*"This is the peer reviewed version of the following article: [*LORENTE S., LOSILLA J. M. & VIVES J. (2018) Instruments to assess the patient comfort during hospitalization: a psychometric review. *Journal of Advanced Nursing*, *74* (5), 1001-1015. doi: 10.1111/jan.13495] ***This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Use of Self-Archived Versions***.

INSTRUMENTS TO ASSESS PATIENT COMFORT DURING HOSPITALIZATION: A PSYCHOMETRIC REVIEW

Sonia LORENTE, Josep-Maria LOSILLA, Jaume VIVES

Universitat Autònoma de Barcelona

Sonia LORENTE, Adjunct lecturer, MSc, Department of Psychobiology and Methodology of Health Science, Universitat Autònoma de Barcelona, UAB; Paediatric Nurse, PNP, Hospital de Terrassa, Consorci Sanitari de Terrassa, CST; Josep-Maria LOSILLA, Associate Professor, PhD, Department of Psychobiology and Methodology of Health Science, Universitat Autònoma de Barcelona, UAB; Jaume VIVES, Associate Professor, PhD, Department of Psychobiology and Methodology of Health Science, Universitat Autònoma de Barcelona, UAB.

This work was supported by the Grant PSI2014-52962-P, Spanish Ministry of Economy and Competitiveness (July, 2015). The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript. No conflict of interest has been declared by the author(s).

All authors have agreed on the final version, and meet at least one of the following criteria recommended by the International Committee of Medical Journal Editors, ICMJE: substantial contributions to conception and design, acquisition of data, analysis and interpretation of data, and drafting the article or revising it critically for important intellectual content. Correspondence concerning this article should be addressed to Jaume Vives, Associate Professor, PhD, Department of Psychobiology and Methodology of Health Science, Universitat Autònoma de Barcelona, UAB. Phone: +34935812331. E-mail: jaume.vives@uab.cat

ABSTRACT

Aim: To analyse the psychometric properties and the utility of instruments used to measure patient comfort, physical, social, psychospiritual and/or environmental, during hospitalization. **Background**: There are no systematic reviews nor psychometric reviews of instruments used to measure comfort, which is considered an indicator of quality in health care associated with quicker discharges, increased patient satisfaction and better cost-benefit ratios for the institution. **Design**: Psychometric review. **Data sources**: MEDLINE, CINAHL, PsycINFO, Web of Knowledge, ProQuest Thesis&Dissertations, Google. **Review methods**: We limited our search to studies published between 1990 and 2015. The psychometric analysis was performed using the COnsensus-based Standards for the selection of health status Measurement INstruments (COSMIN), along with the Quality Criteria for Measurement Properties. The utility of the instruments was assessed according to their cost-efficiency, acceptability and educational impact. Protocol registration in PROSPERO, CRD42016036290. **Results**: Instruments reviewed showed moderate methodological quality, and their utility was poorly reported. Thus, we cannot recommend any questionnaire without reservations, but the Comfort Scale, the General Comfort Questionnaire and their adaptations in adults and elderly patients, the Psychosocial Comfort Scale and the Incomfort des Patients de Reanimation are the most recommendable instruments to measure comfort. **Conclusions**: The methodology of the studies should be more rigorous and authors should adequately report the utility of instruments. This review provides a strategy to select the most suitable instrument to assess patient comfort according to their psychometric properties and utility, which is crucial for nurses, clinicians, researchers and institutions. **Keywords**: Psychometric Review, Validity of instruments, Patient comfort, Assessment tools, Utility of instruments, Nursing.

SUMMARY STATEMENT

Why is this review needed?

- There are no psychometric reviews of the health instruments used to assess patient comfort during hospitalization, being the comfort an indicator of quality in health care.
- Assessing the psychometric properties and the utility of the instruments measuring patient comfort is essential to improve the quality of care.

What are the key findings?

- Most instruments measuring patient comfort were valid and reliable, but no instrument can be completely recommended according to their utility, i.e., cost-efficiency, acceptability and educational impact.
- Methodologies used in studies on the measurement properties of instruments assessing patient comfort should be more accurate and authors should adequately report their utility.

How should the findings be used to influence policy/practice/research/education?

- Findings categorize the comfort instruments according to their measurement properties, allowing carers to select the most suitable comfort questionnaire in current clinical scenarios, which is essential for researchers and clinicians.
- Findings also provide a strategy to develop the most suitable comfort measurement instrument in accordance with their psychometric properties, utility and purpose, which is crucial for researchers and patients.
- Selecting the most appropriate instrument to assess comfort may improve the health care provided and patient satisfaction as well as reducing the institution costs.

INSTRUMENTS TO ASSESS PATIENT COMFORT DURING HOSPITALIZATION: A PSYCHOMETRIC REVIEW

INTRODUCTION

Measuring patient experience is essential to improving the quality of care provided in different health settings (WHO 2000, 2004, 2013). To date, different health care instruments have been developed to assess patient experience during hospitalization (Bruyneel et al., 2017; Manary, Boulding, Staelin, & Glickman, 2013; Pettersen, Veenstra, Guldvog, & Kolstad, 2004). In this sense, different systematic reviews have been developed to assess the measurement properties of these instruments, which are normally used to assess patient experiences related to pain reduction, care, hospital environment, or communication with professionals, to name a few (Beattie, Lauder, Atherton, & Murphy, 2015; Ellis-Smith et al., 2016). Nevertheless, there are no systematic reviews of the health instruments used to assess patient comfort during hospitalization, with comfort considered a direct indicator of quality of health care (NQMC 2002; Kolcaba 2013). Given that improving the physical, social, psychospiritual and/or environmental patient comfort is associated with quicker discharges, fewer readmissions, increased patient satisfaction and stronger cost-benefit ratios for the institution (Kolcaba, 2013; 2001), it is essential for nurses, clinicians, researchers and institutions to know how the different instruments assessing comfort perform. With this purpose, a psychometric review to assess the validity (e.g. construct validity or content validity) and the reliability (e.g. inter-rater reliability) of instruments assessing patient comfort was conducted. In addition, since the instruments need to have high utility if they are to be used in the real-world practice, the cost-efficiency (e.g. cost of obtaining a sample), acceptability (e.g. suitability from the patient perspective), and educational impact (e.g. utility of the collected data) of each instrument measuring comfort were also assessed, according to the utility matrix proposed by Beattie et al. (2015).

