Article:
Abahussin, AA orcid.org/0000-0002-7831-1445, West, RM orcid.org/0000-0001-7305-3654, Wong, DC orcid.org/0000-0001-8117-9193 et al. (1 more author) (2019) PROMs for pain in adult cancer patients: a systematic review of measurement properties. Pain Practice, 19 (1). pp. 93-117. ISSN 1530-7085

https://doi.org/10.1111/papr.12711

MRS. ASMA ABAHUSSIN (Orcid ID : 0000-0002-7831-1445)

Article type : Review

Title: PROMs for pain in adult cancer patients: a systematic review of measurement properties.

Authors: Asma A Abahussin[1, 2], MSc; Robert M West[1], PhD; David C Wong[1], DPhil; Lucy E Ziegler[1], PhD.

[1] Leeds institute of Health Sciences, School of Medicine, University of Leeds, Leeds, United Kingdom

[2] Biomedical Technology Department, College of Applied Medical Sciences, King Saud University, Riyadh, Saudi Arabia

Corresponding author:

Asma Abahussin

Leeds Institute of Health Sciences

Level 10 Worsley Building

Clarendon Way

University of Leeds

Leeds LS2 9NL

Tel: +441133434572

Email: umaab@leeds.ac.uk

Key words: PROMs, Pain Measurement, Cancer, Adult, Psychometrics, Measurement properties, Systematic review.

Abstract

Context: Pain is one of the most devastating symptoms for cancer patients. One-third of patients who experience pain do not receive effective treatment. A key barrier to effective pain management is lack of routine measurement and monitoring of pain.

Patient-Reported Outcome Measures (PROMs) are recommended for measuring cancer pain. However, evidence to guide the selection of the most appropriate measure to identify and monitor cancer pain is limited. A systematic review of measurement properties of PROMs for pain in cancer patients is needed to identify the best validated measure for adoption to an electronic platform.

Objectives: Systematically review measurement properties of PROMs used for adult cancer patients to measure pain and, as a secondary goal, investigate the evidence of validated mobile health (mHealth) applications used to measure pain (registration number: CRD42017065575).

Methods: Medline, EMBASE and CINAHL were systematically searched in March 2018 for studies examining measurement properties for PROMs for pain in adult cancer patients. Both of the methodological quality of the studies and their results were appraised using the COSMIN checklist and specific measurement properties criteria respectively.

Results: Sixteen studies evaluating eight instruments were included. No studies using a PROM in a mHealth application were identified. The methodological quality of the measurement properties ranged between poor and fair. No instrument showed strong positive evidence for all the evaluated measurement properties. Based on the available evidence, the Brief Pain Inventory-Short Form (BPI-SF) had the strongest evidence to support its selection for the measurement of cancer pain.

Conclusion: The BPI-SF was the best performing measure across all proprieties evaluated through COSMIN. Better quality validation studies of PROMs for cancer pain are needed to explore the full range of measurement properties. Utilising mHealth applications for measuring pain for cancer patients is an innovative approach worth of further investigation.

Key words: PROMs, Pain Measurement, Cancer, Adult, Psychometrics, Measurement properties, Systematic review.

Introduction

Pain is one of the most devastating symptoms for patients throughout the cancer trajectory [1]. The World Health Organisation's (WHO) analgesic three-step ladder is the definitive clinical principle for cancer pain management [2]. It has been used since 1986, and involves a stepwise approach to analgesic prescriptions for cancer pain [3, 4]. Despite improvements in pain management with this strategy [5, 6], some cancer patients still experience high levels of pain in situations where it is possible to reduce suffering. Around 25% to 33% of cancer patients receive insufficient pain management [7, 8]. Two systematic reviews that assessed the quality of pain management in adult cancer patients revealed modest improvements in pain management between 2008 and 2014, but stated that one third of patients who experience pain are under-treated [9, 10].

Inadequate pain assessment is considered a significant barrier to sufficient pain management [11, 12]. Therefore, pain management guidelines emphasise routine and systematic pain assessment including documenting a detailed pain history and medication efficacy [12, 13]. Measuring and documenting pain every 4 hours for each patient for 5 weeks improved pain assessment (from 42% to 71%) and pain management (from 59% to 97%) for patients in intensive care units (ICUs) including oncologic surgery ICU [14]. Assessing and recording pain regularly may be appropriate given that a patient is more likely to forget such details about pain after receiving care. In particular, retrospective scores of pain are known to be significantly higher than contemporaneous diary report of pain [15].

Patient-Reported Outcome Measures (PROMs) are recommended for measuring cancer pain [16]. PROMs are frequently presented as questionnaires that are completed by patients to measure health-related constructs [17]. Various PROMs are used for assessing pain in patients with cancer, including the Brief Pain Inventory (BPI) [18-21], the McGill Pain Questionnaire (MPQ) [22-24], the Numerical Rating Scale (NRS) [1, 25-27], and the Visual Analogue Scale (VAS) [26-28]. With the increasing number of these instruments, it is more challenging to identify which is the most appropriate one for use in a clinical or research setting. Systematic reviews of measurement properties are useful for critically appraising and comparing the content and measurement properties of all available tools measuring a specific construct to understand their strengths and limitations and make an informed choice [29, 30].

COSMIN (the COnsensus-based Standards for the selection of health Measurement INstruments) is a committee that aims to improve the selection of health-related measurement instruments by forming tools and guidelines, based on international experts' consensus, for conducting or assessing a systematic review of measurement properties [29-31]. The need for such standards was identified in 2009 through reviewing systematic review studies of health status measurement instruments. The methods used to assess the quality of the studies and the quality of the results differed widely, and the methodological quality of such reviews should be improved [30]. The COSMIN tools include a protocol for systematic reviews of measurement instruments, a

checklist to assess the methodological quality of studies and quality criteria for measurement properties [31-35]. Indeed, the quality of both validation studies and their results (measurement properties) is important to appraise instruments for a health construct. If the methodological quality of a study is inappropriate, the results cannot be trusted and the quality of the instrument remains unclear [32, 36]. Since measurement properties are essential for evaluating an instrument [30, 37], and they are not clearly defined in the literature, the COSMIN group provides international definitions [31, 38].

A systematic review of the measurement properties of the established tools for measuring pain in cancer patients is crucial to compare them and identify the best validated pain measure for this population. Very few reviews of this type have been conducted previously. Jensen (2003) conducted a review of the validity and reliability of pain measures for adult cancer patients, but it has several limitations [39]. The review was not systematic and is now out-dated. Publications were from a three-year period only and their methodological quality was not clearly considered. There are systematic reviews of instruments related to the holistic symptoms of cancer [40] and Quality-of-Life (QoL) [41]. For both types of study there has been no rigorous assessment or any consideration of the quality of the studies and their results. These instruments are not specific for pain in cancer patients. Indeed, QoL instruments usually measure multiple constructs and general health perceptions; and are not designed to be a specific pain measurement instrument, although they do contain symptom items [30, 40]. In addition, the WHO analgesic ladder is designed specifically for cancer patients [3, 4]. This implies that treatment of cancer pain requires an approach specific to cancer patients and that cancer patients are distinct from non cancer patients in the type of pain they experience. This relates largely to the meaning of the pain; cancer pain is often interpreted as an indicator of disease progression and the association of pain with life threatening disease is a likely to account for why pain interference is ranked higher in cancer patients compared to non cancer patients even when pain intensity is the same [42]. As a consequence, it is logical to have a specific pain tool for cancer patients.

Mobile health (mHealth) is an innovative and timely method for health monitoring and intervention in the home setting, which can be utilised in measuring pain, compared to existing approaches. This is because of the increasing use of mobile and smart device applications and social media [43, 44]. mHealth has been described as the use of *"mobile computing, medical sensor, and communications technologies for health care"* [45]. The rapid increase in mobile device use has been associated with similar expansion in the field of mHealth. For example, mHealth interventions including smartphone applications (apps) have been utilised in supporting self-management and medication adherence for asthma [46], Parkinson's disease [47], lower back pain [48] and chronic conditions [49]. For pain assessment and management for cancer patients, only a few mHealth interventions have been attempted [1, 50-52].

The aim of the current study is to systematically review the measurement properties of PROMs for pain used for adult cancer patients following the COSMIN framework and recommendations. A secondary goal is to investigate the evidence of validated mHealth applications or mobile electronic tools used to measure pain for adult cancer patients.

Methods

The systematic review[1] was conducted according to the guidelines from the Cochrane Handbook for Systematic Reviews [36], in combination with the protocol for systematic reviews of measurement properties recommended by COSMIN panel [31, 53].

Search strategy

In accordance with the Cochrane guidelines, the PICO (Population, Intervention, Comparison, and Outcome) concepts were applied to the research question. Adult (≥ 18 years of age) cancer patients were considered as the population (P); the intervention (I) was PROMs used to measure pain; the comparison (C) concept was not applicable to the research question since this would require an unmeasured arm; while the measurement properties defined by COSMIN were considered the outcomes (O) for this systematic review. There are nine measurement properties grouped within three domains: reliability (internal consistency, reliability, and measurement error), validity (content validity, construct validity (or hypotheses testing), structural validity, cross-cultural validity, and criterion validity), and responsiveness [38] (see Appendix 1 for properties' definitions). The criterion validity was excluded for this review since no gold standards exist either for PROM instruments [54] or for measurements of pain [55, 56].

Medline (Ovid from 1996), EMBASE (Ovid from 1996), and CINAHL (EBSCO from 1981) electronic databases were searched in March 2018. A search strategy was designed and performed using search terms that had been carefully specified after several consultations with a librarian and an information specialist. In accordance with the Cochrane review guidelines, a combination of index terms, such as MeSH in Medline, and free-text terms for each identified PICO concept was searched combined by the conjunction 'OR'. Then, the search results for all the concepts were combined by the conjunction 'AND'. To focus the search to retrieve PROM tools, the Oxford filter for PROMs developed by the PROMs Group [29, 57] was used. The search was restricted to English language publications, with no time limitation. An example of the detailed search strategy applied on Medline is illustrated in Appendix 2. Additional papers were identified by manually searching the reference lists of the included primary studies and key review studies as well as searching forward referencing of these studies using the Web of Science database.

