) recommendations. The protocol for this systematic review was registered in the Prospective Register of Systematic Reviews (PROSPERO) (registration number CRD42020201363).

Search strategy

The search was carried out in the following databases: Cosmin Databases, Cochrane Library, PsycINFO, SciELO, Cinahl and Medline up to September 2020 (updated June 2021) (Table 1). The search strategy and the inclusion and exclusion criteria were designed based on Terwee et al. (2009). The inclusion criteria were: validated questionnaires in physicians or medical students to assess skills and knowledge in palliative care, analysing at least one psychometric property. The exclusion criteria were: evaluation of intervention programmes, assessment of services, and languages other than English, Spanish or German. All the articles identified were included in the Rayyan bibliographic resource manager for screening and eligibility and to detect duplications. Electronic searches were scrutinised by two independent reviewers for selection, eligibility and inclusion in the review. The two reviewers independently assessed titles and abstracts, and full texts. A third reviewer was involved in discordant decisions, and discussion ensued to reach consensus. Finally, inclusion or exclusion decisions were documented following PRISMA guidelines.

Insert Table 1 around here

Data extraction

The two reviewers independently retrieved information from the selected articles based on the following categories: author, year, population of validation, country of validation, language, instrument description, and psychometric properties analysed.

Quality appraisal

According to Mokkink et al. (2018), the selected studies were evaluated by two authors separately with the latest version of the COSMIN Risk of Bias checklist. The risk of bias checklist was structured in 11 sections: instrument development, content validity, structural validity, internal consistency, cross-cultural validity/measurement invariance, reliability, measurement error, criterion validity, hypothesis testing for construct validity and responsiveness. Each item is scored as very good, adequate, doubtful, or inadequate. The lowest score obtained for any item in the box defines the measurement attribute's overall methodological quality score (Mokkink et al. 2018).

The order of importance of the psychometric properties is as follows: content validity is the most important property, followed by the internal structure of an instrument (including the measurement properties structural validity, internal consistency, and cross-cultural validity/measurement invariance), and, finally, the remaining properties: reliability, measurement error, criterion validity, hypotheses testing for construct validity and responsiveness (COSMIN Taxonomy of Measurement Properties).

Data summary

The psychometric properties of the questionnaires were assessed based on COSMIN criteria: methodological quality, quality of results, and quality of evidence. Methodological quality was evaluated based on the COSMIN Risk of Bias Checklist (Mokkink et al. 2018). Quality of results was assessed based on the updated criteria of Prinsen et al. as follows: positive, indeterminate, inconsistent or negative (Prinsen et al. 2018). Quality of evidence was evaluated and graded in four levels: high, moderate, low and very low, using the Grading of Recommendations Assessment, Development, and

Evaluation (GRADE) (Guyatt et al. 2008). Methodological quality, quality of results and quality of evidence were assessed independently by two authors (M.L-G and J.M-M), with a third author (L.R) to reach a consensus in discordant decisions.

Results

The search strategy yielded 1330 articles; 4 of them were eliminated due to duplicates. After reading the title and abstract, 1228 articles were excluded as they did not meet the inclusion criteria, meaning 98 studies were selected for a full-text reading. Of these, 19 did not show psychometric properties, 41 were not measuring instruments, and 25 were not validated in physicians or medical students. The two reviewers' agreement was very good (Kappa Index= 0.80, I.C 95% 0.63-0.97). One additional questionnaire was included in the updated search carried out in June 2021. Finally, a total of 14 studies based on questionnaires to measure the skills and knowledge of physicians or medical students in palliative care were included in the systematic review for analysis of their psychometric properties (Figure 1).

Insert Figure 1 around here

The 14 studies included in the current review are shown in Table 2. The population groups for validation were: four studies in general physicians (Yamamoto et al. 2013; Mosich et al. 2017; Connolly et al. 2018; Al-Ansari et al. 2019), three in medical students (Fetz et al. 2017; Loera et al. 2018; Gryschek et al. 2020), and two in paediatric specialists (Wool and Northam 2011; Brock et al. 2015); four studies for both

physicians and nurses (Nakazawa et al. 2009; Thi Thanh Vu et al. 2019; Vidal Serrano et al. 2019; Lopez-Garcia et al. 2020).

Six studies were developed from Europe (Mosich et al. 2017; Fetz et al. 2017; Loera et al. 2018; Connolly et al. 2018; Vidal Serrano et al. 2019; Lopez-Garcia et al. 2020), two in Japan (Nakazawa et al. 2009; Yamamoto et al. 2013), three in the United States (Wool and Northam 2011; Brock et al. 2015; Montagnini et al. 2021), one in Brazil (Gryschek et al. 2020), one in Kuwait (Al-Ansari et al. 2019), and one in Vietnam (Thi Thanh Vu et al. 2019). All studies were published in English except two: one in Spanish (Vidal Serrano et al. 2019) and one in German (Mosich et al. 2017).

With regard to the response design of the questionnaires, two questionnaires (Palliative Care Knowledge Test (PCKT) (Nakazawa et al. 2009), Palliative Care Knowledge Questionnaire for PEACE (PEACE-Q) (Yamamoto et al. 2013)) and a section of the Palliative Care Attitude and Knowledge (PCAK) questionnaire (Al-Ansari et al. 2019) used a dichotomous response (true/false). The Self-Efficacy in Palliative Care (SEPC) instrument is evaluated using a visual analogue scale (Gryschek et al. 2020). The rest of the questionnaires used a Likert scale of at least four levels (Wool and Northam 2011; Brock et al. 2015; Mosich et al. 2017; Fetz et al. 2017; Loera et al. 2018; Connolly et al. 2018; Al-Ansari et al. 2019; Thi Thanh Vu et al. 2019; Vidal Serrano et al. 2019; Montagnini et al. 2021).

The instruments used to measure skills and knowledge of physicians or medical students in palliative care were based on different conceptual models and theoretical frameworks (Table 3) to assess different constructs. Among the 14 instruments identified in this review, 12 were original versions. In terms of the conceptual model, the PCKT was developed according to the WHO definition of palliative care and its application in the early stages of illness, while the PCAK was based on a committee of

palliative medicine specialists with clinical, educational and training experience. The instruments assess palliative care knowledge in physicians or medical students on different constructs, with knowledge, attitudes and barriers to palliative care being the most repeated constructs in the questionnaires. (Table 3).

Insert Table 2 around here

Insert table 3 around here

Table 4 shows the psychometric properties of the analysed palliative care questionnaires based on their methodological quality (Mokkink et al. 2018), quality of results (Prinsen et al. 2018), and quality of evidence (Guyatt et al. 2008). According to COSMIN methodology (risk of bias checklist), the main properties were content validity, internal structure (structural validity and internal consistency), and other measurement properties (reliability, measurement error, criterion validity, hypotheses testing for construct validity and responsiveness) (Mokkink et al. 2018). The hypothesis testing psychometric property was not assessed in any of the analysed studies, and for this reason it is not shown in the results for this systematic review.

The Palliative Care Knowledge Test (PCKT) questionnaire was assessed by Nakazawa et al. (Nakazawa et al. 2009). It consists of 20 items to measure knowledge of palliative care in physicians and nurses. With regard to quality of results, internal consistency (KR-20=0.81) and reliability intraclass correlation coefficient (ICC= 0.88) properties were positive, and methodological quality was rated as "very good". Quality of evidence was "high" for the analysed properties (content validity, structural validity, internal consistency, and reliability). However, content validity in methodological

quality scored "appropriate". The structural validity also scored "appropriate" due to sample size. The Palliative Care Knowledge Test Spanish Version (PCKT-SV) reported "positive" internal consistency (KR-20=0.741) in quality of results, and "very good" in cross-cultural adaptation of methodological quality (Lopez-Garcia et al. 2020).