Background

Patient comfort is considered an individualized and holistic experience, a source of patient satisfaction and well-being. The concept of comfort is historically associated with nursing. Different theories and definitions of comfort have been developed, but most have been restricted to physical connotations, such as pain control. However, the Theory of Comfort, by Kolcaba (1992), associated the concept of comfort with strengthening, encouragement, aid and support, and provided a theoretical significance for comfort in nursing. According to this theory, comfort is defined as the immediate experience of being strengthened by the need for relief (the experience of a patient whose specific comfort need is addressed), ease (the state of calmness or contentment), and transcendence (the state in which the patient rises above pain or problems), met in physical, psychospiritual, sociocultural, and environmental contexts; much more so than the absence of pain or other physical discomforts. When the three kinds of comfort are combined with the four contexts of experience, a twelve-cell grid is created, which is useful for assessing the comfort needs of patients and families (e.g. privacy, pain control, information about clinical procedures, anxiety, or noisy environment), for planning interventions to address those needs, evaluating the effectiveness of those interventions to enhance the comfort, and for measuring the desired outcome in research and practice (Kolcaba, 2013). In this context, the General Comfort Questionnaire (GCQ) is probably the first instrument specifically developed to assess patient comfort as a holistic experience, registered as a multidisciplinary outcome indicator of quality in health care in the National Quality Measures Clearinghouse (Kolcaba 1992, 2013; NQMC 2002). Since then, a large range of instruments to assess patient comfort have been developed, adapted or validated, either within the Theory of Comfort or other theoretical frameworks.

To our knowledge, there are no psychometric reviews assessing the performance of comfort instruments to date. For this reason, this study aims to examine the psychometric properties

and utility of each questionnaire used to measure patient comfort in order to select the most valid, reliable and useful instrument for nurses, clinicians and researchers to use in present-day health care.

THE REVIEW

Aims

1. Identify health instruments measuring patient comfort as a holistic experience during hospitalization.
2. Systematically review the measurement properties and interpretability of each comfort instrument to measure patient comfort.
3. Examine the utility of each comfort instrument according to their cost efficiency, acceptability and educational impact in different health care settings.
4. Classify the different comfort instruments according to their measurement properties, and utility.

Design

According to the protocol of this review (Lorente, Vives & Losilla, 2017), we conducted a psychometric review, applying the COnsensus-based Standards for the selection of health status Measurement INstruments (COSMIN) to assess the methodological quality of studies on measurement properties (Mokkink et al., 2012; Terwee et al., 2012) and the Quality Criteria for Measurement Properties to assess the quality of instruments (Terwee et al. 2007). This review was registered in PROSPERO, CRD42016036290.

Search methods

We aimed to identify published instruments to assess patient comfort during hospitalization. We defined different combinations of keywords (using Mesh and other thesauruses, where

available), in relation to the concept (e.g., comfort, theory), setting (hospitalization or admission) and instruments (e.g., questionnaires, scales) (Supplementary file 1). The search was carried out between 1990 and 2015, the findings were restricted by language (English), and the following databases were included: Medical Literature Analysis and Retrieval System (MEDLINE), by ProQuest, Cumulative Index to Nursing and Allied Health Literature (CINAHL), by EBSCOhost, Psychological Information (PsycINFO), by APA PsycNET, Thesis & Dissertations, by ProQuest, and ISI Web of Knowledge (WoS, Web of Science CORE), by Thomson Reuters. To include grey literature, we also searched records in Google, and reviewed up to 400 links. In addition, search alerts in CINAHL & PsycINFO were set until December 2016.

Inclusion criteria

Time frame

From 1990**,** with the development of the first instrument to assess patient comfort framed within the Theory of Comfort, by Kolcaba in 1992, until 2015.

Study type

Studies developing or validating questionnaires and/or scales measuring holistic comfort during hospitalization. Protocols, guidelines, conceptual frameworks, narrative reviews, intervention studies, report opinions and qualitative studies, were excluded.

Age group

We included all age groups (new-borns, toddlers, children, teenagers, young adults, middle-aged adults and elderly people).

Context

We evaluated instruments used to measure comfort in different healthcare settings where the patient was hospitalized due to acute illnesses (e.g. pneumonia, urinary tract infection), chronic pathologies (e.g. psychiatric illnesses, dementia), surgical interventions or childbirth. Therefore, the following settings were included: general paediatric wards, general adult wards, delivery rooms and maternity wards, operating rooms, Intensive Care Units (ICU), Paediatric Intensive Care Units (PICU), Neonatal Intensive Care Units (NICU), postoperative areas, psychiatric wards and hospices.

Instruments used to assess patient comfort

We included instruments developed with scales and/or subscales with closed-ended items, as Likert, Visual Analogic Scales (VAS) and/or Numerical Rating Scales (NRS) specifically designed for measuring patient comfort as a holistic experience during hospitalization. Instruments measuring comfort during screening or diagnostic tests were excluded (e.g. colonoscopy, bronchoscopy, angiography, mammography, injections, biopsies, cystoscopy, cytology, fertility treatments), as well as those instruments assessing comfort exclusively by physiological parameters (i.e. heart rate and/or blood pressure) or those which measure the effect of a specific intervention (e.g. warm blanket vs. classical blanket or midazolam vs. fentanyl).

Search outcomes

The results of the search strategy were reported according to the PRISMA flow diagram (Figure 1). A total of 2995 references were identified through databases search, plus 20 references from Google. After removing duplicates, 2843 titles and abstracts were screened. A reviewer applied the inclusion criteria to all titles and abstracts. If no decision could be made based solely on title and abstract alone, the full paper was retrieved. The inclusion criteria were checked independently by two review authors and discrepancies were resolved

through discussion (with a third author where necessary). After the assessment of 103 full-texts for eligibility, 35 articles were selected for inclusion, and 14 additional articles were identified and retained by the references search and citation alerts. A total of 49 articles were included in the psychometric review.