[1] The systematic review is registered with registration number CRD42017065575 on PROSPERO, an international database of prospectively registered systematic reviews in health and other fields produced by Centre of Reviews and Dissemination (CRD) at University of York.

Eligibility criteria and selection of articles

Studies were selected based on the following criteria: a validation published primary study for a PROM used to specifically measure pain that reported one or more of its measurement properties, and the instrument was administered on adult (≥ 18 years of age) patients with a definite cancer of any type. Studies that included patients with diseases other than cancer were excluded unless the results for the cancer patients were presented separately. PROMs that were specific to certain cancer type were excluded as the review aimed to select the best pain measurement tool for the wider cancer population. PROMs that were general, that is measuring pain within other constructs, or were indirectly related to pain, such as QoL, disease symptoms, and treatment satisfaction instruments, were excluded. Furthermore, the review excluded studies that validated measures based on measures of our interest, RCTs (randomised control trials) or other longitudinal studies, as recommended by the COSMIN protocol. Such studies usually provide indirect evidence and it is difficult to assess validity or responsiveness. No hypotheses regarding the validity or responsiveness of the instrument of interest are formulated and verified in these studies [31, 53].

The results from the searched databases were accumulated in reference manager software (EndNote X7) where any duplicate articles were removed. The studies' titles and abstracts were scanned against the specified inclusion and exclusion criteria before the articles were read as full texts and re-examined for eligibility. The selection of studies was conducted independently through the two stages by two reviewers (AA and LZ), and any disagreements between them were resolved through discussion.

Quality assessment of the studies

The methodological quality of the included studies was evaluated and rated using the COSMIN checklist, which has a 4-point rating scale [32, 33]. The studies were rated by AA and LZ. The checklist consists of nine boxes representing the nine measurement properties defined by the COSMIN panel. Each box has 5–18 items describing whether a study on a measurement property meets the standard for appropriate methodological quality. A score (poor, fair, good, excellent) was given to each item based on the level of adherence to a specific standard. The overall score for each measurement property was specified by considering the lowest score awarded to any item in the checklist box associated with the property. For example, if one item in the internal consistency box was graded as poor for a study, the overall methodological quality of this property was rated as poor. Each measurement property was rated separately.

Quality assessment of the instruments

The quality of the measurement properties of the instrument was assessed using the modified version of the quality criteria for good measurement properties published by the COSMIN panel [29, 34]. The possible rating specified by these criteria for a measurement property is 'positive', 'indeterminate', or 'negative' (see Appendix 3). The measurement properties were assessed by AA and LZ.

Best evidence synthesis

The strength of the evidence for the measurement properties for each tool was determined by considering the following: the number of studies, their methodological quality, and the consistency and quality of the results [31, 53]. The evidence of a measurement property is considered: (1) strong (positive or negative), when consistent findings were derived from multiple studies (at least two) of good methodological quality, or by one study of excellent methodological quality; (2) moderate (positive or negative), when consistent findings were derived from multiple studies of fair methodological quality, or from one study of good methodological quality; (3) limited (positive or negative), when findings were derived from one study of fair methodological quality; (4) conflicting, when conflicting results were found in two or more studies; and (5) unknown, when findings were derived only from studies of poor methodological quality. AA and LZ attributed the level of evidence for the tools.

Data extraction

Two groups of data were extracted from each study and reported in tables. The first group, study characteristics, encompasses general information about the study and the evaluated instrument. This includes the author(s), year of publication, characteristics of the population among which the instrument was evaluated (disease, gender, mean age, settings, country and language), and general features of the instrument as described by the study (name, construct, number of items and version). The second data group, instruments' measurement properties, represents the results of the measurement properties of the tool reported by the study. All the necessary data were extracted by AA and LZ.

Results

Study selection

The PRISMA diagram [58] in Figure 1 shows the results from the literature search and the selection process. Sixteen validation studies of pain measurement instruments met the eligibility criteria. The review did not identify any studies that used a pain measurement instrument in a smartphone or tablet application oriented to adult cancer patients.

Study characteristics

Eight pain measurement instruments were evaluated by the included studies. The characteristics of the studies and instruments are illustrated in Table 1. Table 2 shows the results for measurement properties as reported by the studies.

Study methodological quality and result quality

The ratings of the methodological quality of the studies and the quality of the results per measurement property and instrument are reported in Table 3. The strength of the evidence for each property per instrument is shown in Table 4. Summaries are provided below.

Brief Pain Inventory (BPI)

The BPI has 15 items, and was evaluated in different languages by eight studies [59-66] making it the most evaluated instrument of all the instruments included in this review. The BPI was designed to measure the severity of pain and its impact on functioning of patients using an 11-point NRS. It also uses a drawing where patients mark the location of their pain, and asks about pain treatment and relief [60].The majority of the studies identified two factors (severity and interference) for the BPI using confirmatory factor analysis (CFA) to support structural validity. The quality of this result was rated negative because the ratio of the variance explained by the first to the second factors was less than 4. This was consistent in multiple fair methodological studies [61-63] leading to moderate negative evidence for structural validity. Indeed, the BPI structure validity was reported by an excellent methodological study [59], but the findings were rated as indeterminate, as the percentage of variance explained by each factor was not provided. Therefore, this result was ignored for evidence synthesis. The evidence for the construct validity was moderate positive. Cross-cultural validity for the BPI was reported only by poor methodological studies [59, 60, 62-66], which had no multiple group CFA or differential item functioning (DIF). The quality of the findings was indeterminate in all the studies for the same reason. This resulted in unknown evidence for the BPI cross-cultural validity. In terms of the BPI reliability, only one study (poor methodological) [59] reported test-retest reliability. This means there is unknown evidence for the BPI reliability property.

On the other hand, internal consistency was addressed by all eight studies and showed strong positive evidence.

BPI-Short Form (BPI-SF)

The BPI-SF has 11 items, and was assessed by three studies [23, 67, 68]. It has the same two subscales as the BPI. This was confirmed by moderate positive evidence for structural validity. The latter was rated as positive (unlike the full version) because the first factor accounted for more than 20% of the variability, and the ratio of the variance explained by the two factors was greater than 4. The evidence for the cross-cultural validity was unknown, as was the case with the original version. The assessment of the construct validity and the internal consistency properties showed moderate positive evidence.

McGill pain questionnaire (MPQ)

One study [69] assessed the MPQ and met the eligibility criteria of the review. The MPQ as described by the study has 4 subscales (sensory, affective, evaluative, and miscellaneous) with 24 items. The methodological quality for almost all the evaluated measurement properties was rated poor mainly because the sample size (N =114) was inadequate (i.e. was not greater than 7 times the number of items). This gave unknown evidence for these properties (see Table 4).

MPQ-Short Form (MPQ-SF)

The MPQ-SF was validated by one study, and has two subscales (sensory and affective) derived from the original MPQ with 15 items [23]. It was unclear how missing items were handled through the analyses of this study, so the methodological quality for the evaluated measurement properties rated as fair. There was limited positive evidence for internal consistency and structural validity and limited negative evidence for construct validity.

MPQ-SF-2

MPQ-SF-2 is an update of the original version (MPQ-SF) that includes neuropathic qualities in addition to the sensory and affective qualities [22]. It has four subscales (continuous, intermittent, neuropathic, and affective) with 24 items. It was assessed by one study, which showed excellent methodological quality for both internal consistency and structural validity. The findings for the latter were rated as negative, as a number of positive rating criteria were not met, including the Tucker-Lewis index (TLI), which was less than 0.95. The methodological quality of the construct validity was rated as fair in this study because no hypotheses were formulated before testing, but it is possible to deduce what was expected. The evidence synthesis of the MPQ-SF-2 resulted in strong positive evidence for internal consistency, strong negative evidence for structural validity, and limited positive evidence for construct validity.

Location-Based Assessment of Sensory Symptoms in Cancer (L-BASIC)

The L-BASIC includes four items and was evaluated by one study [70]. The study tested three measurement properties with fair methodological quality for internal consistency and reliability and poor methodological quality for construct validity. It was not clear how missing data were handled, and no information was given about the measurement properties of the comparator instruments. The evidence synthesis for the L-BASIC resulted in limited positive evidence for internal consistency and reliability, and unknown evidence for construct validity.

Cancer Pain Inventory (CPI)

The CPI comprises 5 subscales (catastrophizing, interference, stoicism, social aspects, and pain medication) with 19 items. Three measurement properties for the instrument were tested by only one study, which was rated with fair methodological quality. No explanation was given of how missing data were handled [71]. The synthesis of evidence showed limited positive evidence for both internal consistency and construct validity and limited negative evidence for structural validity, as the ratio of the variance explained by the first to the second factor was less than 4.

Brief 4-week pain diary

The brief 4-week pain diary has 7 items in 2 subscales (pain and pain impact on quality of life). It was evaluated by one study [72], which tested internal consistency with poor methodological quality and tested construct validity with fair methodological quality. The study was rated thus because no factor analysis was performed, and there was no explanation of how missing items were handled. The synthesis of evidence resulted in unknown evidence for internal consistency and limited positive evidence for construct validity.

Discussion

PROMs play an increasingly significant role in monitoring symptoms in cancer patients and can facilitate improvements in quality of life through timely identification and management of symptoms. They promote communication between patients and health professionals and enhance patients' involvement in care and treatment planning and decision [73]. Research evidence indicated that the use of PROMs as part of routine clinical care for cancer patients increases patient satisfaction with care [73], improves symptom management and overall quality of life [74, 75] leading to less frequent hospitalisation and better survival rates [74].

The current study mainly aimed to review the validated PROMs used to measure pain as one of the significant symptoms for cancer. This provides healthcare professionals with an evidence based selected instrument. The review found 3398 studies from which 3373 studies were excluded on the title and abstract screening stage to end with 25 studies. Nine further studies were excluded at the full text screening stage due to the reasons detailed in Figure 1. Sixteen studies evaluated eight pain measurement instruments were included in the review. These studies were conducted in various countries so the languages of the instruments were also heterogeneous (see Table 1). The studies and their results for the measurement properties were systematically reviewed and appraised using the COSMIN checklist and good measurement properties criteria proposed by the COSMIN group respectively (see Table 3). The strength of evidence was identified, based on the COSMIN best evidence synthesis guidelines, for each of the evaluated measurement properties per instrument, as shown in Table 4.