Perinatal Palliative Care Perceptions and Practice Barriers Scale (PPCPBS) were assessed in the study by Wool and Northam (Wool and Northam 2011); its main aim is to measure barriers and perceptions of perinatal palliative care among clinicians using 55 items (plus 14 demographic items). These authors analysed three psychometric properties. With regard to quality of results, content validity, structural validity and internal consistency (α=0.77) scored "positive", and methodological quality "very good". The quality of evidence rating was "high" for content validity and structural validity, and "moderate" for internal consistency.

Yamamoto et al. evaluated the Palliative Care Knowledge Questionnaire for PEACE (PEACE-Q)(Yamamoto et al. 2013), which consists of 33 items to measure physicians' knowledge of palliative care. In methodological quality, the content validity was rated as "very good" because they used the Delphi method (Fitch et al. 2001). Structural validity, internal consistency, and reliability were also rated as "very good". Quality of results was "positive" for internal consistency (KR-20=0.87) and reliability (ICC= 0.84); however, structural validity obtained a score of "inconsistent". All the properties analysed in this questionnaire obtained "high" quality of evidence, except for content validity, which was "moderate" (Yamamoto et al. 2013).

The Paediatric Palliative Care Questionnaire (PPCQ) developed by Brock et al. (Brock et al. 2015) contains 39 items to measure self-efficacy, knowledge and adequacy of prior medical training among paediatric fellowship physicians. The PPCQ reported a

"positive" score in internal consistency (α=0.84-0.95) and "negative" in reliability (K=0.63) for quality of results.

The Palliative Kompetenztest für Ärzte (PKT) (Mosich et al. 2017) was designed by joining the PEACE-Q questionnaire (Yamamoto et al. 2013) translated into German and the items checking specific self-efficacy expectations in palliative care in the Bonn Palliative Knowledge Test (Pfister et al. 2011) transformed into physician's duties. The PKT contains 52 items and assesses palliative care knowledge and self-efficacy in physicians. With regard to methodological quality, content validity was rated "very good", while internal consistency and reliability were rated as "doubtful". Furthermore, cross-cultural validity was rated as "inadequate" in methodological quality; no confirmatory factor analysis was performed. Quality of evidence obtained "moderate" for content validity, internal consistency and reliability. Cross-cultural reported a "low" score in this criterion (Mosich et al. 2017).

The Palliative Care Education and Practice Questionnaire Germany (PCEP-GR) questionnaire (Fetz et al. 2017) contains 36 items to measure knowledge in palliative care in medical students. The PCEP-GR obtained "negative" in structural validity (Kaiser Meyer Olkin 0.81) and in internal consistency (α=0.66-0.83), and "indeterminate" in reliability (Spearman Brown 0.90) for quality of results. Methodological quality was "very good" for these properties and "moderate" for quality of evidence.

The Frommelt Attitude Toward the Care of the Dying Scale (FATCOD-B) was developed by Loera et al. (Loera et al. 2018), containing 30 items to evaluate medical students' attitudes towards care for the dying and normative beliefs about the dying and family members. With regard to methodological quality, content validity was rated as "very good", internal consistency as "doubtful" and reliability as "inadequate". Quality

of results scored "negative" in structural validity and internal consistency, and "indeterminate" in reliability. Quality of evidence scored "moderate" in structural validity, and "low" in internal consistency and reliability.

The Palliative Care Competence Framework Questionnaire (PPCFQ) (Connolly et al. 2018) contains 27 items to assess palliative care knowledge, attitudes and behaviour in physicians. The instrument scored methodological quality as "inadequate" in content validity and reliability. Quality of results scored "positive" in content validity and internal consistency (α =0.944-0.959), and "indeterminate" in reliability.

With regard to methodological quality, the PCAK (Al-Ansari et al. 2019) contains 7 items assessing attitude, self-efficacy and knowledge of palliative care in physicians working in primary care and general hospitals. PCAK scored "very good" in internal consistency and reliability, but "doubtful" in content validity and structural validity. Quality of results scored "positive" in internal consistency (α=0.824), while content validity and reliability (Person test-retest; r=0.95) scored "indeterminate". Quality of evidence scored "low" in content validity and structural validity.

The Palliative Care Difficulties Scale (PCDS) validated by Vidal Serrano et al. (Vidal Serrano et al. 2019) is a translation into Spanish of the original questionnaire (Nakazawa et al. 2009) validated in physicians; it contains 15 items assessing difficulties in palliative care related to communication, support, symptom relief and coordination between nurses and physicians. With regard to methodological quality, internal consistency and cross-cultural scored "very good", while content validity, structural validity, and reliability scored "doubtful". Quality of results, scored "positive" in internal consistency (α=0.87), but "indeterminate" in reliability (Spearman's Rho=0.81) and cross-cultural validity. Quality of evidence scored "low" in content

validity and structural validity, and "moderate" in cross-cultural criteria (Vidal Serrano et al. 2019).

The psychometric properties of the SEPC questionnaire were measured in two studies (Mason and Ellershaw 2004; Gryschek et al. 2020), containing 23 items assessing palliative care self-efficacy competence in medical students. Quality of results of structural validity obtained different scores in both studies: "indeterminate" in Mason and Ellershaw's (2004) study performed with a varimax rotation analysis of SEP; and "negative" in the Gryschek et al.'s (2020) study performed by means of a confirmatory factor analysis. In both studies, internal consistency scored "very good" for methodological quality, and "high" for quality of evidence. Quality of results scored "positive" in internal consistency in both studies, measured with Cronbach's alpha (α =0.92 – 0.93) (Mason and Ellershaw 2004) and (α =0.82 – 0.97) (Gryschek et al. 2020).

The End-of-Life Care Questionnaire (EOL-Q) (Montagnini et al. 2021) contains 28 items assessing self-perceived knowledge, attitudes and behaviours in palliative care for health professionals. The EOL-Q is a version of the original End Of Life Intensive Care Unit (EOL-UCI) (Montagnini et al. 2021). With regard to methodological quality, the questionnaire scored "very good" in internal consistency, "adequate" in content validity and "doubtful" in structural validity. Quality of results scored "positive" in content validity and in internal consistency (α =0.93) (Montagnini et al. 2021).

Insert Table 4 around here

Discussion

The assessment of knowledge and skills of physicians or medical students in terms of palliative care is essential in order to ensure high-quality palliative care at the end of life (Artioli et al. 2019). Several validated questionnaires are available for

assessing knowledge and skills in palliative care and choosing the most appropriate instrument is essential.

The assessment of knowledge and skills in palliative care will allow us to streamline and improve the process in clinical, teaching and research settings, focusing on training (Artioli et al. 2019). Most studies on undergraduate palliative care suggest that many newly graduated physicians feel unprepared to perform end-of-life care (Hesselink et al. 2010; DiBiasio 2016). In fact, Morrison et al. recommended that competencies in palliative care should be included in the educational guidelines for future physicians at international level (Morrison et al. 2007). These competencies include communication, pain and symptom management, quality of life approach, care, coordination, and participation and collaboration in an interdisciplinary team. Self-assessment of acquired competencies has been shown to develop physicians' and nurses' palliative care skills and improve quality of patient care (Desbiens et al. 2012). Using questionnaires on knowledge and skills in palliative care allows limitations in clinical practice to be identified. It would therefore be necessary to select questionnaires with the highest reliability and validity appropriate to the clinical context in order to identify existing educational needs quickly and reliably (DiBiasio 2016). The questionnaires are usually applied at the end of the undergraduate or postgraduate training programmes (Fetz et al. 2017; Ioshimoto et al. 2020) or at the beginning and end of clinical rotations (Mason and Ellershaw 2004). Others fill out the questionnaires in the professional context despite the hectic nature of clinical practice (Al-Ansari et al. 2019; Lopez-Garcia et al. 2020).