----------Insert Figure 1 here or near here----------

Quality appraisal

Quality assessment of the studies

The methodological quality of included studies was assessed by using the COnsensus-based Standards for the selection of health status Measurement INstruments (COSMIN) (Mokkink et al., 2010). The COSMIN checklist consists of 12 criterion, referred to as "boxes", to evaluate whether general requirements of the study on measurement properties are met. We used eight boxes to assess the quality of the studies on internal consistency, reliability, measurement error, structural validity, content validity (including face validity), construct validity, cross-cultural validity, and responsiveness. The criterion validity was not rated since no gold standard for instruments to measure comfort exists (Mokkink et al., 2012; Terwee et al., 2012) (See supplementary File 2 for domains and definitions). Methodological quality of every measurement property was appraised using the four point score (poor, fair, good, excellent), according to the "worse score counts" algorithm (Terwee, Mokkink & Patrick, 2011). For instance, if item 3 in Reliability box was scored as fair (i.e. moderate sample), and other items of the box were scored as good or excellent, the overall methodological quality of Reliability box was fair. Therefore, if any item was scored as poor (i.e. small sample), the overall methodological quality of the box was scored as poor. Two reviewers independently assessed the methodological quality of studies and discrepancies were identified and resolved through discussion (with a third author where necessary).

Quality assessment of the instruments

The quality of the results for each study was assessed on the basis of Quality Criteria for Measurement Properties proposed by Terwee et al. (2007). These criteria consist of rating the content validity, internal consistency, construct validity, reproducibility, responsiveness, floor and ceiling effects and interpretability. Criterion validity was not assessed. The quality of the results was rated as "+" (positive), "-" (negative), "?" (doubtful), or "0" (no information) (See supplementary File 3 for domains, definitions and ratings). Two reviewers independently rated the quality of results, and discrepancies were identified and resolved through discussion (with a third author where necessary).

Data abstraction

Extracted information of each selected instrument included: author, year, country, outcome measures, purpose/use, number of items, response categories, type of patients; characteristics (theoretical framework, validity tests conducted and results, reliability tests conducted and results, responsiveness tests conducted and results, response rate, sample size, setting, respondents population and demographics, level of expertise required for scoring, time required to completion, mode of administration). Once the information was extracted, we categorised the instruments according to the measurement model (reflective and/or formative) taking into account the considerations detailed in Coltman et al. (2008) (See supplementary File 4), and we reported different aspects related to the utility (cost-efficiency, acceptability and educational impact) according to Van der Vleuten´s Utility criteria (Van der Vleuten, 1996, cited in Beattie *et al.* 2015) (See supplementary File 5). Some authors of eligible studies were contacted to provide missing or additional data when necessary.

Synthesis

The extracted information related to the instruments characteristics was reported in a table designed *ad ho*c. Results of the methodological quality assessment were synthetized in a table, according to the COSMIN domains (Mokkink et al. 2012) and the Quality Criteria of Measurement Properties proposed by Terwee et al. (2007). The extracted information related to the utility (cost-efficiency, acceptability and educational impact) was synthetized in a table designed according to Van der Vleuten´s Utility Index Matrix (Van der Vleuten, 1996, cited in Beattie et al. 2015).

RESULTS

Characteristics of instruments

Instruments assessing patient comfort were taken from 1991 to 2016, and were developed or adapted in the following countries: United States of America, United Kingdom, the Netherlands, Brazil, Spain, Thailand, China, Germany, Italy, Portugal, Norway, Sweden, France and Israel. Table 1 shows instruments assessing comfort, that are listed and described in the next sections: Comfort of paediatric patients, Comfort of adult patients, and Comfort in elderly patients with dementia (See supplementary File 6 for further details).

----------Insert Table 1 here or near here----------

Comfort of paediatric patients

The comfort of paediatric patients was generally assessed by instruments based on the observation of specific emotions and physiological parameters. The Comfort Scale (CS) was the most common scale, with 8 items and two dimensions that account for the 84% of the variance, originally developed by Ambuel et al. (1992) to assess infant distress in the Paediatric Intensive Care Unit (PICU). The CS was based upon the concept of psychological distress, defined as a multidimensional response to internal or external aversive stimuli, that may include fear, discomfort, anxiety and pain. The measurement model used to design this

instrument was reflective, with items concerned with facial expressions, muscle tension, heart rate and blood pressure.

Different authors validated the CS in PICU (Bear & Ward-Smith, 2006; Brunow de Carvalho et al. 1999; Courtman, Wardurgh, & Petros, 2003; Froom et al., 2008; Ista et al., 2005; Lamas et al., 2008; Nievas, Spentzas, & Bogue, 2014; Triltsch et al., 2005; Tschiedel et al. 2015). Some others authors also assessed the comfort of neonates by using the CS (Blauer, 1996; van Dijk et al., 2000; Wielenga et al., 2004; Franck et al., 2011; Cury, Martinez & Carlotti, 2013), Comfort Neo Scale (Monique van Dijk et al., 2009), Modified Comfort Scale (Lee & Young, 2005; Gjerstad et al., 2008), and Adapted Comfort Scale (Caljouw et al., 2007). We also identified the Comfort Behavioural Scale (CBS) as an adaptation of the CS, with 6 items and one dimension (Johansson & Kokinsky, 2009; de Jong et al., 2010, 2012; Amigoni et al., 2012; Bai et al., 2012; Valkenburg, 2012; Boerlage et al., 2012, 2014; da Costa Silva et al., 2013; Tristão et al., 2013; Andersen et al., 2015). Lastly, the Paediatric Perioperative Comfort Instrument (PPCI), 7 items (Moriber, 2009a), which assess the comfort of paediatric patients after surgical procedures was developed. This questionnaire was based upon the Theory of Comfort, by Kolcaba (1992), and the measurement model to design this instrument was reflective. Reflective items were related to facial expressions and verbal indicators.