Internal consistency was assessed in all the included instruments. Construct validity and structural validity were the second most frequently evaluated measurement properties. Cross-cultural validity was evaluated by ten studies, seven of which were about BPI. Reliability was addressed in only three instruments. The remaining measurement properties, that is, measurement error, content validity, and responsiveness, were not tested in any instrument (Table 2). 52/60 of the methodological quality of the evaluated measurement properties ranged between poor and fair quality. The low ratings were generally due to insufficient sample sizes, vague or not previously formulated hypotheses, lack of information regarding the handling of missing items or regarding the constructs being measured by the comparator instruments or their measurement properties, internal consistency statistics not being calculated for each subscale individually, and multiple-group CFA not being performed for translated instruments.

The evidence synthesis presented in Table 4 showed that no instrument had strong positive evidence for all the evaluated measurement properties. Therefore, no strong recommendation can be derived from the available evidence in relation to identifying a fully validated pain measurement instrument for adult cancer patients. Based on the available evidence, the BPI-SF is the best evaluated instrument, as it shows moderate positive evidence in internal consistency, construct validity, and structural validity whereas none of the other instruments showed comparable evidence. Indeed, the full BPI and MPQ-SF-2 showed stronger positive evidence for internal consistency compared to the BPI-SF. On the other hand, the BPI showed negative structural validity as reported by several fair methodological studies; this resulted in moderate negative evidence while the BPI-SF had moderate positive evidence. The MPQ-SF-2 also showed negative structural validity in addition to inadequate evidence for the other measurement properties indicating that the BPI-SF has greater validity.

The results of the review should be interpreted with caution. It should not be presumed that the instruments for which it was not possible to establish adequate validity are invalid. Typically, there was insufficient evidence to establish their validity; information was missing and the quality of

the available research was inadequate. Therefore, more validation studies of better quality are needed to address all the measurement properties of the identified instruments and to reveal more about their quality.

The review did not identify any studies that used pain measurement instruments in a smartphone or tablet application for adult cancer patients, which establishes the valuable opportunity for future researches in this area. In fact, a study aiming to develop a system that utilises an app based on a validated pain measurement scale to record pain has been initiated. The results of the current study inform the choice of the scale.

Strengths and limitations

The systematic review was informed by Cochrane guidelines and COSMIN protocol. This approach has added to the robustness of the study. Cochrane is particularly tailored to systematic reviews of RCTs (randomised control trial) studies and this study is oriented to measurement properties studies. As the former is well accredited and has more structured search strategy guidelines, it was used in line with the protocol to ensure advantages of both to achieve the aim of the review. COSMIN recommends building the search strategy in combination with a search filter for finding studies on measurement properties [29, 76]. However, when this was piloted, it retrieved far less relevant studies compared to using Cochrane guidelines for constructing the search strategy and filtering the search using the Oxford filter. This may be because the COSMIN filter for measurement properties studies is designed and validated for PubMed [76] and is not validated for the databases used in this review [29].

Using the Oxford filter is a probable explanation for not identifying any mHealth applications using PROMs for pain. Whilst there may be mHealth applications for pain, the filter successfully excluded them because they did not use PROMs.

The study has some limitations. The assessment of the studies and measurement properties was limited in some instances by lack of information available in some papers. In these instances, no further information was sought from the original authors. In addition, the review was restricted to English language publications only; we acknowledge that validation studies may have been published in other languages which may provide further insight into these tools.

Conclusion

Sixteen studies were identified but little evidence of thorough evaluation of pain tools. Given the extent of current published evidence, the BPI-SF is the most appropriate instrument. More validation studies of better quality are desired.

Utilising mHealth applications for measuring pain for cancer patients is an innovative approach worth of further investigation. A study in this area is established.

Disclosures and Acknowledgements

This study is funded by King Saud University, Saudi Arabia. The funder is gratefully acknowledged. The authors declare no conflicts of interest.

References

1. Besse KT, Faber-te Boveldt ND, Janssen GH, et al. Pain Assessment with Short Message Service and Interactive Voice Response in Outpatients with Cancer and Pain: A Feasibility Study. *Pain Pract*. 2016;**16:**320-6.
2. Raphael J, Ahmedzai SH, Barrie J, et al. *The British Pain Society's Cancer pain management: a perspective from the British Pain Society, supported by the Association for Palliative Medicine and the Royal College of General Practitioners*. 2010. Available from: https://www.britishpainsociety.org/static/uploads/resources/files/book_cancer_pain.pdf. Accessed
3. Gao W, Gulliford M, Bennett MI, Murtagh FE, Higginson IJ. Managing cancer pain at the end of life with multiple strong opioids: a population-based retrospective cohort study in primary care. *PLoS One*. 2014;**9:**e79266.
4. WHO. *World Health Organisation: Cancer Pain Relief*. 1986. Available from: http://apps.who.int/iris/bitstream/10665/43944/1/9241561009_eng.pdf. Accessed
5. Azevedo Sao Leao Ferreira K, Kimura M, Jacobsen Teixeira M. The WHO analgesic ladder for cancer pain control, twenty years of use. How much pain relief does one get from using it? *Support Care Cancer*. 2006;**14:**1086-93.
6. Ventafridda V, Tamburini M, Caraceni A, De Conno F, Naldi F. A validation study of the WHO method for cancer pain relief. *Cancer*. 1987;**59:**850-6.
7. Vuong S, Pulenzas N, DeAngelis C, et al. Inadequate pain management in cancer patients attending an outpatient palliative radiotherapy clinic. *Support Care Cancer*. 2016;**24:**887-92.

8. Mitera G, Zeiadin N, Kirou-Mauro A, et al. Retrospective assessment of cancer pain management in an outpatient palliative radiotherapy clinic using the Pain Management Index. *J Pain Symptom Manage*. 2010;**39:**259-67.
9. Greco MT, Roberto A, Corli O, et al. Quality of cancer pain management: an update of a systematic review of undertreatment of patients with cancer. *J Clin Oncol*. 2014;**32:**4149-54.
10. Deandrea S, Montanari M, Moja L, Apolone G. Prevalence of undertreatment in cancer pain. A review of published literature. *Ann Oncol*. 2008;**19:**1985-91.
11. Kwon JH. Overcoming barriers in cancer pain management. *J Clin Oncol*. 2014;**32:**1727-33.
12. Shute C. The challenges of cancer pain assessment and management. *Ulster Med J*. 2013;**82:**40-2.
13. National Comprehensive Cancer Network. *National Comprehensive Cancer Network: Adult cancer pain.* 2017. Available from: https://www.nccn.org/professionals/physician_gls/PDF/pain.pdf. Accessed 23 April, 2017.
14. Erdek MA, Pronovost PJ. Improving assessment and treatment of pain in the critically ill. *Int J Qual Health Care*. 2004;**16:**59-64.
15. Lewandowski AS, Palermo TM, Kirchner HL, Drotar D. Comparing diary and retrospective reports of pain and activity restriction in children and adolescents with chronic pain conditions. *Clin J Pain*. 2009;**25:**299-306.
16. Stewart J. The challenges of cancer pain assessment. *Ulster Med J*. 2014;**83:**44-6.
17. Gilbert A, Sebag-Montefiore D, Davidson S, Velikova G. Use of patient-reported outcomes to measure symptoms and health related quality of life in the clinic. *Gynecol Oncol*. 2015;**136:**429-39.
18. Chow E, Ding K, Parulekar WR, et al. Revisiting classification of pain from bone metastases as mild, moderate, or severe based on correlation with function and quality of life. *Support Care Cancer*. 2016;**24:**1617-23.
19. Dalton JA, Higgins MK, Miller AH, Keefe FJ, Khuri FR. Pain Intensity and Pain Interference in Patients With Lung Cancer: A Pilot Study of Biopsychosocial Predictors. *Am J Clin Oncol*. 2015;**38:**457-64.
20. Montague L, Green CR. Cancer and breakthrough pain's impact on a diverse population. *Pain Med*. 2009;**10:**549-61.
21. Cleeland CS, Ryan KM. Pain assessment: global use of the Brief Pain Inventory. *Ann Acad Med Singapore*. 1994;**23:**129-38.
22. Gauthier LR, Young A, Dworkin RH, et al. Validation of the short-form McGill pain questionnaire-2 in younger and older people with cancer pain. *J Pain*. 2014;**15:**756-70.
23. Shin H, Kim K, Young Hee K, Chee W, Im EO. A comparison of two pain measures for Asian American cancer patients. *West J Nurs Res*. 2007;**29:**545-60.
24. Melzack R. The McGill Pain Questionnaire: major properties and scoring methods. *Pain*. 1975;**1:**277-99.
25. Bortolussi R, Zotti P, Matovic M, et al. A phase II study on the efficacy and safety of procedural analgesia with fentanyl buccal tablet in cancer patients for the placement of indwelling central venous access systems. *Support Care Cancer*. 2016;**24:**1537-43.
26. Haefeli M, Elfering A. Pain assessment. *Eur Spine J*. 2006;**15:**S17-24.