To our knowledge, this is the first systematic review based on palliative care questionnaires that assess psychometric properties to measure skills and knowledge in physicians or medical students based on the COSMIN methodology. Validation

following the COSMIN recommendations gives us an optimal score in the four domains and their respective measurement properties: reliability (internal consistency, reliability and measurement error), validity (content validity, construct validity and criterion validity), responsiveness, and interpretability. In this sense, construct validity has three aspects: structural validity, cross-cultural validity and hypothesis testing (Mokkink et al. 2010). Questionnaires with a greater number of psychometric properties analysed and with higher scores allow a higher level of reliability and repeatability in the measures, which in practical terms translates into ensuring we are assessing objectively and fulfilling the aim of the questionnaire (Tsang et al. 2017). This systematic review has shown that PEACE-Q (Yamamoto et al. 2013) and PCKT (Nakazawa et al. 2009) were the instruments with the highest scores for methodological quality, quality of results, and quality of evidence using the statistical test recommended in the COSMIN criteria.

The PCAK and the PCDS questionnaires had a “doubtful” score in the methodological quality in content validity due to two reasons, the first one related to professional perspective according to relevance and comprehension of the items (Al-Ansari et al. 2019), and the second one because it is not clear if the group or interview meetings are recorded and transcribed verbatim (Vidal Serrano et al. 2019). PCAK (Al-Ansari et al. 2019) and PCDS (Vidal Serrano et al. 2019) had "doubtful" methodological quality in structural validity because the factor analysis was 5 times the number of items but in fewer than 100 participants. The EOL-Q (Montagnini et al. 2021) also had a "doubtful" score due to not providing clear information on some of the aspects for performing the principal component factor analysis. The studies on PCEP-GR (Fetz et al. 2017) and FATCOD-B (Loera et al. 2018) obtained "negative" quality of results in the internal consistency property, because the value was offered in ranges and the reference was the lower one. Other questionnaires obtained high values in properties, but based on other

statistical tests different from COSMIN criteria. In this sense, PKT (Mosich et al. 2017), PCEP-GR (Fetz et al. 2017), PCCFQ (Connolly et al. 2018), PCAK (Al-Ansari et al. 2019), PCDS (Vidal Serrano et al. 2019) and FATCOD-B (Loera et al. 2018) obtained high reliability scores based on statistical tests different from ICC or Kappa index, meaning quality of results was scored as "indeterminate". Three of the analysed instruments only assessed two psychometric properties. The PPCQ (Brock et al. 2015) and the PKT (Mosich et al. 2017) analysed internal consistency and reliability, and the SEPC questionnaire (Mason and Ellershaw 2004) structural validity and internal consistency.

This systematic review also reveals that, according to COSMIN criteria, the studies analysed did not assess or report all the recommended properties. Increasing sample size, comparing the results of the participants in the questionnaire to the results of a previously validated questionnaire of similar characteristics, and performing more in-depth statistical analyses such as intraclass correlation coefficient or item response theory will increase the validity of the questionnaires analysed (Mokkink et al. 2010).

Cross-cultural translation is of particular relevance since clinical practice guidelines and questionnaires must be adapted to it (Soikkeli-Jalonen et al. 2020). In this systematic review, three questionnaires were cross-cultural adaptations from the original, with the PCKT-SV (Lopez-Garcia et al. 2020) having the highest score.

Whether an instrument displays appropriate measurement properties is not a fixed attribute, but rather is dependent on the context and population being studied. Studies using questionnaires that have not been validated in the population of interest may be subject to measurement error, and conclusions cannot be drawn with total confidence (Dowrick et al. 2015).

Strength and limitations

This is the first time that a systematic review has been carried out based on the COSMIN methodology in order to identify the properties of questionnaires for evaluating knowledge and skills in palliative care among physicians and medical students (Mokkink et al. 2018; Prinsen et al. 2018). COSMIN is one of the most powerful tools for assessing psychometric properties and selecting assessment instruments (Prinsen et al. 2018). The current study limited the search to articles published in Spanish, English and German, meaning articles in other languages may have been missed. However, most of the studies were published in English. Despite the meticulous search strategy used, some validated questionnaires may have been overlooked in the current study.

Conclusion

In this systematic review, the questionnaires with the greater number of psychometric properties analysed and with the highest psychometrics scores were Palliative Care Knowledge Questionnaire for PEACE (Yamamoto et al. 2013) and Palliative Care Knowledge Test (Nakazawa et al. 2009), according to the COSMIN methodology. These instruments should be the preferred choice to assess palliative care skills in physicians. It is recommended to perform in-depth studies following COSMIN criteria in order to improve the psychometric properties of all the questionnaires analysed. This systematic review also highlights the need for cross-cultural validation of instruments to assess knowledge in palliative care in various contexts.

Declarations of interest statement

Funding

Not applicable.

Conflict of interest / Competing Interest

Not applicable.

Authors Contribution

- Substantial contributions to the conception or design of the work; or the acquisition, analysis, or interpretation of data for the work: all authors
- Drafting the work or revising it critically for important intellectual content: all authors
- Final approval of the version to be published: all authors
- Agreement to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved: all authors
- Gave final approval to the submitted paper: all authors

Acknowledgements

Not applicable.

Ethical Approval

Ethical approval not required. The systematic review was registered in PROSPERO (CRD42020201363).

References

Abu-Saad Huijer H, Dimassi H. 2007. Palliative care in Lebanon: knowledge, attitudes and practices of physicians and nurses. J Med Liban. 55(3):121–128.

Al-Ansari AM, Suroor SN, AboSerea SM, Abd-El-Gawad WM. 2019. Development of palliative care attitude and knowledge (PCAK) questionnaire for physicians in Kuwait. BMC Palliat Care. 18(1):49. https://doi.org/10.1186/s12904-019-0430-9

Artioli G, Bedini G, Bertocchi E, Ghirotto L, Cavuto S, Costantini M, Tanzi S. 2019. Palliative care training addressed to hospital healthcare professionals by palliative care specialists: a mixed-method evaluation. BMC Palliative Care. 18(1):88. https://doi.org/10.1186/s12904-019-0476-8

Bing-Jonsson PC, Bjørk IT, Hofoss D, Kirkevold M, Foss C. 2015. Competence in advanced older people nursing: development of "nursing older people--competence evaluation tool." Int J Older People Nurs. 10(1):59–72. https://doi.org/10.1111/opn.12057

Block SD. 2002. Medical education in end-of-life care: the status of reform. J Palliat Med. 5(2):243–248. https://doi.org/10.1089/109662102753641214

Brock KE, Cohen HJ, Popat RA, Halamek LP. 2015. Reliability and Validity of the Pediatric Palliative Care Questionnaire for Measuring Self-Efficacy, Knowledge, and Adequacy of Prior Medical Education among Pediatric Fellows. J Palliat Med. 18(10):842–848. https://doi.org/10.1089/jpm.2015.0110

Centeno C, Garralda E, Carrasco JM, den Herder-van der Eerden M, Aldridge M, Stevenson D, Meier DE, Hasselaar J. 2017. The Palliative Care Challenge: Analysis of Barriers and Opportunities to Integrate Palliative Care in Europe in the View of National Associations. J Palliat Med. 20(11):1195–1204. https://doi.org/10.1089/jpm.2017.0039

Centeno C, Noguera A, Lynch T, Clark D. 2007. Official certification of doctors working in palliative medicine in Europe: data from an EAPC study in 52 European countries. Palliat Med. 21(8):683–687. https://doi.org/10.1177/0269216307083600

Centeno C, Pons-Izquierdo JJ, Lynch T, Donea O, Rocafort J, Clark D. 2013. EAPC Atlas of Palliative Care in Europe 2013 - Cartographic Edition [Internet]. [place unknown]: EAPC (European Association for Palliative Care); [accessed 2021 Feb 16]. https://dadun.unav.edu/handle/10171/29290

Connolly M, McLean S, Guerin S, Walsh G, Barrett A, Ryan K. 2018. Development and Initial Psychometric Properties of a Questionnaire to Assess Competence in Palliative Care: Palliative Care Competence Framework Questionnaire. Am J Hosp Palliat Care. 35(10):1304–1308. https://doi.org/10.1177/1049909118772565

COSMIN Taxonomy of Measurement Properties. COSMIN [Internet]. [accessed 2022 Feb 19]. https://www.cosmin.nl/tools/cosmin-taxonomy-measurement-properties/

Desbiens J-F, Gagnon J, Fillion L. 2012. Development of a shared theory in palliative care to enhance nursing competence. J Adv Nurs. 68(9):2113–2124. https://doi.org/10.1111/j.1365-2648.2011.05917.x

DiBiasio E. 2016. Palliative and End-of-Life Care Education Among Alpert Medical School Students. R I Med J (2013). 99(4):20–25.