Comfort of adult patients

The comfort of adult patients was evaluated in different health settings. The General Comfort Questionnaire (GCQ) (Kolcaba, 1992), with 48 items and twelve dimensions that accounted for a variance of 63.4%, was developed to assess the comfort of adult patients in medical and surgical wards, and was based upon the Theory of Comfort (Kolcaba, 1992). The measurement model to design this instrument was mixed, reflective and formative. Reflective items were related to fear, anxiety, or information provided, while formative items were

related to environmental sources of discomfort, such as light, furniture or noise. We found some adaptations of the GCQ, as the Childbirth Comfort Questionnaire (CCQ), 14 items to assess the comfort of the women during childbirth (Durnell, 2003), and Psychiatric In-Patients Comfort Scale (PICS), 38 items to assess the comfort of patients with psychiatric disorders (Alves-Apóstolo et al. 2007). The Psychosocial Comfort scale, 21 items, to assess the comfort of adults in medical and surgical wards (Yen, 1994). This questionnaire was also based on the Theory of Comfort, and the measurement model was reflective, with items concerned to anxiety, trust in the medical team or communication with professionals.

Different instruments were developed within other frameworks to assess the comfort of adult patients. The Pain Discomfort Scale (PDS), which consists of 10 items, was developed by Jensen, Karoly & Harris (1991) to assess the discomfort of patients with chronic pain. This questionnaire was based on the multidimensional nature of pain, and was designed as a reflective instrument, composed of items related to fear, helplessness or annoyance. The Incomforts des Patients de REAnimation (IPREA) includes 16 items to assess the comfort of the patients admitted to the Intensive Care Unit (ICU) by Kalfon *et al.* (2010). This tool was based on the conceptualization of stress and post-traumatic stress, and was designed as a formative questionnaire, with items concerned with environmental sources of discomfort, such as noise, light or furniture. The CS, was validated to assess the comfort of adults admitted to ICU (Ashkenazy & DeKeyser-Ganz, 2011), and the Psychiatric Discomfort Scale (PDS), with 23 items, was used to assess the comfort of the patients with psychiatric disorders (Betemps, 1999).

Comfort of elderly patients with dementia

The comfort of elderly patients with dementia was assessed by diverse instruments based upon the conceptualization of the discomfort as the observation of specific emotions and

body language, defined as negative and/or physical state in response to physical problems and/or environmental conditions. All of them were formative questionnaires. Formative items concerned with environmental sources of discomfort, such as noise, light or professional procedures, were included in this questionnaire. We identified the Discomfort in Alzheimer Type (DS-DAT), 9 items, (Dello Russo et al., 2008; Hurley, Volicer, Hanrahan, Houde, & Volicer, 1992); Discomfort Scale, 16 items (Morrison et al., 1998); Discomfort Behaviour Scale (DBS), 17 items (Stevenson et al., 2006); Comfort Assessment Dying with Dementia, 14 items (CAD EOLD) (Volicer, Hurley & Blasi, 2001; Kiely et al., 2006) and Source of Discomfort Scale (SODS), 20 items (Cohen-Mansfield et al. 2013). In addition, we found the Hospice Comfort Questionnaire (HCQ), 49 items (Novak et al. 2001; Tanatwanit, 2011) and the End of Life Comfort Planning Questionnaire, 28 items (Oliveira et al. 2016), as adaptations from the GCQ. The former instrument measurement model was mixed (formative and reflective), and the latter was reflective.

Methodological quality

Table 2 shows the results of the methodological quality of studies on measurement properties, COSMIN, and the quality assessment of instruments, Quality Criteria of Measurement Properties. Some specific considerations are synthetized.

----------Insert Table 2 here or near here----------

Comfort of paediatric patients

Content validity was only tested in two studies (Ambuel et al., 1992; Moriber, 2009), and the quality of methodology was considered as “excellent” according to the COSMIN criteria. However, the CS, by Ambuel et al. (1992), also obtained a “doubtful” score because the target population and the item selection process were not adequately reported. Reliability was tested in numerous studies, and most were rated as “fair” and “doubtful”, because authors

assessed the inter-rater and/or test-retest reliability of continuous scores by using correlations (Ambuel et al. 1992; Bear & Ward-Smith, 2006; Lee & Young, 2005) or t-tests (Boerlage et al. 2012). Instrument structure, according to their factor analysis, was rated as "good" or "excellent", although it was only tested by six authors (Ambuel et al., 1992; de Jong et al., 2012; Franck et al., 2011; Moriber, 2009; Valkenburg, 2012; van Dijk et al., 2000). Construct validity (convergent and/or discriminative validity) was generally rated as "fair", but "doubtful", when the hypothesis about expected correlations between scores of instruments was not formulated. Responsiveness was also rated as "fair" and "doubtful", in most cases when the hypothesis about expected differences, the Smallest Detectable Change (SDC) or the Minimal Important Change (MIC), were not stated. Lastly, interpretability obtained a "doubtful" score when the comfort scores of patients were not adequately described per groups and subgroups.

Comfort of adult patients

Overall methodological quality of instruments was rated as "good", and most studies obtained a "positive" score on Quality Criteria of Measurement Properties. Nevertheless, some studies obtained a "doubtful" score in construct validity and responsiveness because the hypothesis and SDC or MIC were not stated.

Comfort of elderly patients with dementia

Content validity was tested in most studies. The methodological quality of content validity was generally rated as "excellent", and the quality of results obtained a "positive" score on Quality Criteria of Measurement Properties. Reliability was rated as "fair" and "doubtful" because authors assessed the inter-rater and/or test-retest reliability of continuous scores by using correlations (Cohen-Mansfield et al., 2013; Hurley et al. 1992; Morrison et al., 1998). The structure of the instruments, according to their factor analysis, was rated as "good",

although it has only been tested by three authors (Oliveira et al., 2016; Stevenson et al., 2006; Volicer et al., 2001). Construct validity (convergent and/or discriminative validity) was rated as “fair” and “doubtful”, when the hypothesis on the expected correlations between scores of instruments was not stated. Cross-cultural validity in translated questionnaires was rated as “poor” in all cases, since the multi-group factor analysis to test the construct invariance was not reported.