27. Hjermstad MJ, Fayers PM, Haugen DF, et al. Studies comparing Numerical Rating Scales, Verbal Rating Scales, and Visual Analogue Scales for assessment of pain intensity in adults: a systematic literature review. *J Pain Symptom Manage*. 2011;**41:**1073-93.
28. Cong Y, Sun K, He X, et al. A Traditional Chinese Medicine Xiao-Ai-Tong Suppresses Pain through Modulation of Cytokines and Prevents Adverse Reactions of Morphine Treatment in Bone Cancer Pain Patients. *Mediators Inflamm*. 2015;**2015:**961635.
29. COSMIN. *COSMIN*. 2017. Available from: http://www.cosmin.nl/index.html. Accessed 9th Jan 2017.
30. Mokkink LB, Terwee CB, Stratford PW, et al. Evaluation of the methodological quality of systematic reviews of health status measurement instruments. *Qual Life Res*. 2009;**18:**313-33.
31. Mokkink LB, Prinsen CA, Bouter LM, Vet HC, Terwee CB. The COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) and how to select an outcome measurement instrument. *Braz J Phys Ther*. 2016;**20:**105-13.
32. Terwee CB, Mokkink LB, Knol DL, et al. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. *Qual Life Res*. 2012;**21:**651-7.
33. Mokkink LB, Terwee CB, Patrick DL, et al. 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*. 2010;**19:**539-49.
34. Terwee CB, Bot SD, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol*. 2007;**60:**34-42.
35. Mokkink LB, Terwee CB, Knol DL, et al. Protocol of the COSMIN study: COnsensus-based Standards for the selection of health Measurement INstruments. *BMC Med Res Methodol*. 2006;**6:**2.
36. Higgins JP, Green S. *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0* 2011. Available from: www.handbook.cochrane.org. Accessed 20 December 2016.
37. Englbrecht M, Tarner IH, van der Heijde DM, et al. Measuring pain and efficacy of pain treatment in inflammatory arthritis: a systematic literature review. *J Rheumatol Suppl*. 2012;**90:**3-10.
38. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol*. 2010;**63:**737-45.
39. Jensen MP. The validity and reliability of pain measures in adults with cancer. *J Pain*. 2003;**4:**2-21.
40. Kirkova J, Davis MP, Walsh D, et al. Cancer symptom assessment instruments: a systematic review. *J Clin Oncol*. 2006;**24:**1459-73.
41. Bonomi AE, Shikiar R, Legro MW. Quality-of-life assessment in acute, chronic, and cancer pain: a pharmacist's guide. *J Am Pharm Assoc (Wash)*. 2000;**40:**402-16.
42. Fordyce W. Pain in cancer and chronic non-cancer conditions: similarities and differences. *Acta Anaesthesiol Scand*. 2001;**45:**1086-9.

43. WeAreSocial. *SOCIAL THINKING: DIGITAL IN 2017: GLOBAL OVERVIEW*. 2017. Available from: https://wearesocial.com/blog/2017/01/digital-in-2017-global-overview. Accessed 24 April, 2017.
44. Deloitte. *Global Mobile Consumer Survey 2016: UK Cut*. 2016. Available from: https://www.deloitte.co.uk/mobileuk/. Accessed 24 April, 2017.
45. Istepanian RSH, Jovanov E, Zhang YT. Guest Editorial Introduction to the Special Section on M-Health: Beyond Seamless Mobility and Global Wireless Health-Care Connectivity. *IEEE Trans Inf Technol Biomed*. 2004;**8:**405-14.
46. Kosse RC, Bouvy ML, de Vries TW, et al. mHealth intervention to support asthma self-management in adolescents: the ADAPT study. *Patient Prefer Adherence*. 2017;**11:**571-77.
47. Lakshminarayana R, Wang D, Burn D, et al. Smartphone- and internet-assisted self-management and adherence tools to manage Parkinson's disease (SMART-PD): study protocol for a randomised controlled trial (v7; 15 August 2014). *Trials*. 2014;**15:**374.
48. Irvine AB, Russell H, Manocchia M, et al. Mobile-Web app to self-manage low back pain: randomized controlled trial. *J Med Internet Res*. 2015;**17:**e1.
49. Whitehead L, Seaton P. The Effectiveness of Self-Management Mobile Phone and Tablet Apps in Long-term Condition Management: A Systematic Review. *J Med Internet Res*. 2016;**18:**e97.
50. Fortier MA, Chung WW, Martinez A, Gago-Masague S, Sender L. Pain buddy: A novel use of m-health in the management of children's cancer pain. *Comput Biol Med*. 2016;**76:**202-14.
51. Stinson JN, Jibb LA, Nguyen C, et al. Construct validity and reliability of a real-time multidimensional smartphone app to assess pain in children and adolescents with cancer. *Pain*. 2015;**156:**2607-15.
52. Jibb LA, Cafazzo JA, Nathan PC, et al. Development of a mHealth Real-Time Pain Self-Management App for Adolescents With Cancer: An Iterative Usability Testing Study [Formula: see text]. *J Pediatr Oncol Nurs*. 2017;**34:**283-94.
53. Terwee CB, de Vet HC, Prinsen CA, Mokkink LB. *Protocol for systematic reviews of measurement properties*. 2011. Available from: http://www.cosmin.nl/images/upload/files/Protocol%20klinimetrische%20review%20version%20nov%202011(1).pdf. Accessed 15 January 2017.
54. Mokkink LB, Terwee CB, Knol DL, et al. The COSMIN checklist for evaluating the methodological quality of studies on measurement properties: a clarification of its content. *BMC Med Res Methodol*. 2010;**10:**22.
55. Julie Hearn, Higginson IJ. Cancer pain epidemiology: a systematic review In: Eduardo D. Bruera, Portenoy RK, eds. *Cancer Pain: Assessment and Management* USA: Cambridge University Press 2003: 19-37.
56. McDowell I, Newell C. *Measuring Health: A Guide to Rating Scales and Questionnaires*. Second edn. New York, USA: Oxford University Press, 1996.
57. University of Oxford. *Patient-Reported Outcome Measures Group*. 2016. Available from: http://phi.uhce.ox.ac.uk/. Accessed 5th Jan, 2017.
58. Moher D, Liberati A, Tetzlaff J, Altman DG, Group P. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Int J Surg*. 2010;**8:**336-41.
59. Aisyaturridha A, Naing L, Nizar AJ. Validation of the Malay Brief Pain Inventory questionnaire to measure cancer pain. *J Pain Symptom Manage*. 2006;**31:**13-21.

60. Ballout S, Noureddine S, Huijer HA, Kanazi G. Psychometric evaluation of the arabic brief pain inventory in a sample of Lebanese cancer patients. *J Pain Symptom Manage*. 2011;**42:**147-54.
61. Caraceni A, Mendoza TR, Mencaglia E, et al. A validation study of an Italian version of the Brief Pain Inventory (Breve Questionario per la Valutazione del Dolore). *Pain*. 1996;**65:**87-92.
62. Ferreira KA, Teixeira MJ, Mendonza TR, Cleeland CS. Validation of brief pain inventory to Brazilian patients with pain. *Support Care Cancer*. 2011;**19:**505-11.
63. Mystakidou K, Mendoza T, Tsilika E, et al. Greek brief pain inventory: validation and utility in cancer pain. *Oncology*. 2001;**60:**35-42.
64. Saxena A, Mendoza T, Cleeland CS. The assessment of cancer pain in north India: the validation of the Hindi Brief Pain Inventory--BPI-H. *J Pain Symptom Manage*. 1999;**17:**27-41.
65. Uki J, Mendoza T, Cleeland CS, Nakamura Y, Takeda F. A brief cancer pain assessment tool in Japanese: the utility of the Japanese Brief Pain Inventory--BPI-J. *J Pain Symptom Manage*. 1998;**16:**364-73.
66. Wang XS, Mendoza TR, Gao SZ, Cleeland CS. The Chinese version of the Brief Pain Inventory (BPI-C): its development and use in a study of cancer pain. *Pain*. 1996;**67:**407-16.
67. Kalyadina SA, Ionova TI, Ivanova MO, et al. Russian Brief Pain Inventory: validation and application in cancer pain. *J Pain Symptom Manage*. 2008;**35:**95-102.
68. Klepstad P, Loge JH, Borchgrevink PC, et al. The Norwegian brief pain inventory questionnaire: translation and validation in cancer pain patients. *J Pain Symptom Manage*. 2002;**24:**517-25.
69. Mystakidou K, Parpa E, Tsilika E, et al. Greek McGill Pain Questionnaire: validation and utility in cancer patients. *J Pain Symptom Manage*. 2002;**24:**379-87.
70. Burkey AR, Kanetsky PA. Development of a novel location-based assessment of sensory symptoms in cancer patients: preliminary reliability and validity assessment. *J Pain Symptom Manage*. 2009;**37:**848-62.
71. Deshields TL, Tait RC, Manwaring J, et al. The Cancer Pain Inventory: preliminary development and validation. *Psychooncology*. 2010;**19:**684-92.
72. Maunsell E, Allard P, Dorval M, Labbe J. A brief pain diary for ambulatory patients with advanced cancer: acceptability and validity. *Cancer*. 2000;**88:**2387-97.
73. Velikova G, Booth L, Smith AB, et al. Measuring quality of life in routine oncology practice improves communication and patient well-being: a randomized controlled trial. *J Clin Oncol*. 2004;**22:**714-24.
74. Basch E, Deal AM, Kris MG, et al. Symptom Monitoring With Patient-Reported Outcomes During Routine Cancer Treatment: A Randomized Controlled Trial. *J Clin Oncol*. 2016;**34:**557-65.
75. Basch E, Jia X, Heller G, et al. Adverse symptom event reporting by patients vs clinicians: relationships with clinical outcomes. *J Natl Cancer Inst*. 2009;**101:**1624-32.
76. Terwee CB, Jansma EP, Riphagen, II, de Vet HCW. Development of a methodological PubMed search filter for finding studies on measurement properties of measurement instruments. *Qual Life Res*. 2009;**18:**1115-23.