Dowrick AS, Wootten AC, Murphy DG, Costello AJ. 2015. "We Used a Validated Questionnaire": What Does This Mean and Is It an Accurate Statement in Urologic Research? Urology. 85(6):1304–1311. https://doi.org/10.1016/j.urology.2015.01.046

Elsner F, Larkin P, Lorenzl S, Paal P. 2017. Breaking Barriers: Towards a WHO policy for strengthening palliative care education for all healthcare professionals. EAPC Blog [Internet]. [accessed 2021 Feb 16]. https://eapcnet.wordpress.com/2017/11/13/breaking-barriers-towards-a-who-policy-for-strengthening-palliative-care-education-for-all-healthcare-professionals/

Ferrell B, Buller H, Paice J, Anderson W, Donesky D. 2019. End-of-Life Nursing and Education Consortium Communication Curriculum for Interdisciplinary Palliative Care Teams. J Palliat Med. 22(9):1082–1091. https://doi.org/10.1089/jpm.2018.0645

Ferris FD, Gómez-Batiste X, Fürst CJ, Connor S. 2007. Implementing quality palliative care. J Pain Symptom Manage. 33(5):533–541. https://doi.org/10.1016/j.jpainsymman.2007.02.033

Fetz K, Wenzel-Meyburg U, Schulz-Quach C. 2017. Validation of the German revised version of the program in palliative care education and practice questionnaire (PCEP-GR). BMC Palliat Care. 16(1):78. https://doi.org/10.1186/s12904-017-0263-3

Fitch K, Bernstein SJ, Aguilar MD, Burnand B, LaCalle JR, Lazaro P, van het Loo M, McDonnell J, Vader J, Kahan JP. 2001. The RAND/UCLA Appropriateness Method User's Manual [Internet]. [accessed 2021 Jun 17]. https://www.rand.org/pubs/monograph_reports/MR1269.html

Gryschek G, Cecilio-Fernandes D, Mason S, de Carvalho-Filho MA. 2020. Assessing palliative care education in undergraduate medical students: translation and validation of the Self-Efficacy in Palliative Care and Thanatophobia Scales for Brazilian Portuguese. BMJ Open. 10(6):e034567. https://doi.org/10.1136/bmjopen-2019-034567

Gutiérrez Sánchez D, Pérez Cruzado D, Cuesta-Vargas AI. 2018. The quality of dying and death measurement instruments: A systematic psychometric review. J Adv Nurs. https://doi.org/10.1111/jan.13687

Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, Schünemann HJ. 2008. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. BMJ. 336(7650):924–926. https://doi.org/10.1136/bmj.39489.470347.AD

Hales S, Zimmermann C, Rodin G. 2010. Review: the quality of dying and death: a systematic review of measures. Palliat Med. 24(2):127–144. https://doi.org/10.1177/0269216309351783

Harris M, Gaudet J, O'Reardon C. 2014. Nursing care for patients at end of life in the adult intensive care unit. Journal of Nursing Education and Practice. 4(6):84. https://doi.org/10.5430/jnep.v4n6p84

Hawley P. 2017. Barriers to Access to Palliative Care. Palliat Care. 10:1178224216688887. https://doi.org/10.1177/1178224216688887

Hesselink BAM, Pasman HRW, van der Wal G, Soethout MBM, Onwuteaka-Philipsen BD. 2010. Education on end-of-life care in the medical curriculum: students' opinions and knowledge. J Palliat Med. 13(4):381–387. https://doi.org/10.1089/jpm.2009.0291

Ioshimoto T, Shitara DI, do Prado GF, Pizzoni R, Sassi RH, de Gois AFT. 2020. Education is an important factor in end-of-life care: results from a survey of Brazilian physicians' attitudes and knowledge in end-of-life medicine. BMC Med Educ. 20(1):339. https://doi.org/10.1186/s12909-020-02253-8

Lange JW, Mager D, Greiner PA, Saracino K. 2011. The ELDER Project: educational model and three-year outcomes of a community-based geriatric education initiative. Gerontol Geriatr Educ. 32(2):164–181. https://doi.org/10.1080/02701960.2011.572056

Loera B, Molinengo G, Miniotti M, Leombruni P. 2018. Refining the Frommelt Attitude Toward the Care of the Dying Scale (FATCOD-B) for medical students: A confirmatory factor analysis and Rasch validation study. Palliat Support Care. 16(1):50–59. https://doi.org/10.1017/S147895151700030X

Lopez-Garcia M, Rubio L, Gomez-Garcia R, Sanchez-Sanchez F, Miyashita M, Medina-Abellan MD, Perez-Carceles MD. 2020. Palliative care knowledge test for nurses and physicians: validation and cross-cultural adaptation. BMJ Support Palliat Care. https://doi.org/10.1136/bmjspcare-2019-002182

Mason S, Ellershaw J. 2004. Assessing undergraduate palliative care education: validity and reliability of two scales examining perceived efficacy and outcome expectancies in palliative care. Med Educ. 38(10):1103–1110. https://doi.org/10.1111/j.1365-2929.2004.01960.x

Miyashita M, Sanjo M, Morita T, Hirai K, Uchitomi Y. 2007. Good death in cancer care: a nationwide quantitative study. Ann Oncol. 18(6):1090–1097. https://doi.org/10.1093/annonc/mdm068

Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, Bouter LM, de Vet HCW. 2010. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. Qual Life Res. 19(4):539–549. https://doi.org/10.1007/s11136-010-9606-8

Mokkink LB, de Vet HCW, Prinsen CAC, Patrick DL, Alonso J, Bouter LM, Terwee CB. 2018. COSMIN Risk of Bias checklist for systematic reviews of Patient-Reported Outcome Measures. Qual Life Res. 27(5):1171–1179. https://doi.org/10.1007/s11136-017-1765-4

Montagnini M, Smith HM, Price DM, Strodtman L, Ghosh B. 2021. An Instrument to Assess Self-Perceived Competencies in End-of-Life Care for Health Care Professionals:

The End-of-Life Care Questionnaire. Am J Hosp Palliat Care.:10499091211005736. https://doi.org/10.1177/10499091211005735

Morrison LJ, Scott JO, Block SD, American Board of Hospice and Palliative Medicine Competencies Work Group. 2007. Developing Initial Competency-based Outcomes for the Hospice and Palliative Medicine Subspecialist: phase I of the hospice and palliative medicine competencies project. J Palliat Med. 10(2):313–330. https://doi.org/10.1089/jpm.2006.9980

Mosich V, Sellner-Pogány T, Wallner J. 2017. PKT - Palliative competence test for physicians : Design and validation of a questionnaire to assess knowledge and specific self-efficacy expectations of physicians in palliative care. Schmerz. 31(4):375–382. https://doi.org/10.1007/s00482-016-0180-z

Nakazawa Y, Miyashita M, Morita T, Umeda M, Oyagi Y, Ogasawara T. 2009. The palliative care knowledge test: reliability and validity of an instrument to measure palliative care knowledge among health professionals. Palliat Med. 23(8):754–766. https://doi.org/10.1177/0269216309106871

Pfister D, Müller M, Müller S, Kern M, Rolke R, Radbruch L. 2011. Validation of the Bonn test for knowledge in palliative care (BPW). Schmerz. 25(6):643–653. https://doi.org/10.1007/s00482-011-1111-7