Utility of instruments

The utility index facilitated the evaluation of the cost-efficiency, acceptability and educational impact of each instrument. Results were showed in Table 3, and some considerations are synthetized.

----------Insert Table 3 here or near here----------

Firstly, the cost-efficiency evaluated the sample size to achieve the required level of reliability according to the purpose of the instrument, administering the time, and the administrative costs of applying the questionnaire and the completion of a reliable sample (e.g. professional training or number of collaborators). Because authors did not report the sample size used to reach the required reliability, we took into account the participants of each study. When authors did not report administrative costs, we considered whether the tool was self-reported or administered by interviewers (some administrative costs), or administered by trained professionals (large administrative costs). The CS, by Ambuel et al. (1992), and associated questionnaires, reported a shorter administering time (2 minutes), but also greater administrative costs, because the questionnaire had to be applied by specifically trained professionals, either nurses or doctors, or both. Secondly, the acceptability evaluated the understanding of the patients and the assessments (missing items and response rate). However, we were unable to evaluate the assessments in most cases, as information related to

missing items and/or response rate was generally not reported, apart from Jensen, Karoly & Harris (1991), van Djik et al. (2000), Novak et al. (2001), Stevenson et al. (2006) and Boerlage et al. (2014). Lastly, the educational impact evaluated the purpose of the instrument, the scoring system and the feedback of the results. We considered that all studies reported evidence of achieved purpose, and the results were applicable in a practical context.

DISCUSSION

To our knowledge, this is the first psychometric review assessing the performance of instruments used to assess patient comfort during hospitalization. With this purpose, we analysed the measurement properties and the use of a wide variety of questionnaires assessing comfort in different settings and age groups. The CS (Ambuel et al. 1992) and associated questionnaires were the most popular instruments measuring comfort in critical paediatric units. The GCQ (Kolcaba, 1992) and associated questionnaires were focused on assessing comfort in diverse areas, as surgical, medical and psychiatric wards or childbirth. Different questionnaires were also developed to assess comfort of elderly patients with dementia (DS-DAT, DBS, CAD EOLD, SODS, Discomfort Scale, HCQ, End of Life).

The methodological quality assessment indicated that several instruments were not as rigorously developed and validated as COSMIN and Quality Criteria of Measurement Properties recommend (Terwee et al. 2007; Mokkink et al. 2012). Poor reporting of item selection process and/or incomplete descriptions of the sample characteristics may limit the content validity (Terwee et al. 2007). The confirmatory factor analysis (CFA) to explore the dimensionality of the questionnaire and to determine the reliability coefficients of internal consistency derived from the measurement model were not always conducted and authors frequently reported results of CFA and reliability analyses from previously published research instead of analysing both in their own samples. This practice may be problematic, as

a measurement model, given that factor loadings and reliability depend on the sample data, its size and missing items (Floyd & Widaman 1995; de Vet et al. 2005; Mokkink et al. 2012). Moreover, the reliability of comfort scores was sometimes analysed by correlations, so the systematic error is not taken into account and agreement is not really assessed. The evaluation of construct validity and sensitivity without testing specific hypothesis may lead to misleading conclusions, since authors may be tempted to offer alternative explanations for low correlations or little mean differences instead of concluding that the questionnaire maybe is not valid for the intended purpose (Terwee et al. 2007; Mokkink et al. 2012). Lastly, when an instrument measuring comfort was translated and cross-culturally adapted, the multi-group factor analysis to test the construct invariance was not performed, increasing the risk of biased results in the comparison of scores (Little & Slegers 2005; Mokkink et al. 2012).

As far as the assessment of the utility of instruments is concerned, the educational impact was excellent, but cost-efficiency and acceptability were difficult to appraise because costs and assessments (missing items and response rate) were poorly reported, as pointed out in Beattie et al. (2015). For this reason, we were unable to rigorously categorize the instruments according to their utility. However, to select the right instrument for an intended purpose and a clinical scenario not only the usability should be considered, but also the attributes of comfort to be measured. In this sense, the self-reported instruments framed within the Theory of Comfort, by Kolcaba (1992), which assessed the comfort of patients in four contexts, most closely reflect the real experience of patients during hospitalization. But some instruments assessing the comfort of the paediatric or elderly patients may lead to biased conclusions, because the family experience is not usually assessed (e.g. CS or SODS). Different attributes of comfort, such as communication with professionals, information about procedures, opinions about the treatment, privacy, etc. need to be assessed even when patients are unable

to refer to comfort needs. In these cases, interviewing the family may be crucial to really assess the comfort of paediatric patients or elderly patients with dementia.

Recommendations for further research

Assessing comfort patient is essential to increase patient satisfaction and institutions. Given that both patient comfort and satisfaction are considered quality indicators of the health care provided (WHO, 2000; NQMC 2002; Kolcaba 2013), instruments aiming to assess these patient experiences should be valid, reliable and useful (Terwee et al. 2007; Keszei et al. 2010). Therefore, recommendations for further research include the systematic use of methodological quality assessment checklists, as COSMIN (Mokkink et al. 2012) and/or Quality Criteria of Measurement Properties (Terwee et al. 2007), and better data reporting including subjects understanding, application costs and assessments.

Strengths and limitations

A strength of our study is the use of the COSMIN (Mokkink et al. 2012) to evaluate methodological quality and the measurement properties of included studies, and the Quality Criteria (Terwee et al. 2007) to assess the quality of instruments, along with the application of the Utility Index Matrix (Beattie et al. 2015) to assess their cost-efficiency, acceptability and educational impact and discuss their usefulness in the real cases. Concerning limitations, it should be noted that only English literature was included and, although most of the studies were peer-reviewed published papers, there were five unpublished thesis and dissertations. Lastly, the poor reporting of some studies made it difficult to categorize the instruments according to their utility, as we had initially aimed to do.