Table 1: Characteristics of the identified studies

Article	Characteristics of the population						Characteristics of the tool			
	Disease	**Sample size (% female)**	**Mean age (SD)**	**Settings**	**Country**	**Language**	**Name**	**Construct**	**Number of subscales and items**	**Version**
[60]	Mixed cancer diagnoses	75 adult oncology patients receiving pain treatment (44%)	Not provided 88% > 45 years	Inpatient and outpatient in departments of a major tertiary care centre in Beirut	Lebanon	Arabic	BPI	Pain	2 subscales with 15 items (severity and interference)	Arabic BPI
[61]	Mixed cancer diagnoses	104 cancer patients with pain (44.5%)	48.6 (not provided)	Inpatient and outpatient in the Pain Therapy and Palliative Care Division of the National Cancer Institute of Milan.	Italy	Italian	BPI	Pain	2 subscales with 15 items (severity and interference)	Italian (BQVD)
[65]	Mixed cancer diagnoses	121 patients experiencing pain (45%)	56 (not provided)	The Saitama Cancer Centre	Japan	Japanese	BPI	Pain	2 subscales with 15 items (severity and interference)	Japanese (BPI-J)
[59]	Mixed cancer diagnoses	- 113 cancer patients with pain (37.2%) - 40 patients re-interviewed	45.7 years (±16.84)	Inpatients and outpatients at Hospital University Sains Malaysia (HUSM) and Hospital Kota	Malaysia	Malay	BPI	Pain	2 subscales with 15 items (severity and interference)	Malay (BPI-M)

				Bharu (HKB), Kelantan						
[62]	Mixed cancer diagnoses	143 cancer patients with pain (61.54%)	57.3 years (±13.28)	Outpatients at Hospital das Clinicas, University of Sao Paulo	Brazil	Brazilian Portuguese	BPI	Pain	2 subscales with 15 items (severity and interference)	Brazilian (BPI-B)
[63]	Mixed cancer diagnoses	220 cancer patients with pain (44%)	61.3 years (±14.84)	Patients at two Greek national cancer centres (Areteion Hospital and the Koropi Health Centre)	Greece	Greek	BPI	Pain	2 subscales with 15 items (severity and interference)	Greek (G-BPI)
[64]	Mixed cancer diagnoses	300 cancer patients with pain (100 bilingual 40% female, 200 monolingual 55% female)	46 years (±13) for bilingual and 85 years (±12) for monolingual	Inpatients and outpatients at cancer referral centre in north India	India	Hindi	BPI	Pain	2 subscales with 15 items (severity and interference)	Hindi (BPI-H)
[66]	Mixed cancer diagnoses	147 cancer patients with pain (42%)	54 years (not provided) ranged from 18 to 86 years	Inpatients and outpatients at three cancer hospitals in Beijing.	China	Chinese	BPI	Pain	2 subscales with 15 items (severity and interference)	Chinese (BPI-C)
[67]	Mixed cancer diagnoses	221cancer patients with pain (62%)	62 years (±14.1)	Inpatients and outpatients at four St. Petersburg hospitals: the City Oncological Centre (surgery,	Russia	Russian	BPI-SF	Pain	2 subscales with 11 items (severity and interference)	Russian (BPI-R)

				radiotherapy, and chemotherapy departments), the Russian Military Medical Academy (surgery, hematology, and clinical immunology departments), District Hospice No. 3, and the City Hospital No. 15, Kirovsky District (hematology unit)						
[68]	Mixed cancer diagnoses	235 cancer patients with pain (44.7%)	Mean and SD not provided (median = 63 years and range 29–89)	Inpatients at University Hospital of Trondheim	Norway	Norwegian	BPI-SF	Pain	2 subscales with 11 items (severity and interference)	Norwegian BPI
[23]	Mixed cancer diagnoses	119 Asian American cancer patients (82.4%)	52.2 years (±10.9)	Patients were recruited through both Internet (Web sites of the internet cancer support groups) and real clinical setting	USA	English	MPQ-SF and BPI-SF	Pain	MPQ-SF: 2 subscales with 15 items (sensory and affective) BPI-SF: 2 subscales with 11 items (severity and interference)	English MPQ-SF and BPI-SF

[22]	Mixed cancer diagnoses	269 advanced cancer patients with pain (%not provided)	Not provided – age ranged from 21 to ≥60	Outpatient clinics at a comprehensive cancer centre and those receiving home palliative care in Toronto, Ontario	Canada	English	MPQ-SF-2	Pain	4 subscales with 24 items (continuous, intermittent, neuropathic and affective)	English MPQ-SF-2
[69]	Mixed cancer diagnoses	- 114 cancer patients with pain (49.1%) - 80 patients repeated the test	62.90 years (±10.38)	Inpatients and outpatients at the Pain Relief and Palliative Care Unit at the University of Athens	Greece	Greek	MPQ	Pain	4 subscales with 24 items (sensory, affective, evaluative, miscellaneous)	Greek (G-MPQ)
[70]	Mixed cancer diagnoses	- 97 Patients with pain or sensory complaints enrolled at clinic visits, during a chemotherapy sessions or at presentation to the clinic. - 39 patients repeated the test.	54.1 (not provided)	The Rena Rowan Breast Cancer Centre, the general oncology clinics and the Penn Pain Medicine Centre of the University of Pennsylvania	USA	English	L-BASIC	Location based sensory symptoms	4 items	English
[71]	Mixed cancer diagnoses	262 patients with pain (62%)	52.1 years (±14.3)	Inpatient and outpatient from in an oncology	USA	English	CPI	Pain concerns	5 subscales with 19 items (Catastrophizing,	English

				clinic of an NCI-designated comprehensive cancer centre in a Midwestern metropolitan area					Interference, Stoicism, Social Aspects, and Pain Medication)	
[72]	Mixed cancer diagnoses	98 Ambulatory advanced cancer patients (50%)	56.7 years (±11.6)	Two hospital-based oncology clinics in Quebec City, Quebec	Canada	French	Brief 4-week pain diary	Pain	7 items in two sections: 1- Pain indicators: pain intensity and the number of opioid rescue doses, 2- Pain impact of quality of life indicator (5 items)	French

Abbreviation: SD = stander deviation, BPI = brief pain inventory, L-BASIC = location-based assessment of sensory symptoms in cancer, CPI = cancer pain inventory, MPQ= McGill pain questionnaire, MPQ-SF = McGill pain questionnaire-short form, BPI-SF = brief pain inventory-short form.

Table 2: Measurement properties of the identified tool

Tool (version)	Article	Internal consistency	Reliability	Construct validity	Structural validity	Cross-cultural validity
BPI (Arabic)	[60]	Cronbach alpha for the two subscales: 0.82 (severity) and 0.92 (interference)		- Positive correlation between pain now item and VAS (r = 0.68, P<0.01). - Negative correlation between pain on average and the item about relief provided by pain treatment (r = - 0.19, P = 0.10) - Higher severity scores in patients with metastases compared with those without metastases (5.7±1.7 vs. 4.9±1.5, P=0.02) - Significant correlation between severity and interference items (r = 0.63, no p-value provided)	PCA with oblimin rotation: 2 factors (1- severity explaining 11.3% and 2- interference explaining 55.8%)	Forward and backward translation method
BPI (Italian)	[61]	Cronbach alpha for both subscales: 0.78		- Positive strong correlation between the composite score for the interference subscale and the Therapy Impact Questionnaire (TIQ) composite based on activity and affect items (r = 0.66.) - Strong correlation between composite score for severity subscale and the pain item in the TIQ (r = 0.45). No P-value provided	PAF with direct oblimin rotation: 2-factor solution (1- interference explaining 31.4 % and 2- severity explaining 14.8%)	

BPI (Japanese)	[65]	Cronbach alpha for both subscales: 0.81			PAF with non-orthogonal (oblimin) rotation: 2-factor solution (1-interference and 2-severity) % of variance explained by each factor is not provided	Forward and backward translation method
BPI (Malay)	[59]	Cronbach alpha for the two subscales: 0.81 (intensity) and 0.88 (interference)	ICC for the two subscales: 0.61 (intensity) and 0.87 (interference)	High negative correlation with KPS (r ranged from -0.52 to -0.73, P<0.001)	PAF with direct oblimin rotation: 2 factors: 1-intensity (4 items), 2-interference (7 items); total variance explained 62.1%	Forward and backward translation method
BPI (Brazilian)	[62]	Cronbach alpha for the two subscales: 0.91 (severity) and 0.87 (interference)		- Positive and moderate to strong correlations with MPQ, and EORTC-QLQ30 pain scale (r ranged from 0.38 to 0.90, P<0.05) - Patients with poor performance status had greater pain than those with high performance status (scores ranged from 6.20 to 6.96, vs. 4.32 to 4.95, P=0.000 to 0.007) - Patients with metastatic disease had higher pain than patients with local or regional disease (scores ranged from 5.26 to 6.04 vs. 4.20 to 4.28, P=0.012 to 0.042)	CFA using structural equation modelling: a model with two dimensions (severity and interference) showed a good fit to data: GFI= 0.82, CFI=0.95, NFI=0.91, NNFI=0.94 and RMSEA= 0.11	Forward and backward translation method
BPI (Greek)	[63]	Cronbach alpha for the two subscales: 0.89		Positive weak to moderate correlation with KPS scale (r	PAF with nonorthogonal (direct oblimin) rotation:	Forward and backward

		(severity) and 0.85 (interference)	ranged from 0.2 to 0.4; P<0.0005 to 0.003)	2 factors 1- interference (7 items) 2- severity (4 items) explained 44% and 19%, respectively of total variance	translation method
BPI (Hindi)	[64]	Cronbach alpha for the two subscales: 0.89 and 0.88 (severity) and 0.91 and 0.78 (interference) for the bilingual and monolingual samples respectively		PAF with nonorthogonal rotation: 2 factors for the bilingual sample: 1- interference (7 items) 2- severity (4 items); 3 factors for the monolingual sample: 1- severity (4 items), 2- mood-related interference (3 items), 3- activity-related interference (3 items) % of variance explained by each factor is not provided	- Forward and backward translation method - Alternate-form reliabilities for the two subscales in the English and Hindi version are 0.88 and 0.95 - The factor structure is similar across both versions
BPI (Chinese)	[66]	Cronbach alpha for the two subscales: 0.89 (severity) and 0.92 (interference)	- Positive moderate correlation with ECOG performance status (r 0.33, P<0.05) - Strong correlation between pain severity and greater pain interference (r 0.60, P<0.05)	PAF with oblimin rotation: 2 factors 1- interference (7 items) 2- severity (4 items) both explained 72% of total variance	Forward and backward translation method
BPI-SF (Russian)	[67]	Cronbach alpha for the two subscales: 0.93 (severity) and 0.95 (interference)	- Patients with poor performance status had greater pain than patients with good performance status	PAF with direct oblimin rotation: 2 factors 1- interference (7 items) 2- severity (4 items)	Forward and backward translation method