Prinsen CAC, Mokkink LB, Bouter LM, Alonso J, Patrick DL, de Vet HCW, Terwee CB. 2018. COSMIN guideline for systematic reviews of patient-reported outcome measures. Qual Life Res. 27(5):1147–1157. https://doi.org/10.1007/s11136-018-1798-3

Rubio L, López-García M, Gaitán-Arroyo MJ, Martin-Martin J, Santos-Amaya I. 2020. Palliative care undergraduate education: Do medical and nursing students need more skills in ethical and legal issues? Medical Hypotheses. 142:110138. https://doi.org/10.1016/j.mehy.2020.110138

Schulman-Green D, Ercolano E, Lacoursiere S, Ma T, Lazenby M, McCorkle R. 2011. Developing and testing a web-based survey to assess educational needs of palliative and end-of-life health care professionals in Connecticut. Am J Hosp Palliat Care. 28(4):219–229. https://doi.org/10.1177/1049909110385219

Sociedad Española de Cuidados Paliativos. [accessed 2021 Feb 16]. https://www.secpal.com/

Soikkeli-Jalonen A, Stolt M, Hupli M, Lemetti T, Kennedy C, Kydd A, Haavisto E. 2020. Instruments for assessing nurses' palliative care knowledge and skills in specialised care setting: An integrative review. J Clin Nurs. 29(5–6):736–757. https://doi.org/10.1111/jocn.15146

Terwee CB, Jansma EP, Riphagen II, de Vet HCW. 2009. Development of a methodological PubMed search filter for finding studies on measurement properties of measurement instruments. Qual Life Res. 18(8):1115–1123. https://doi.org/10.1007/s11136-009-9528-5

Thi Thanh Vu H, Hoang Nguyen L, Xuan Nguyen T, Thi Hoai Nguyen T, Ngoc Nguyen T, Thi Thu Nguyen H, Trung Nguyen A, Pham T, Tat Nguyen C, Xuan Tran B,

et al. 2019. Knowledge and Attitude Toward Geriatric Palliative Care among Health Professionals in Vietnam. Int J Environ Res Public Health [Internet]. [accessed 2020 Oct 14] 16(15). https://doi.org/10.3390/ijerph16152656

Tsang S, Royse CF, Terkawi AS. 2017. Guidelines for developing, translating, and validating a questionnaire in perioperative and pain medicine. Saudi J Anaesth. 11(Suppl 1):S80–S89. https://doi.org/10.4103/sja.SJA_203_17

Vidal Serrano S, López Alonso SR, Fernández Ojeda MDR, Lacalle Remigio JR. 2019. Validation of the "Palliative Care Difficulties Scale (PCDS)" questionnaire in the spanish population about health professionals´ difficulties in palliative care. Rev Esp Salud Publica. 93.

Weissman DE, Blust L. 2005. Education in palliative care. Clin Geriatr Med. 21(1):165–175, ix. https://doi.org/10.1016/j.cger.2004.08.004

Wool C, Northam S. 2011. The Perinatal Palliative Care Perceptions and Barriers Scale Instrument©: development and validation. Adv Neonatal Care. 11(6):397–403. https://doi.org/10.1097/ANC.0b013e318233809a

Yaakup H, Eng TC, Shah SA. 2014. Does Clinical Experience Help Oncology Nursing Staff to Deal with Patient Pain Better than Nurses from other Displines? Knowledge and Attitudes Survey Amongst Nurses in a Tertiary Care in Malaysia. Asian Pacific Journal of Cancer Prevention. 15(12):4885–4891. https://doi.org/10.7314/APJCP.2014.15.12.4885

Yamamoto R, Kizawa Y, Nakazawa Y, Morita T. 2013. The palliative care knowledge questionnaire for PEACE: reliability and validity of an instrument to measure palliative care knowledge among physicians. J Palliat Med. 16(11):1423–1428. https://doi.org/10.1089/jpm.2013.0112

Table 1. Search strategy.

Table 2. Study characteristics and description of the palliative care questionnaires

Table 3. Conceptual model and theoretical frameworks of the palliative care questionnaires.

Table 4. COSMIN psychometric properties of the palliative care questionnaires.

Figure 1. Study selection flow diagram

Table 1. Search strategy

1	("Knowledge"[tiab] OR "attitude"[tiab] OR "skills"[tiab] OR "competences"[tiab]
2	"Palliative care"[MeSH] OR "terminal care[MeSH]" OR "end of life care "[tiab] "terminally ill"[MeSH] OR "advanced illness"[tiab] OR " terminal illness"[tiab] OR "end of life"[tiab]
3	(Instrument[tiab] OR instruments[tiab] OR measure [tiab] OR measures [tiab] OR questionnaire[tiab] OR questionnaires[tiab] OR scale[tiab] OR scales[tiab] OR tool[tiab] OR tools[tiab] OR survey [tiab] OR test [tiab]
4	Instrumentation[sh] OR methods[sh] OR "Validation Studies"[pt] OR "Comparative Study"[pt] OR "psychometrics"[MeSH] OR psychometr*[tiab] OR clinimetr*[tw] OR clinometr*[tw] OR "outcome assessment (health care)"[MeSH] OR "outcome assessment"[tiab] OR "outcome measure*"[tw] OR "observer variation"[MeSH] OR "observer variation"[tiab] OR "Health Status Indicators"[Mesh] OR "reproducibility of results"[MeSH] OR reproducib*[tiab] OR "discriminant analysis"[MeSH] OR reliab*[tiab] OR unreliab*[tiab] OR valid*[tiab] OR coefficient[tiab] OR homogeneity[tiab] OR homogeneous[tiab] OR "internal consistency"[tiab] OR (cronbach*[tiab] AND (alpha[tiab] OR alphas[tiab])) OR (item[tiab] AND (correlation*[tiab] OR selection*[tiab] OR reduction*[tiab])) OR agreement[tiab] OR precision[tiab] OR imprecision[tiab] OR "precise values"[tiab] OR test-retest[tiab] OR (test[tiab] AND retest[tiab]) OR (reliab* [tiab] AND (test[tiab] OR retest[tiab])) OR stability[tiab] OR interrater[tiab] OR inter-rater[tiab] OR intrarater[tiab] OR intra-rater[tiab] OR intertester[tiab] OR inter-tester[tiab] OR intratester[tiab] OR intra-tester[tiab] OR interobserver[tiab] OR inter-observer[tiab] OR intraobserver[tiab] OR intraobserver[tiab] OR intertechnician[tiab] OR inter-technician[tiab] OR intratechnician[tiab] OR intra-technician[tiab] OR interexaminer[tiab] OR inter-examiner[tiab] OR intraexaminer[tiab] OR intra-examiner[tiab] OR interassay[tiab] OR inter-assay[tiab] OR intraassay[tiab] OR intra-assay[tiab] OR interindividual[tiab] OR inter-individual[tiab] OR intraindividual[tiab] OR intra-individual[tiab] OR interparticipant[tiab] OR inter-participant[tiab] OR intraparticipant[tiab] OR intra-participant[tiab] OR kappa[tiab] OR kappa's[tiab] OR kappas[tiab] OR repeatab*[tiab] OR ((replicab*[tiab] OR repeated[tiab]) AND (measure[tiab] OR measures[tiab] OR findings[tiab] OR result[tiab] OR results[tiab] OR test[tiab] OR tests[tiab])) OR generaliza*[tiab] OR generalisa*[tiab] OR concordance[tiab] OR (intraclass[tiab] AND correlation*[tiab]) OR discriminative[tiab] OR "known group"[tiab] OR "factor analysis"[tiab] OR "factor analyses"[tiab] OR dimension*[tiab] OR subscale*[tiab] OR (multitrait[tiab] AND scaling[tiab] AND (analysis[tiab] OR analyses[tiab])) OR "item discriminant"[tiab] OR "interscale correlation*"[tiab] OR error[tiab] OR errors[tiab] OR "individual variability"[tiab] OR (variability[tiab] AND (analysis[tiab] OR values[tiab])) OR (uncertainty[tiab] AND (measurement[tiab] OR measuring[tiab])) OR "standard error of measurement"[tiab] OR sensitiv*[tiab] OR responsive*[tiab] OR ((minimal[tiab] OR minimally[tiab] OR clinical[tiab] OR clinically[tiab]) AND (important[tiab] OR significant[tiab] OR detectable[tiab]) AND (change[tiab] OR difference[tiab])) OR (small*[tiab] AND (real[tiab] OR detectable[tiab]) AND (change[tiab] OR difference[tiab])) OR "meaningful change" [tiab] OR "ceiling effect"[tiab] OR "floor effect"[tiab] OR "Item response model"[tiab] OR IRT[tiab] OR Rasch[tiab] OR "Differential item functioning"[tiab] OR DIF[tiab] OR "computer adaptive testing"[tiab] OR "item bank"[tiab] OR "cross-cultural equivalence"[tiab]
5	#1 AND #2 AND #3 AND #4
6	"Protocol"[Publication Type] OR "addresses"[Publication Type] OR "biography"[Publication Type] OR "case reports"[Publication Type] OR "comment"[Publication Type] OR "editorial"[Publication Type] OR "congresses" [Publication Type] OR "consensus development conference"[Publication Type] OR "consensus development conference"[Publication Type] OR "practice guideline"[Publication Type]) OR "suffering from"[tiab] OR "animals"[MeSH]
7	#5 NOT #6
8	FILTER: Language (English, Spanish, and German)
9	FILTER: Period (1980–2020)