CONCLUSIONS

Measuring patient comfort is a good practice to improve the health care provided and to increase patient satisfaction in present-day clinical scenarios. With this purpose, a number of

instruments assessing comfort have been developed and validated across different settings and circumstances, most of which are valid and reliable. However, there is no instrument that can currently be wholly recommended among the questionnaires we have reviewed. The CS and the CBS were the most adequate questionnaires to assess the comfort of children in critical areas. However, the moderate methodological quality and the cost of training professionals should be taken into account when these instruments are applied. The GCQ, and their adaptations to assess both the comfort of adult patients and elderly patients with dementia, reported overall good methodological quality, and they are probably the best self-reported questionnaires, to assess the comfort of the patient in diverse scenarios, although more validations with different samples are needed. The Psychosocial Comfort scale and IPREA also reported good methodological quality, being the most recommendable instruments to assess the comfort of the patient with surgical procedures and medical diagnoses or admitted to critical areas, respectively. Nevertheless, the former showed poor reporting on costs and assessments, and the latter reported floor effects.

The findings of this psychometric review additionally provide a strategy to select and develop the most appropriate instrument to assess comfort patient in accordance to their purpose, psychometric properties, and utility. Firstly, the measurement model, which is reflective and/or formative, should be considered when a new instrument is developed, as the measurement model defines the construct of interest, either comfort or discomfort. In addition, the questionnaires devised from a mixed theoretical model, formative and reflective, allow the identification of different sources of comfort/discomfort as well as the diagnosis of the level of comfort/discomfort, respectively. So, these mixed questionnaires may be useful for nurses, clinicians and researchers to assess the comfort needs of patients and families, to plan interventions to address those needs (e.g. pain reduction, hand massage, coaching or guided imagery to reduce the anxiety), and to evaluate the effectiveness of those interventions

to enhance comfort and well-being. Secondly, questionnaires assessing comfort must be adequately validated and their reliability should be correctly established to ensure their methodological quality. In this sense, the correct validation of translated and cross-culturally adapted questionnaires is also required to ensure the invariance of the construct. Otherwise, the results may lead to biased conclusions. Lastly, instruments measuring patient comfort should report cost-efficiency, acceptability and educational impact, which are crucial in the daily practice, both for professionals and patients.

References

Alves-Apóstolo, J. L., Kolcaba, K., Cruz-Mendes, A., & Calvário-Antunes, M. T. (2007). Development and psychometric evaluation of the Psychiatric In-patients Comfort Scale (PICS). *Enfermeria Clinica*, *17*(1), 17–23. http://doi.org/10.1016/S1130-8621(07)71760-6

Ambuel, B., Hamlett, K. W., Marx, C. M., & Blumer, J. L. (1992). Assessing distress in pediatric intensive care environments: the COMFORT scale. *Journal of Pediatric Psychology*, *17*(1), 95–109. http://doi.org/10.1093/jpepsy/17.1.95

Amigoni, A., Mozzo, E., Brugnaro, L., Gentilomo, C., Stritoni, V., Michelin, E., & Pettenazzo, A. (2012). Assessing sedation in a pediatric intensive care unit using Comfort Behavioural Scale and Bispectral Index: these tools are different. *Minerva Anestesiologica*, *78*(3), 322–329.

Andersen, R. D., Bernklev, T., Langius-Eklof, A., Nakstad, B., & Jylli, L. (2015). The COMFORT behavioural scale provides a useful assessment of sedation, pain and distress in toddlers undergoing minor elective surgery. *Acta Paediatrica*, *104*(9), 904–909. http://doi.org/10.1111/apa.13045

Ashkenazy, S., & DeKeyser-Ganz, F. (2011). Assessment of the reliability and validity of the Comfort Scale for adult intensive care patients. *Heart & Lung*, *40*(3), e44-51. http://doi.org/10.1016/j.hrtlng.2009.12.011

Bai, J., Hsu, L., Tang, Y., & Van Dijk, M. (2012). Validation of the COMFORT Behavior Scale and the FLACC Scale for Pain Assessment in Chinese Children after Cardiac Surgery. *Pain Management Nursing*, *13*(1), 18–26. http://doi.org/10.1016/j.pmn.2010.07.002

Bear, L. A., & Ward-Smith, P. (2006). Interrater reliability of the COMFORT scale. *Pediatric Nursing*, *32*(5), 427.

Beattie, M., Lauder, W., Atherton, I., & Murphy, D. J. (2015). Instruments to measure patient experience of health care quality in hospitals: a systematic review. *Systematic Reviews*, *3*(1), 4. http://doi.org/10.1186/2046-4053-3-4

Betemps, E. J. (1999). A Self-Administered Instrument to Measure Psychiatric Discomfort of Persons With Serious Mental Illness. *Psychiatric Services*, *50*(1), 107–108. http://doi.org/10.1176/ps.50.1.107

Blauer, T. (1996). *A simultaneous comparison of three neonatal pain scales during common newborn intensive care unit procedures*. University of Utah.

Boerlage, A. A., Ista, E., de Jong, M., Tibboel, D., & van Dijk, M. (2012). The COMFORT behavior scale: is a shorter observation period feasible?. *Pediatric Critical Care Medicine*, *13*(2), e124–e125. http://doi.org/10.1097/PCC.0b013e3182192d92

Boerlage, A. A., Ista, E., Duivenvoorden, H. J., de Wildt, S. N., Tibboel, D., & van Dijk, M. (2014). The COMFORT behaviour scale detects clinically meaningful effects of analgesic and sedative treatment. *European Journal of Pain*, *19*(4), 473–479. http://doi.org/10.1002/ejp.569

Brunow de Carvalho, W., Lucas da Silva, P. S., Paulo, C. S., Fonseca, M. M., & Belli, L. A. (1999). Comparison between the Comfort and Hartwig sedation scales in pediatric patients undergoing mechanical lung ventilation. *São Paulo Medical Journale*, *117*(5), 192–196.