			(scores ranged from 3.2 to 3.6 vs. 1.5 to 1.7, P<0.001)	explained 73.5% and 6.5%, respectively of total variance	
BPI-SF (Norwegian)	[68]	Cronbach alpha for the three subscales: 0.87 (severity), 0.92 (physical interference) and 0.91 (psychological interference)	Positive strong correlations with intensity and influence items in EORTC-QLQ30 (r ranged from 0.62 to 0.70; P<0.001)	PAF with oblimin rotation: 3-factor model:1- physical interference (3 items), 2- psychological interference (4 items), 3- pain severity (4 items) explained 82% of total variance	Forward and backward translation method
BPI-SF and MPQ-SF (English)	[23]	Cronbach alpha for total scale: 0.94 for the MPQ-SF, and 0.97 for the BPI-SF	- Positive moderate correlation between the total pain scores from the MPQ-SF and the BPI-SF and the usage of pain medications (r = 0.25 and 0.42 respectively, P<0.01) - Higher mean total pain scores in the pain medication group than the no pain medication group for the MPQ-SF (9.30, 3.86) and the BPI-SF (48.00, 20.56) - For MPQ-SF, the differences of the total scores and the sensory dimension scores between the two groups were not significant	PCA with varimax rotations: - 2 factors for MPQ-SF: 1- sensory (11 items explaining 57.06 % of the variance), 2- affective (4 items explaining 8.66 of the variance) - 2 factors for BPI-SF:1- sensory (4 items explaining 71.96 % of the variance), 2- reactive (7 items explaining 10.35 % of the variance	
MPQ-SF-2 (English)	[22]	Cronbach alpha for total scale: 0.89 for younger group and 0.93 for older group	- For both groups, positive strong correlation with BPI average pain (r=0.67, 0.55, P≤0.01)	CFA using SEM with maximum likelihood estimation: Model fit assessed with the SRMR	

				- Moderate correlation with CES-D (r=0.27, 0.35), negative with Pain Relief (r=-0.34, -0.30) and with SF-36 physical health QOL (r=-0.23, -0.32) and with KPS (r=-0.25, -0.29) all at P≤0.01	=0.09, RMSEA =0.07, CFI=0.77, TLI=0.74 and AIC=927.39	
MPQ (Greek)	[69]	Cronbach alpha for the descriptor scale 0.96	Positive weak to moderate correlation between pre and post treatment for PRI,PPI and NWC in the scale items (r ranged from 0.23 to 0.44; P 0.0005 to 0.045)	- Positive moderate to strong correlation between PRI,PPI and NWC in the scale items (r ranged from 0.42 to 0,92; P<0.0001) - Significant difference (P<0.005) between pre-treatment and post-treatment scores except for the PRI-evaluative item - Significant difference (P<0.05) on the scale scores between patient with different performance status	PCA with varimax rotation: 2 factors: 1-sensory, affective and evaluative, 2-miscellaneous; explained 75% and 20.2% of total variance respectively	Forward and backward translation method
L-BASIC (English)	[70]	Cronbach alpha for total scale: 0.74	- Using Kappa statistic, good strict and relaxed agreement of location (k=0.76; 95% CI=0.66-0.86) and descriptor categories (k=0.80; 95% CI=0.70-0.89) used by a given patient without clinical change (n 32) at the two time points (interval median=14 days)	- Significant correlation between the global body score and every item on the BPI pain (r range: 0.47-0.61) and functional interference subscales (r range: 0.22-0.49). - Significant correlations between worst body part score and every BPI item pain subscale item (r range: 0.54-0.60) and functional interference subscale item		

		- strong correlation between intensity values (r=0.72) and unpleasantness values (r=0.66) for 109 concordance body parts reported at test retest - Moderate correlations for the number of adjectival descriptors used per body part (r=0.48) and the mean adjectival descriptor severity weights (r=0.31) between test and retest. - Fair strict body part-to-body part agreement of adjectival descriptor categories used to score the 109 matched regions between the two time points (k=0.56; 95% CI=0.50-0.62)	except pain-related interference with appetite (r range: 0.35-0.53) - Only a significant correlation between worst body part score and the pain item on the physical well-being subscale (r=0.33) of the FACT-G - Difference in the sensory qualities of the three distinct clinical constructs (head and neck cancer, breast cancer related upper extremity lymphedema, and chemotherapy related neuropathy) derived from difference in frequency of descriptor category among the three (X^2=223; P<0.001) - Significant correlation with the worst, least, average, and right now pain items on the BPI (r range: 0.59-0.67) for head (n=11) and neck (n=14) patients - No significant associations with items on the BPI or FACT-G for patients with upper extremity lymphedema secondary to breast cancer treatment (n=27) as well as for patients with neuropathy (n=32); no r or P values

			provided.	
CPI (English)	[71]	Cronbach alpha for total scale: 0.82, for the five subscales: 0.82 (Catastrophizing); 0.70 (Interference with Functioning); 0.62 (Stoicism); 0.51 (Social Aspects); and 0.63 (Pain Medication).	-Moderate correlations between the CPI subscales and each the measures of pain (BPI: r=0.38 for severity and r=0.45 for interference), pain disability (PDI: r=0.42), and depressed mood (CES-D: r=0.55) for all P<0.01. - Strong correlation between the CPI Interference subscale and each of the SOPA Disability subscale (r=0.6), the BPI Interference subscale (r=0.53), and the PDI Sum score (r=0.56) for all P<0.01. - Multiple positive correlations between Catastrophizing subscale and measures of disability and distress (r=0.45,0.48,0.40; P<0.01) and strongly negative correlation with expectations for a medical cure (r=-0.40; P<0.01) - Positive and moderate correlation between the Social Aspect subscale and the Depression measure (r=0.39) and negative with expectations for a medical cure (r=-0.18) P<0.01 -In general, low correlation between the Stoicism and	PCA with varimax rotation: 5 factors: 1- Catastrophizing (25.7% of variance), 2- Interference (10.4% of variance), 3- Stoicism (7.5% of variance), 4- Social Aspects (6.5% of variance), and Pain Medication (5.9% of variance)

			Medication subscales and the other measures subscales (r ranged from 0.0 to 0.24) - A significant difference between inpatients (mean=2.08) and outpatients (mean=1.74) scores for the CPI subscale measuring interference with functioning (P<0.05) - A significant difference between patients reporting higher pain (≥4 on BPI) (HP) from patients with lower levels of pain (LP) in scoring significantly higher on each of the CPI subscales: Catastrophizing (HP =2.24, LP=1.75, P<0.001), Interference (HP =2.35, LP=1.69, P<0.001), Stoicism (HP =2.26, LP=2.05, P<0.05), Social Aspects (HP =1.76, LP=1.50, P<0.001), Pain Medication (HP =2.15, LP=1.87, P<0.05)
Brief 4-week pain diary (French)	[72]	Cronbach alpha for the pain quality of life impact indicator ranged from 0.87 to 0.92	- Positive strong correlations between pain intensity and both rescue doses and pain impact on quality of life indicators (r ranged from 0.64 to 0.73 in Weeks 1–4, P <0.01) - Significant difference between patients' pain levels

		in relation to change in pain impact on quality of life indicators (p value ranged from 0.0004 to 0.05) - Positive correlation between each of: pain intensity, rescue doses, and pain impact on quality of life and the EORTC-QLQ30 symptom scales (r ranged from 0.30 to 0.73; P<0.01) (with the exception of appetite loss); and negatively correlated with the EORTC-QLQ30 functioning scales (r ranged from -0.26 to -0.62; P <0.05)

Note: The measurement error, content validity and responsiveness properties were deleted from the table because they were not evaluated by any study.

Abbreviation: BPI = brief pain inventory, PCA = principal component analysis, CFA = confirmatory factor analysis, L-BASIC = location-based assessment of sensory symptoms in cancer, FACT-G: Functional Assessment of Cancer Therapy-General, CPI = cancer pain inventory, CES-D = Centre for Epidemiological Studies—Depression scale, PDI = Pain Disability Index, SOPA = Survey of Pain Attitudes, MPQ-SF = McGill pain questionnaire-short form, SEM = structural equation modelling, SRMR = standardized root mean square residual, RMSEA = root mean square error of approximation, CFI = comparative fit index, TLI = Tucker- Lewis index, AIC = Akaike information criterion, KPS = Karnofsky performance status, EORTC-QLQ30 = the European Organization for Research and Treatment of Cancer Core Quality of Life Questionnaire, BPI-SF = brief pain inventory-short form, PAF = principal axis factor analysis, ICC = intraclass correlation coefficient, MPQ = McGill pain questionnaire, GFI = goodness- of-fit index, NFI = normed fit index, NNFI = non-normed fit index, CFI = comparative fit index, RMSEA = root-mean-square error of approximation, PPI = present pain index, PRI = pain rating index, NWC = number of words chosen, ECOG = Eastern Cooperative Oncology Group.

Table 3: Methodological quality of the studies and quality of results reported per measurement property and tool.

Tool	Article	Internal consistency		Reliability		Construct validity		Structural validity		Cross-cultural validity	
		M[1]	Q[2]	M	Q	M	Q	M	Q	M	Q
BPI	[60]	Poor	+			Poor	+	Poor	-	Poor	?
	[61]	Good	+			Poor	+	Fair	-		
	[65]	Good	+					Fair	?	Poor	?
	[59]	Excellent	+	Poor	-	Fair	+	Excellent	?	Poor	?
	[62]	Fair	+			Fair	+	Fair	-	Poor	?
	[63]	Fair	+			Poor	+	Fair	-	Poor	?
	[64]	Fair	+					Fair	?	Poor	?
	[66]	Fair	+			Poor	+	Fair	?	Poor	?
BPI-SF	[67]	Fair	+			Poor	+	Fair	+	Poor	?
	[68]	Good	+			Fair	+	Good	?	Poor	?
	[23]	Fair	?			Fair	+	Fair	+		
MPQ-SF	[23]	Fair	+			Fair	-	Fair	+		
MPQ-SF-2	[22]	Excellent	+			Fair	+	Excellent	-		
MPQ	[69]	Poor	-	Fair	?	Poor	+	Poor	-	Poor	?
L-BASIC	[70]	Fair	+	Fair	+	Poor	+				
CPI	[71]	Fair	+			Fair	+	Fair	-		
Brief 4-week pain diary	[72]	Poor	+			Fair	+				

Note: The measurement error, content validity, and responsiveness properties were deleted from the table because they were not evaluated by any study.