Table 2. Study characteristics and description of the palliative care questionnaires.

Study (author and year)	Population	Original language/translation	Instrument Description	Measurement properties	Validity/reliability	COSMIN Score
Palliative Care Knowledge Test (PCKT)						
PCKT (Nakazawa et al. 2009)	Japan: General physicians and nurses (n=797)	Japan	20 items in 5 domains including 'philosophy', 'pain', 'dyspnoea', 'psychiatric problems' and 'gastrointestinal problems' were selected. Dichotomy response: true/false	1. Structural Validity 2. Internal Consistency 3. Reliability	Structural Validity Internal consistency: Kuder-Richardson formula 20: 0.81 Reliability: Kappa coefficient 0.26-0.74 Intraclass correlation: test-retest examination 0.88	1. Structural Validity: Adequate 2. Internal Consistency: Very good 3. Reliability: Very good
PCKT-Spanish Version (PCKT-SV) (Lopez-Garcia et al. 2020)	Spanish: Physicians and nurses (n=561)	English/Spanish		1. Content Validity 2. Internal Consistency 3. Cross-Cultural	Internal Consistency: Kuder-Richardson formula 20: 0.741	1. Content Validity: Very good 2. Internal Consistency: Doubtful 3. Cross Cultural: Very good
Perinatal Palliative Care Perceptions and Practice Barriers Scale (PPCPBS)						
PPCPBS (Wool and Northam 2011)	USA: Physicians perinatal medicine (n=264)	English	Perceptions scale: 23 items + Barriers scale: 22 items +14 demographics items. 6-point Likert: strongly disagree – strongly agree.	1. Content Validity 2. Structural Validity 3. Internal Consistency	Structural validity: Kaises-Mayer Olkin=0.88 Internal consistency Cronbach´s α= 0.77.	1. Content Validity: Very good 2. Structural validity: Very good 3. Internal Consistency: Very good
Palliative Care Knowledge Questionnaire for PEACE (PEACE-Q)						
PEACE-Q (Yamamoto et al. 2013)	Japan: Physicians not palliative medicine (n=434)	Japanese	33 items across the following 9 domains: (1) philosophy of palliative care, (2) cancer pain, (3) side effects of opioids, (4) dyspnea, (5) nausea and vomiting, (6) psychological distress, (7) delirium, (8) communication regarding palliative care, and (9) community-based palliative care. Dichotomic response: true/false	1. Content Validity 2. Structural Validity 3. Internal consistency 4. Reliability	Structural Validity: Item Response Theory ranged from -2.76 to 0.29, discrimination of 0.69 to 2.67. Internal Consistency: Kuder-Richardson Formula 20=0.87; Item Response Theory: difficulty in all items =-2.76 to 0.29, with a discrimination = 0.69 to 2.67. Reliability: Intraclass correlation=0.84; kappa coefficient =-0.04 to 0.85, and was 0.3 or less for 14 items.	1. Content Validity: Adequate 2. Structural Validity: Very good 3. Internal consistency: Very good 4. Reliability: Very good
Pediatric Palliative Care Questionnaire (PPCQ)						
PPCQ (Brock et al. 2015)	U.S.A: Pediatric fellows (pediatric cardiology, critical care, hematology/ oncology, and neonatal-perinatal medicine	English	Self-reporting: Knowledge.39 items (3 subscales): self-efficacy (23), knowledge (10), fellow's perceived adequacy of prior medical education (6) (5-point Likert-scale)	1. Content Validity 2. Internal consistency 3. Reliability	Internal consistency: Cronbach's alpha: Self-efficacy=0.95; Medical education Self-reportings=0.84. Reliability: "*Subscale Self effycacy"* intraclass correlation coefficient=0.78 and Self-efficacy kappa=0.61"Subscale Adequacy of Medical Education Summary Score" intraclass correlation	1. Content Validity: Adequate 2. Internal consistency: Doubtful 3. Reliability: Doubtful

	programs) (n=147)				coefficient=0.85. "Subscale Adequacy of Medical Education Summary Score kappa=0.63	
Palliative competence test for physicians (PKT)						
PKT (Mosich et al. 2017)	Austria: Physicians (n=105)	German	2 subscales: 34 Knowledge (True/false) Expectation of self-efficacy(spSWE) (likert 0-3: 3 points for each "true" answer; for "fairly correct" 2; for "barely true" 1 point and for "not true" 0 points)	1. Content validity 2. Internal consistency 3. Cross cultural 4. Reliability	Internal Consistency Cronbach´s alpha 0.767 for Knowledge Test and for 0. 919 spSWE. Reliability: test-retest Pearson autocorrelation 0.99	1- Content validity Very good 2- Internal consistency Doubtful 3- Cross cultural Inadequate 4- Reliability Doubtful
Palliative Care Education and Practice Questionnaire Germany (PCEP-GR)						
PCEP-GR(Fetz et al. 2017)	German: Physicians students (n=680)	Germany	36 items subscales: preparation to provide palliative care, attitudes towards palliative care, self-estimation of competence in communication with dying patients and their relatives and self-estimation of knowledge and skills in palliative care. 5-point Likert-scale	1. Content Validity 2. Reliability 3. Internal consistency 4. 3-Structural Validity	Reliability: Spearman-Brown coefficient=0 .90; Internal Consistency Cronbach's alpha " Preparation to provide palliative care"=0,83/"Attitudes towards palliative ca2= 0,66/ "Self-estimation of competence in communication with dying patients and their relatives"= 0,82/"Self-estimation of knowledge and skills in palliative care"=0,75; Structural Validity: Kaiser-Meyer-Olkin criterion=0 .81.	1- Content Validity Adequate 2- Reliability Inadequate 3- Internal consistency Very good 4- Structural Validity Very good
Frommelt Attitude Toward the Care of the Dying Scale (FATCOD-B)						
FATCOD-B(Loera et al. 2018)	Italy: Physicians students (n=200)	Italian	Two dimensions "attitude about care of the dying person"/ "normative beliefs about the dying person and family members" =30 items in two dimensional construct. 5-point Likert scale	1. Structural Validity 2. Internal consistency 3. Reliability	Structural Validity: Rash Model Internal Consistency: Cronbach's alpha coefficient:0.80/0.53 Reliability Index: 0.98/0.99	1. Structural Validity: Very Good 2. Internal Consistency: Doubtful 3. Reliability: Inadequate
Palliative Care Competence Framework Questionnaire (PCCFQ)						
PCCFQ (Connolly et al. 2018)	Ireland: Physician (n=328) All specialties in adult medical care and direct patient contact .	English	3 subscales and a possible range of 1 to 5 for each subscale (5 indicates a more positive response). 10 Questions assessed knowledge, 6 questions assessed attitudes, and 11 assessed behaviour.	1. Content validity 2. Internal Consistency 3. Reliability	Internal Consistency: knowledge subscale (Cronbach α>.944), the attitude subscale (Cronbach α > .862), and the behaviour subscale (Cronbach α > .959). Reliability: Person r 0.634 – 0.624 between subscales	1. Content validity Inadequate 2. Internal Consistency: Very Good 3. Reliability: Inadequate
Palliative care attitude and knowledge (PCAK)						
PCAK (Al-Ansari et al. 2019)	Kuwait: Physicians (n=232) primary care clinics and general hospitals	Arabic	37 items test in 3 subscales: 1- Attitude (11 items): strongly dis-agree (1), disagree (2), not sure (3), agree (4), strongly agree (5); 2- Self-knowledge: 5 points likert (5) for	1. Content Validity 2. Structural Validity 3. Internal Consistency 4. Reliability 5. Criterion Validity	Structural Validity: Kaiser-Meyer-Olkin 0.699 Internal consistency Cronbach's alpha 0.681-0.806 Test retest reliability=0,95	1. Content Validity: Doubtful 2. Structural Validity: Doubtful 3. Internal consistency:

			excellent response,(4) for very good, (3) for good, (2) for weak and (1) for none; 3-Basic knowledge for true or false		Criterion Validity r=0.394-0.893	Very good 4. Reliability: Very good 5. Criterion Validity: Adequate
Palliative Care Difficulties Scale (PCDS)						
PCDS (Vidal Serrano et al. 2019)	Spain: Nurses and physicians (n=118)	English/Spanish	15 items: 5 domains (communication in multidisciplinary teams, communication with patients and family, expert support, alleviating symptoms and community coordination) PCDS total score of the responses in each domain, and thus difficulty subscores ranged from 3 to 15. Likert scale (1 = completely disagree, 2 = disagree, 3 = neither agree nor disagree, 4 = agree, 5 = fully agree).	1. Content Validity 2. Structural validity 3. Internal consistency 4. Cross-cultural 5. Reliability	Structural Validty: Kaiser-Meyer-Olkin= 0.76 and Barlett Sphericity <0.01, principal component analysis with Varimax rotation. Internal Consistency: Cronbach= 0.87; Reliability: test-retest Spearman Rho=0.81;	1. Content Validity: Doubtful 2. Structural validity: Doubtful 3. Internal consistency: Very good 4. Cross-cultural: Very good 5. Reliability Doubtful
Self-Efficacy in Palliative Care; (SEPC)						
SEPC (Mason and Ellershaw 2004)	United Kingdom: Physicians students (n=217)	English	23 items; three factors: (a) communication (b) patient management and (c) multidisciplinary tea working. Based on visual analogue scale ranging from “very anxious” to “very confident” (0-10cm). Higher values indicating higher confidence in that specific task.	1.Content Validity 2.Structural Validity Internal Consistency	Structural Validity: varimax rotated identified 3 factors, accounted 68.2% of the variance on pretest and 74.7% on post-test. Internal Consistency: Cronbach´s= 0.92-0.93	1. Content Validity Inadequate 2. Structural Validity Adequate 3. Internal Consistency Very Good
SEPC- Brazilian Version (SEPC-BR) (Gryschek et al. 2020)	Brazil: Physicians students (n=111)	English/Portuguese		1.Structural Validity 2.Internal Consistency 3.Cross-Cultural Validity	Structural Validity: confirmatory factor analysis replicating original factors (0.51-0.90) Internal Consistency: Cronbach´s: 0.82-0.97 and composite reliability: 0.82-0.96)	1. Structural Validity: Very good 2. Internal Consistency: Very good 3. Cross Cultural Validity: Doubtful
End of Life Care Questionnaire (EOL-Q)						
EOL-Q(Montagnini et al. 2021)	United States_ healthcare professionals from multiple units (n=1,197)	English	28 items; three domains: self-perceived knowledge, self-perceived attitudes and self-perceived behaviors. 5-point Likert scales ranging from 1=strongly disagree to 50strongly agree.	1. Content Validity 2. Structural Validity 3. Internal Consistency	Structural Validity: principal axis factoring method (9.47-1.13) Internal Consistency: Cronbach´s: 0.93	1. Content Validity: Adequate. 2. Structural Validity: Doubtful 3. Internal Consistency: Very good

PCKT: Palliative Care Knowledge Test; PPCPBS: Perinatal Palliative Care Perceptions and Practice Barriers Scale; PEACE-Q: Palliative Care Knowledge Questionnaire for PEACE; PPCQ: Pediatric Palliative Care Questionnaire; PKT: Palliative competence test for physicians; PCEP-GR: Palliative Care Education and Practice Questionnaire Germany; FATCOD-B: Frommelt Attitude Toward the Care of the Dying Scale; PCCFQ: Palliative Care Competence Framework Questionnaire; PCAK: Palliative care attitude and knowledge; PCDS: Palliative Care Difficulties Scale; SEPC: Self-Efficacy in Palliative Care; PC: Palliative Care; EOL-Q: End of Life Care Questionnaire

Table 3. Conceptual model and theoretical frameworks of the palliative care questionnaires.

Instrument	Construct	Conceptual model and theoretical framework
PCKT (Nakazawa et al. 2009)	Knowledge of palliative care.	Assessment to measure the efficacy of palliative care educational programs. Instrument not limited to terminal care and is more in accord with the definition by WHO that palliative care is applicable early in the course of illness.
PPCPBS (Wool and Northam 2011)	Barriers to perinatal palliative care. Perceptions of perinatal palliative care.	Neonatal palliative care trends, prenatal diagnostics, and emerging issues in cases of life limiting fetal conditions organizing using a principle-based ethical conceptual framework.
PEACE-Q (Yamamoto et al. 2013)	Knowledge of palliative care.	Measurement tool to quantify knowledge level of physicians about broader areas of palliative care, by which the effect of education programs could be measured.
PPCQ (Brock et al. 2015)	Self-Assessment, medical education assessment and barriers to Palliative Care.	Based on a literature search on pediatric palliative care, palliative care simulation, communication tool, validity and reliability. Adapted of previous survey on palliative care and reviewed by a panel of expert.
PKT (Mosich et al. 2017)	Knowledge of palliative care and self-efficacy.	Map all the thematic areas of palliative education that are associated with the recommendations of the specialized palliative societies for the teaching of palliative care for physicians
PCEP-GR (Fetz et al. 2017)	Knowledge of palliative care	Based on the evaluation of Undergraduate Palliative Care Education program for medical students. Theory-based based four subscales: preparation to provide palliative care, attitudes towards palliative care, self-estimation of competence in communication with dying patients and their relatives and self-estimation of knowledge and skills in palliative care
FATCOD-B (Loera et al. 2018)	Attitude of toward the Care of the Dying: "attitude about care of the dying person"/ "normative beliefs about the dying person and family members".	Assess attitudes toward caring for terminally ill persons and their families. In order for the tool to be applicable to attitudes of all students from a variety of programs of study,
PCCFQ (Connolly et al. 2018)	Knowledge, attitudes, and behavior to palliative care.	The framework´s face and content validity were assured by a conducting literature review that was used to map competences. Framework was developed and assessed by 105 practitioners.
PCAK (Al-Ansari et al. 2019)	Attitude, self-efficacy and knowledge of palliative care.	Committee members of palliative medicine specialist experienced in clinical, education and training to other physicians identified aspects of knowledge and attitude were central to palliative care practice. Defined palliative care knowledge as "understanding of death, and dying, symptom management, medications and any intervention needed for those patients care". Self-efficacy was defined as "people beliefs about their abilities to make designated levels of performance that implement influence over events that affect their lives while attitude as a system of beliefs and knowledge that everyone has got or has learned through his lifetime (Bandura 1994)
PCDS (Vidal Serrano et al. 2019)	Difficulties in palliative care relating to communication, support, alleviating symptoms and coordination.	Identify factors associated with self-reported practices and difficulties with palliative care in order to improve curricula of palliative care educational interventions.
SEPC (Mason and Ellershaw 2004)	Self-Efficacy competence of palliative care.	Developed from a questionnaire originally designed to assess the effect of a preceptorship programs in palliative care for Australian undergraduate medical students. The anchors of the scale have been chosen according to the theoretical underpinnings of social cognitive theory.
EOL-Q (Montagnini et al. 2021)	Self-perceived knowledge, attitudes and behaviors in palliative care	Developed and modified to palliative care from the original Scale of End of Life Care in the Intensive Unit .