Bruyneel, L., Coeckelberghs, E., Buyse, G., Casteels, K., Lommers, B., Vandersmissen, J., … Vanhaecht, K. (2017). Validation of the Child HCAHPS survey to measure pediatric

inpatient experience of care in Flanders. *European Journal of Pediatrics*. http://doi.org/10.1007/s00431-017-2919-7

Caljouw, M. A. A., Kloos, M. A. C., Olivier, M. Y., Heemskerk, I. W., Pison, W. C. R., Stigter, G. D., & Verhoef, A.-M. J. H. (2007). Measurement of pain in premature infants with a gestational age between 28 to 37 weeks : Validation of the adapted COMFORT scale. *Journal of Neonatal Nursing*, *13*(1), 13–18. http://doi.org/10.1016/j.jnn.2006.11.007

Cohen-Mansfield, J., Thein, K., Marx, M., Dakheel-Ali, M., & Jensen, B. (2013). Sources of Discomfort in Persons With Dementia. *JAMA Internal Medicine*, *173*(14), 1378–1379. http://doi.org/10.1001/jamainternmed.2013.6480.Author

Coltman, T., Devinney, T. M., Midgley, D. F., & Venaik, S. (2008). Formative versus reflective measurement models: Two applications of formative measurement. *Journal of Business Research*, *61*(12), 1250–1262. http://doi.org/10.1016/j.jbusres.2008.01.013

Courtman, S. P., Wardurgh, A., & Petros, A. J. (2003). Comparison of the bispectral index monitor with the Comfort score in assessing level of sedation of critically ill children. *Intensive Care Medicine*, *29*(12), 2239–2246.

Cury, M. R. J., Martinez, F. E., & Carlotti, A. P. C. P. (2013). Pain assessment in neonates and infants in the post-operative period following cardiac surgery. *Postgraduate Medical Journal*, *89*(1048), 63–7. http://doi.org/10.1136/postgradmedj-2012-130808

da Costa Silva, C., Osório Alves, M. M., dos Santos El Halal, M. G., dos Santos Pinheiro, S., & Carvalho, P. R. A. (2013). A comparison of gradual sedation levels using the Comfort-B scale and bispectral index in children on mechanical ventilation in the pediatric intensive care unit. *Revista Brasileira de Terapia Intensiva*, *25*(4), 306–311. http://doi.org/10.5935/0103-507X.20130052

de Jong, A., Baartmans, M., Tuinebreijer, W., & van Loey, N. E. E. (2010). Reliability , validity and clinical utility of three types of pain behavioural observation scales for young children with burns aged 0 – 5 years. *Pain*, *150*(3), 561–7. http://doi.org/10.1016/j.pain.2010.06.016

de Jong, A., Tuinebreijer, W., Bremer, M., van Komen, R., Middelkoop, E., & van Loey, N. E. E. (2012). Construct validity of two pain behaviour observation measurement instruments for young children with burns by Rasch analysis. *Pain*, *153*(11), 2260–6. http://doi.org/10.1016/j.pain.2012.07.021

de Vet, H. C. W., Adèr, H. J., Terwee, C. B., & Pouwer, F. (2005). Are factor analytical techniques used appropriately in the validation of health status questionnaires? A systematic review on the quality of factor analysis of the SF-36. *Quality of Life Research*, *14*(5), 1203–1218. http://doi.org/10.1007/s11136-004-5742-3

Dello Russo, C., Di Giulio, P., Brunelli, C., Dimonte, V., Villani, D., Renga, G., & Toscani, F. (2008). Validation of the Italian version of the Discomfort Scale - Dementia of Alzheimer Type. *Journal of Advanced Nursing*, *64*(3), 298–304. http://doi.org/10.1111/j.1365-2648.2008.04810.x

Durnell, K. (2003). *Exploring the Presence of Comfort within the Context of Childbirth. Thesis*. University of Michigan.

Ellis-Smith, C., Evans, C., Bone, A., Henson, L., Dzingina, M., Kane, P., & Higginson, I. (2016). Measures to assess commonly experienced symptoms for people with dementia in long-term care settings: A systematic review. *BMC Medicine*, *14*, 38–49. http://doi.org/10.1186/s12916-016-0582-x

Floyd, F. J., & Widaman, K. F. (1995). Factor Analysis in the Development and Refinement of Clinical Assessment Instruments. *Psychological Assessment*, *7*(3), 286–299.

http://doi.org/10.1037/1040-3590.7.3.286

Franck, L. S., Ridout, D., Howard, R., Peters, J., & Honour, J. W. (2011). A comparison of pain measures in newborn infants after cardiac surgery. *Pain*, *152*(8), 1758–1765. http://doi.org/10.1016/j.pain.2011.03.017

Froom, S. R., Malan, C. A., Mecklenburgh, J. S., Price, M., Chawathe, M. S., Hall, J. E., & Goodwin, N. (2008). Bispectral Index asymmetry and COMFORT score in paediatric intensive care patients. *British Journal of Anaesthesia*, *100*(5), 690–696. http://doi.org/http://dx.doi.org/10.1093/bja/aen035

Gjerstad, A. C., Wagner, K., Henrichsen, T., & Storm, H. (2008). Skin conductance versus the modified COMFORT sedation score as a measure of discomfort in artificially ventilated children. *Pediatrics*, *122*(4), e848-53. http://doi.org/10.1542/peds.2007-2545

Hurley, A. C., Volicer, B. J., Hanrahan, P. A., Houde, S., & Volicer, L. (1992). Assessment of discomfort in advanced Alzheimer patients. *Research in Nursing & Health*, *15*(5), 369–77. http://doi.org/10.1002/nur.4770150506

Ista, E., M, van D., Tibboel, D., M, de H., Aneja, R., van Dijk, M., … Aneja, R. (2005). Assessment of sedation levels in pediatric intensive care patients can be improved by using the COMFORT "behavior" scale. *Pediatric Critical Care Medicine*, *6*(1), 58–92 8p. http://doi.org/10.1097/01.PCC.0000149318.40279.1A

Jensen, M. P., Karoly, P., & Harris, P. (1991). Assessing the affective component of chronic pain: Development of the pain discomfort scale. *Journal of Psychosomatic Research*, *35*(2–3), 149–154. http://doi.org/10.1016/0022-3999(91)90069-Z

Johansson, M., & Kokinsky, E. (2009). The COMFORT behavioural scale and the modified FLACC scale in paediatric intensive care. *Nursing in Critical Care*, *14*(3), 122–130.

http://doi.org/10.1111/j.1478-5153.2009.00323.x

Keszei, A., Novak, M., & Streiner, D. L. (2010). Introduction to health measurement scales. *Journal of Psychosomatic Research*, *68*(4), 319–323. http://doi.org/10.1016/j.jpsychores.2010.01.006

Kiely, D. K., Volicer, L., Teno, J., Jones, R. N., Prigerson, H. G., & Mitchell, S. L. (2006). The Validity and Reliability of Scales for the Evaluation of End-of- Life Care in Advanced Dementia. *Alzheimer Disease and Associated Disorders*, *20*(3), 176–181.