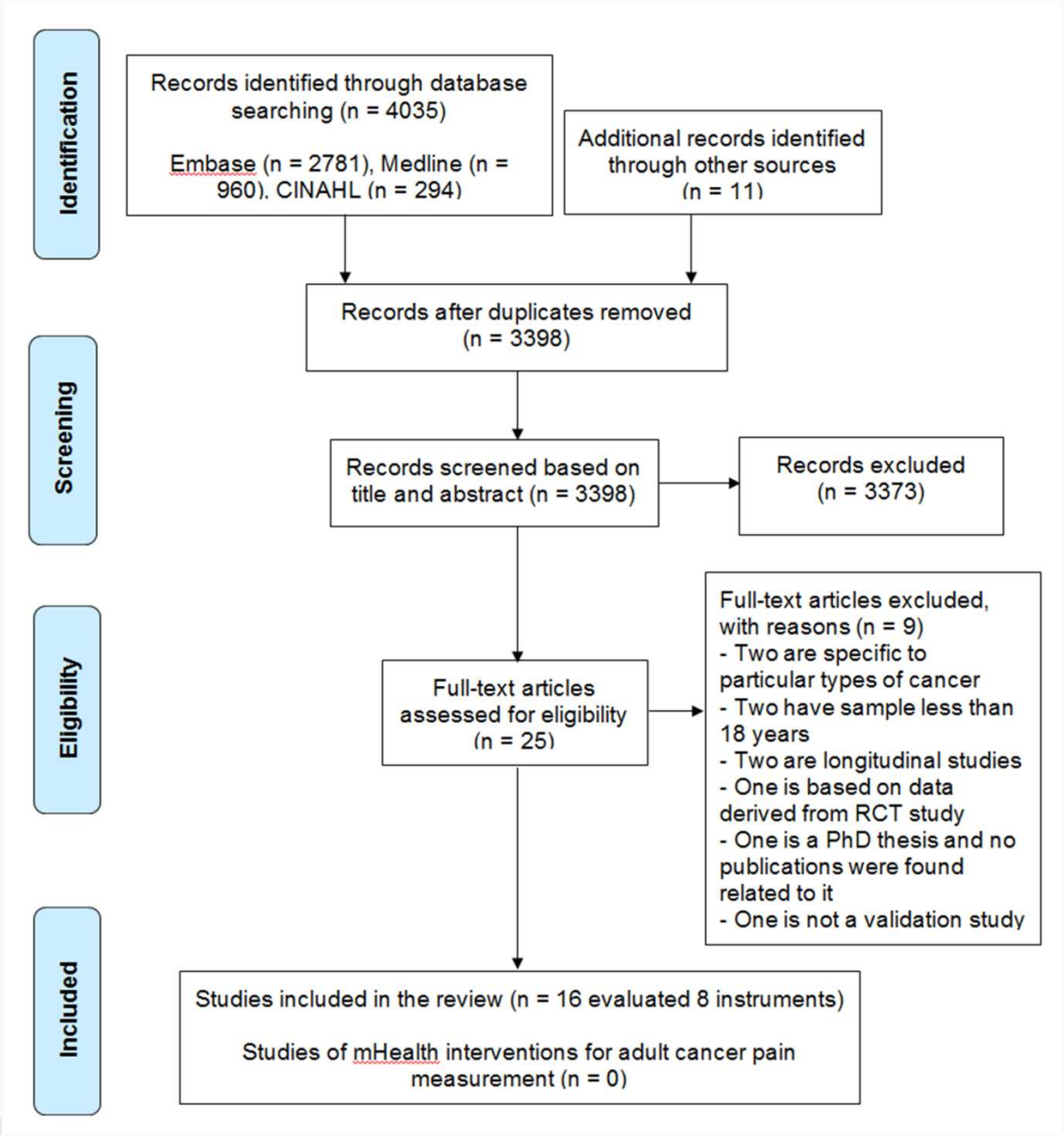
[1] M = Methodological quality of the study rated as excellent, good, fair and poor; [2] Q = Quality of the results rated: + = positive rating, ? = indeterminate rating, - = negative rating.

Abbreviation: BPI = brief pain inventory, L-BASIC = location-based assessment of sensory symptoms in cancer, CPI = cancer pain inventory, MPQ-SF = McGill pain questionnaire-short form, BPI-SF = brief pain inventory-short form

Table 4: Evidence synthesis of the measurement properties of the cancer pain measurement tools.

Tool	Internal consistency	Reliability	Construct validity	Structural validity	Cross-cultural validity
BPI	+++	?	++	--	?
BPI-SF	++	NA	++	++	?
MPQ-SF	+	NA	-	+	NA
MPQ-SF-2	+++	NA	+	---	NA
MPQ	?	?	?	?	?
L-BASIC	+	+	?	NA	NA
CPI	+	NA	+	-	NA
Brief 4-week pain diary	?	NA	+	NA	NA

Note: +++/--- = strong positive/ negative evidence; ++/-- = moderate positive/negative evidence; +/- = limited positive/negative evidence; +/- = conflicting findings; ? = unknown evidence; NA = no information available.



Appendices

Appendix 1

COSMIN's definitions of the measurement properties.

Domain	Measurement property	Definition from Mokkink *et al* [38]
Reliability		*"The degree to which the measurement is free from measurement error."*
Reliability (extended definition)		*"The extent to which scores for patients who have not changed are the same for repeated measurement under several conditions: for example, using different sets of items from the same HR-PROs (internal consistency), over time (testeretest) by different persons on the same occasion (interrater) or by the same persons (i.e., raters or responders) on different occasions (intrarater)."*
	Internal consistency	*"The degree of the interrelatedness among the items."*
	Reliability	*"The proportion of the total variance in the measurements which is because of ''true'' differences among patients."*
	Measurement error	*"The systematic and random error of a patient's score that is not attributed to true changes in the construct to be measured."*
Validity		*"The degree to which an HR-PRO instrument measures the construct(s) it purports to measure."*
	Content validity	*"The degree to which the content of an HR-PRO instrument is an adequate reflection of the construct to be measured."*
	Construct validity (or hypotheses testing)	*"The degree to which the scores of an HR-PRO instrument are consistent with hypotheses (for instance with regard to internal relationships, relationships to scores of other instruments, or differences between relevant groups) based on the assumption that the HR-PRO instrument validly*

		measures the construct to be measured."
	Structural validity	*"The degree to which the scores of an HR-PRO instrument are an adequate reflection of the dimensionality of the construct to be measured."*
	Cross-cultural validity	*"The degree to which the performance of the items on a translated or culturally adapted HR-PRO instrument are an adequate reflection of the performance of the items of the original version of the HR-PRO instrument."*
	Criterion validity	*"The degree to which the scores of an HR-PRO instrument are an adequate reflection of a ''gold standard''."*
Responsiveness	Responsiveness	*"The ability of an HR-PRO instrument to detect change over time in the construct to be measured."*

38. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. J Clin Epidemiol. 2010;63:737-45.

Appendix 2

Search strategy applied on Medline (Ovid 1996 to March 2018) as an example.

Concept	Search terms (number of records)
Population	1 exp Neoplasms/ (1844343) 2 (carcin*or cancer* or tumo?r* or neoplasm* or adenocarcino*).tw. (1022908) 3 exp Adult/ (4229928) 4 adult.tw. (435643) 5 1 or 2 (2056450) 6 3 or 4 (4459633) 7 5 and 6 (919871)
Intervention	8 *Pain Measurement/ (10338) 9 ((pain* or sore* or hurt* or discomfort* or uncomfort* or cramp* or irritat* or analges*) adj3 (measur* or assess* or scal* or scor* or rat* or self report* or self management or self rat* or validated measurement or evaluat* or quantif* or Inventory or inventories)).tw. (72972) 10 exp Pain/ (240689) 11 exp Analgesia/ (22756) 12 8 or 9 or 10 or 11 (280293) 13 exp Software/ (117330) 14 (tool* or electronic* or device* or app* or machine or instrument* or questionnaire*).tw. (4286417) 15 13 or 14 (4330910) 16 12 and 15 (99102)
Comparison	Not applicable
Outcome	17 (validat* or validity or reliab* or objectivit* or clinimetric* or evaluation or responsive* or (psychometr* and propert*) or (cronbach* and alpha) or correlation or coefficient or internal consistency or Cohen* kappa or test retest or variability or standard error of measurement or sensitivity or specificity or hypotheses testing).tw. (2311433) 18 ((minimal* or meaning* or detectabl* or important* or effectiv* or relevant*) and (difference* or change* or improv* or shift* or alteration* or deterioration* or respons* or efficacy or effectiveness)).tw. (1624912) 19 exp Psychometrics/ (52213) 20 exp observer variation/ (34804) 21 17 or 18 or 19 or 20 (3591050)
Combining Population, Intervention & Outcome	22 7 and 16 and 21 (2803)
Adding the Oxford filter for PRO measures	23 (HR-PRO or HRPRO or HRQL or HRQoL or QL or QoL).ti,ab. or quality of life.mp. or (health index* or health indices or health profile*).ti,ab. or health status.mp. or ((patient or self or child or parent or carer or proxy) adj (appraisal* or appraised or report or reported or reporting or rated or rating* or based or assessed or assessment*)).ti,ab. or ((disability or function or functional or functions or subjective or utility or utilities or wellbeing or well being) adj2 (index or indices or instrument or instruments or measure or

	measures or questionnaire* or profile or profiles or scale or scales or score or scores or status or survey or surveys)).ti,ab. (479242) 24 22 and 23 (1018)
Restricting to English language publications	25 limit 24 to English language (960)