PCKT: Palliative Care Knowledge Test; PPCPBS: Perinatal Palliative Care Perceptions and Practice Barriers Scale; PEACE-Q: Palliative Care Knowledge Questionnaire for PEACE; PPCQ: Pediatric Palliative Care Questionnaire; PKT: Palliative competence test for physicians; PCEP-GR: Palliative Care Education and Practice Questionnaire Germany; FATCOD-B: Frommelt Attitude Toward the Care of the Dying Scale; PCCFQ: Palliative Care Competence Framework Questionnaire; PCAK: Palliative care attitude and knowledge; PCDS: Palliative Care Difficulties Scale; SEPC: Self-Efficacy in Palliative Care; PC: Palliative Care; EOL-Q: End of Life Care Questionnaire

Table 4. COSMIN psychometric properties of the palliative care questionnaires.

Instrument	Content validity			Structural Validity			Internal Consistency			Cross Cultural validity/ measurement invariance			Reliability			Criterion Validity		
	M	Q	QE	M	Q (methods)	QE	M	Q α Cronbach, KR-20	QE	M	Q	QE	M	Q r Pearson, Anova, Kappa,	QE	M	Q r Prearon, Spearman´s rho, p value	QE
PCKT (Nakazawa et al. 2009)	A	?	H	A	+ IRT	H	V	+ KR-20= 0.81	H				V	+ ICC=0.88	H			
PCKT-SV (Lopez-Garcia et al., 2020)	V	+	H				D	+ KR-20=0.741	M	V	+	H						
PPCPBS (Wool and Northam 2011)	V	+	H	V	+ PCA	H	V	+ α =0.77	M									
PEACE-Q (Yamamoto et al. 2013)	V	±	M	V	± IRT	H	V	+ KR-20= 0.87	H				V	+ ICC= 0.84	H			
PPCQ (Brock et al. 2015)							D	+ α =0.84-0.95	M				D	- K=0.63	L			
PKT (Mosich et al. 2017)	V	+	M				D	+ α =0.76	M	I	-	L	D	? Pearson r=0.99	M			
PCEP-GR (Fetz et al. 2017)				V	- Kaiser 0.81	M	V	- α =0.66-0.83	M				I	? Spearman-Brown=0.90	M			
FATCOD-B (Loera et al. 2018)				V	- Rash	M	D	- α =0.53-0.80	L				I	? Reliability Index: 0.98-0.99	L			
PCCFQ (Connolly et al. 2018)	I	+	M				V	+ α =0.944-0.959	H				I	? Person r = 0.624-0.634-	VL			
PCAK (Al-Ansari et al. 2019)	D	?	L	D	- Kaiser 0.699	L	V	+ α =0.824	M				V	? Pearson r (test retest) =0.95	H			
PCDS (Vidal Serrano et al. 2019)	D	±	L	D	- PCA	L	V	+ α =0.87	H	V	?	M	D	? Sperman Rho=0.81	H			
SEPC (Mason and Ellershaw 2004)				A	? Varimax Rotation	M	V	+ α =0.92 – 0.93	H									
SEPC -BR(Gryschek et al. 2020)				V	- CFA	H	V	+	H	D	?	M				V	- r= -0.499	L

								α =0.82 – 0.97										
EOL-Q (Montagnini et al. 2021)	A	+	M	D	? Principal Axis Factoring	L	V	+ α = 0.90 - 0.93	H									

PCKT: Palliative Care Knowledge Test; PPCPBS: Perinatal Palliative Care Perceptions and Practice Barriers Scale; PEACE-Q: Palliative Care Knowledge Questionnaire for PEACE; PPCQ: Pediatric Palliative Care Questionnaire; PKT: Palliative competence test for physicians; PCEP-GR: Palliative Care Education and Practice Questionnaire Germany; FATCOD-B: Frommelt Attitude Toward the Care of the Dying Scale; PCCFQ: Palliative Care Competence Framework Questionnaire; PCAK: Palliative care attitude and knowledge; PCDS: Palliative Care Difficulties Scale; SEPC: Self-Efficacy in Palliative Care; EOL-Q: End of Life Care Questionnaire

M: Methodological quality of the study rated as V = Very good, A = Adequate, D = Doubtful, I = Inadequate, empty boxes = not reported.

Q: Quality of the results rated as + = positive rating, ? = indeterminate rating, ± = inconsistent rating, - = negative rating, empty boxes = not reported.

QE: Quality of evidence rated as H = High, M = Moderate, L = Low, VL = Very low.

CFA: Confirmatory Factor Analysis.

IRT: Item Response Theory

PAA Principal Axis Factoring.

PCA: Principal Component Analysis

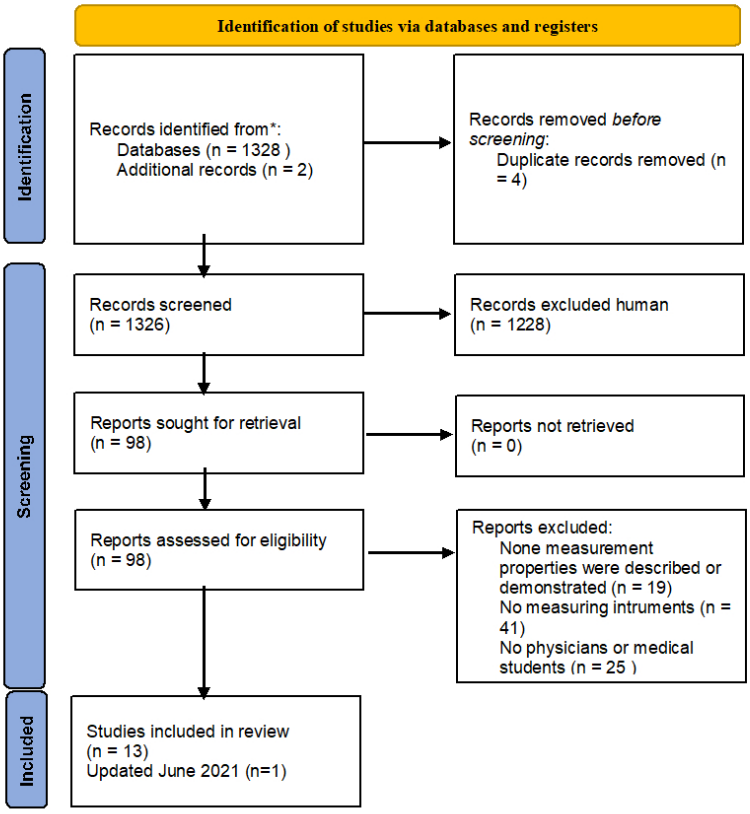

Figure 1: Study selection Flowdiagram

Study selection Flowdiagram

677x646mm (38 x 38 DPI)