Search strategy used on Embase (Ovid 1996 to March 2018):

1 exp Neoplasms/ (2969135)

2 (carcin*or cancer* or tumo?r* or neoplasm* or adenocarcino*).tw. (1666438)

3 exp Adult/ (5284274)

4 adult.tw. (667901)

5 1 or 2 (3313902)

6 3 or 4 (5551551)

7 5 and 6 (1175115)

8 *Pain Measurement/ (331)

9 ((pain* or sore* or hurt* or discomfort* or uncomfort* or cramp* or irritat* or analges*) adj3 (measur* or assess* or scal* or scor* or rat* or self report* or self management or self rat* or validated measurement or evaluat* or quantif* or Inventory or inventories)).tw. (130140)

10 exp Pain/ (971696)

11 exp Analgesia/ (116744)

12 8 or 9 or 10 or 11 (1039377)

13 exp Software/ (79567)

14 (tool* or electronic* or device* or app* or machine or instrument* or questionnaire*).tw. (6561576)

15 13 or 14 (6594765)

16 12 and 15 (370305)

17 (validat* or validity or reliab* or objectivit* or clinimetric* or evaluation or responsive* or (psychometr* and propert*) or (cronbach* and alpha) or correlation or coefficient or internal consistency or Cohen* kappa or test retest or variability or standard error of measurement or sensitivity or specificity or hypotheses testing).tw. (3670061)

18 ((minimal* or meaning* or detectabl* or important* or effectiv* or relevant*) and (difference* or change* or improv* or shift* or alteration* or deterioration* or respons* or efficacy or effectiveness)).tw. (2628261)

19 exp Psychometrics/ (70994)

20 exp observer variation/ (17021)

21 17 or 18 or 19 or 20 (5688685)

22 7 and 16 and 21 (11418)

23 (HR-PRO or HRPRO or HRQL or HRQoL or QL or QoL).ti,ab. or quality of life.mp. or (health index* or health indices or health profile*).ti,ab. or health status.mp. or ((patient or self or child or parent or carer or proxy) adj (appraisal* or appraised or report or reported or reporting or rated or rating* or based or assessed or assessment*)).ti,ab. or ((disability or function or functional or functions or subjective or utility or utilities or wellbeing or well being) adj2 (index or indices or instrument or instruments or measure or measures or questionnaire* or profile or profiles or scale or scales or score or scores or status or survey or surveys)).ti,ab. (836690)

24 22 and 23 (2876)

25 limit 24 to english language (2781)

Search strategy used on CINAHL (EBSCO 1981 to March 2018):

MY

EBSCOhost

Monday, March 19, 2018 10:27:13 AM

#	Query	Limiters/Expanders	Last Run Via	Results
S33	S21 AND S31	Limiters - English Language Search modes - Boolean/Phrase	Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - CINAHL	294
S32	S21 AND S31	Search modes - Boolean/Phrase	Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - CINAHL	295
S31	S23 OR S24 OR S25 OR S26 OR S27 OR S30	Search modes - Boolean/Phrase	Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - CINAHL	166,231
S30	S22 OR S28 OR S29	Search modes - Boolean/Phrase	Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - CINAHL	111,178
S29	AB (disability or function or functional or functions or subjective or utility or utilities or wellbeing or well being) n2 (index or indices or instrument or instruments or measure or measures or questionnaire* or profile or profiles or scale or scales or score or scores or status or survey or surveys)	Search modes - Boolean/Phrase	Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - CINAHL	30,777
S28	AB (patient or self or child or parent or carer or proxy) n1 (appraisal* or appraised or report or reported or reporting or	Search modes - Boolean/Phrase	Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - CINAHL	83,702

	rated or rating* or based or assessed or assessment*)			
S27	AB health status	Search modes - Boolean/Phrase	Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - CINAHL	17,241
S26	AB (health index* or health indices or health profile	Search modes - Boolean/Phrase	Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - CINAHL	6,059
S25	AB quality of life	Search modes - Boolean/Phrase	Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - CINAHL	54,826
S24	AB QoL	Search modes - Boolean/Phrase	Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - CINAHL	6,970
S23	AB QL	Search modes - Boolean/Phrase	Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - CINAHL	176
S22	AB HR-PRO or HRPRO or HRQL or HRQoL	Search modes - Boolean/Phrase	Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - CINAHL	3,643
S21	S7 AND S16 AND S20	Search modes - Boolean/Phrase	Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - CINAHL	617
S20	S17 OR S18 OR S19	Search modes - Boolean/Phrase	Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - CINAHL	474,761

S19	(MM "Psychometrics")	Search modes - Boolean/Phrase	Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - CINAHL	2,185
S18	AB ((minimal* or meaning* or detectabl* or important* or effectiv* or relevant*) and (difference* or change* or improv* or shift* or alteration* or deterioration* or respons* or efficacy or effectiveness))	Search modes - Boolean/Phrase	Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - CINAHL	244,635
S17	AB (validat* or validity or reliab* or objectivit* or clinimetric* or evaluation or responsive* or (psychometr* and propert*) or (cronbach* and alpha) or correlation or coefficient or internal consistency or Cohen* kappa or test retest or variability or standard error of measurement or sensitivity or specificity or hypotheses testing)	Search modes - Boolean/Phrase	Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - CINAHL	283,917
S16	S12 AND S15	Search modes - Boolean/Phrase	Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - CINAHL	41,913
S15	S13 OR S14	Search modes - Boolean/Phrase	Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - CINAHL	786,201
S14	AB tool* or electronic* or device* or app* or machine or instrument* or questionnaire*	Search modes - Boolean/Phrase	Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - CINAHL	618,061
S13	(MH "Software+")	Search modes - Boolean/Phrase	Interface - EBSCOhost Research Databases	251,111

			Search Screen - Advanced Search Database - CINAHL	
S12	S8 OR S9 OR S10 OR S11	Search modes - Boolean/Phrase	Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - CINAHL	133,853
S11	(MH "Analgesia+")	Search modes - Boolean/Phrase	Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - CINAHL	8,163
S10	(MH "Pain+")	Search modes - Boolean/Phrase	Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - CINAHL	116,376
S9	((pain* or sore* or hurt* or discomfort* or uncomfort* or cramp* or irritat* or analges*) n3 (measur* or assess* or scal* or scor* or rat* or self report* or self management or self rat* or validated measurement or evaluat* or quantif* or Inventory or inventories))	Limiters - Abstract Available Search modes - Boolean/Phrase	Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - CINAHL	36,968
S8	(MM "Pain Measurement")	Search modes - Boolean/Phrase	Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - CINAHL	4,749
S7	S5 AND S6	Search modes - Boolean/Phrase	Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - CINAHL	98,192
S6	S1 OR S2	Search modes - Boolean/Phrase	Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - CINAHL	244,321

S5	S3 OR S4	Search modes - Boolean/Phrase	Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - CINAHL	954,466
S4	AB adult*	Search modes - Boolean/Phrase	Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - CINAHL	136,853
S3	(MH "Adult+")	Search modes - Boolean/Phrase	Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - CINAHL	906,221
S2	AB carcin*or cancer* or tumo?r* or neoplasm* or adenocarcino*	Search modes - Boolean/Phrase	Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - CINAHL	15,929
S1	(MH "Neoplasms+")	Search modes - Boolean/Phrase	Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - CINAHL	240,806

Appendix 3

The quality criteria for good measurement properties modified from Terwee *et al. [34]* and available on http://www.cosmin.nl/

Measurement property	Rating*	Criteria
Internal consistency	+	At least limited evidence for unidimensionality or positive structural validity AND Cronbach's alpha(s) ≥ 0.70 and ≤ 0.95
	?	Not all information for '+' reported OR conflicting evidence for unidimensionality or structural validity OR evidence for lack of unidimensionality or negative structural validity
	-	Criteria for '+' not met
Reliability	+	ICC or weighted Kappa ≥ 0.70
	?	ICC or weighted Kappa not reported
	-	Criteria for '+' not met
Measurement error	+	SDC or LoA < MIC
	?	MIC not defined
	-	Criteria for '+' not met
Content validity	+	All items refer to relevant aspects of the construct to be measured AND are relevant for the target population AND are relevant for the purpose of the measurement instrument AND together comprehensively reflect the construct to be measured
	?	Not all information for '+' reported
	-	Criteria for '+' not met
Construct validity (or hypotheses testing)	+	At least 75% of the results are in accordance with the hypotheses
	?	No correlations with instrument(s) measuring related construct(s) AND no differences between relevant groups reported
	-	Criteria for '+' not met
Structural validity	+	Unidimensionality: EFA: First factor accounts for at least 20% of the variability AND ratio of the variance explained by the first to the second factor greater than 4 OR Bi-factor model: Standardized loadings on a common factor >0.30 AND correlation between individual scores under a bi-factor and unidimensional model >0.90

		Structural validity: CFI or TLI or comparable measure >0.95 AND (Root Mean Square Error of Approximation (RMSEA) <0.06 OR Standardized Root Mean Residuals (SRMR)<0.08)
	?	Not all information for '+' reported
	-	Criteria for '+' not met
Cross-cultural validity	+	No important differences found between language versions in multiple group factor analysis or DIF analysis
	?	Multiple group factor analysis AND DIF analysis not performed
	-	One or more criteria for '+' not met
Criterion validity	+	Convincing arguments that gold standard is "gold" AND correlation with gold standard ≥ 0.70
	?	Not all information for '+' reported
	-	Criteria for '+' not met
Responsivenes	+	At least 75% of the results are in accordance with the hypotheses
	?	No correlations with changes in instrument(s) measuring related construct(s) AND no differences between changes in relevant groups reported
	-	Criteria for '+' not met

CFI = comparative fit index; DIF = differential item functioning; EFA= exploratory factor analysis; ICC = intraclass correlation coefficient; LoA = limits of agreement; MIC = minimal important change; SDC = smallest detectable change; TLI = Tucker-Lewis index

* + = positive rating, ? = indeterminate rating, – = negative rating

34. Terwee CB, Bot SD, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol*. 2007;**60:**34-42.