Kolcaba, K. (1992). Holistic comfort: operationalizing the construct as a nurse-sensitive outcome. *Advances in Nursing Science*, *15*(1), 1–10.

Kolcaba, K. (2001). Evolution of the mid range theory of comfort for outcomes research. *Nursing Outlook*, *49*(2), 86–92.

Kolcaba, K. (2013). Comfort. In Lippincott Williams & Wilkins (Ed.), *Middle Range Theories. Application to nursing research.* (Third edit, p. 372). Philadelphia.

Lamas, A., López-Herce, J., Sancho, L., Mencía, S., Carrillo, A., Santiago, M. J., & Martínez, V. (2008). Assessing sedation in critically ill children by bispectral index, auditory-evoked potentials and clinical scales. *Intensive Care Medicine*, *34*(11), 2092–2099. http://doi.org/10.1007/s00134-008-1198-1

Lee, W. K., & Young, B. W. Y. (2005). Measuring the sedation level of mechanically ventilated infants by a modified COMFORT scale. *Hong Kong Journal of Paediatrics*, *10*(3), 189–195.

Little, T. D., & Slegers, D. W. (2005). Factor Analysis: Multiple Groups. *Behavioral Science*, *2*, 617–623. http://doi.org/10.1002/0470013192.bsa221

Lorente S. , Vives J. & Losilla J. M. (2017) Instruments to assess the patient comfort during hospitalization: a psychometric review protocol. *Journal of Advanced Nursing, 73* (3), 735-741. doi: 10.1111/jan.13180

Manary, M. P., Boulding, W., Staelin, R., & Glickman, S. W. (2013). The Patient Experience and Health Outcomes. *New England Journal of Medicine, 368*(3), 201–203. http://doi.org/10.1056/NEJMp1211775

Mokkink, L., Terwee, C., Patrick, D., Alonso, J., Stratford, P., Knol, D. L., … de Vet, H. C. W. (2010). The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: An international Delphi study. *Quality of Life Research, 19*(4), 539–549. http://doi.org/10.1007/s11136-010-9606-8

Mokkink, L., Terwee, C., Patrick, D., Alonso, J., Stratford, P. W., Knol, D. L., … de Vet, H. C. W. (2012). *The COSMIN checklist manual*. Retrieved from http://www.cosmin.nl/images/upload/files/COSMIN checklist manual v9.pdf

Moriber, N. (2009). *Evaluating the Reliability and Validity of the Pediatric Perioperative Comfort Instrument: A Pschometric Study. Thesis*. Rush University.

Morrison, R. S., Ahronheim, J. C., Morrison, G. R., Darling, E., Baskin, S. A., Morris, J., … Meier, D. E. (1998). Pain and discomfort associated with common hospital procedures and experiences. *Journal of Pain and Symptom Management*. Netherlands: Elsevier Science. http://doi.org/10.1016/S0885-3924(97)00261-3

Nievas, I. F. F., Spentzas, T., & Bogue, C. W. (2014). SNAP II index: an alternative to the COMFORT scale in assessing the level of sedation in mechanically ventilated pediatric patients. *Journal of Intensive Care Medicine, 29*(4), 225–228. http://doi.org/http://dx.doi.org/10.1177/0885066613475422

Novak, B., Kolcaba, K., Steiner, R., & Dowd, T. (2001). Measuring comfort in caregivers and patients during late end-of-life care. *American Journal of Hospice & Palliative Care*, *18*(3), 170–80 3p 1p. http://doi.org/10.1177/104990910101800308

Oliveira, S. M., Caldeira, S. M. A., Amado, J. C., & Kolcaba, K. (2016). Cultural adaptation and validation of the Portuguese End of Life Spiritual Comfort Questionnaire in Palliative Care patients. *Porto Biomedical Journal*, *In press*, 6–11. http://doi.org/10.1016/j.pbj.2016.08.003

Pettersen, K. I., Veenstra, M., Guldvog, B., & Kolstad, A. (2004). The Patient Experiences Questionnaire: Development, validity and reliability. *International Journal for Quality in Health Care*, *16*(6), 453–463.

Stevenson, K. M., Brown, R. L., Dahl, J. L., Ward, S. E., & Skemp, M. (2006). The Discomfort Behaviour Scale: A Measure of Discomfort in the Cognitively Impaired Based on the Minimum Data Set 2.0. *Research in Nursing & Health*, *29*, 576–587. http://doi.org/10.1002/nur.20168

Tanatwanit, Y. (2011). *Comfort as experienced by Thai older patients with advanced cancer. Thesis*.

Terwee, C., Bot, S., de Boer, M., van der Windt, D. A. W. M., Knol, D. L., Dekker, J., … de Vet, H. C. W. (2007). Quality criteria were proposed for measurement properties of health status questionnaires. *Journal of Clinical Epidemiology*, *60*(1), 34–42. http://doi.org/10.1016/j.jclinepi.2006.03.012

Terwee, C., Mokkink, L., Knol, D., Ostelo, R. W. J. G., Bouter, L. M., & de Vet, H. C. W. (2012). Rating the methodological quality in systematic reviews of studies on measurement properties: A scoring system for the COSMIN checklist. *Quality of Life Research*, *21*(4), 651–657. http://doi.org/10.1007/s11136-011-9960-1

Terwee, C., Mokkink, L., & Patrick, D. (2011). COSMIN checklist with 4-point scale. Retrieved from http://www.cosmin.nl

Triltsch, A. E., Nestmann, G., Orawa, H., Moshirzadeh, M., Sander, M., Grosse, J., … Spies, C. D. (2005). Bispectral index versus COMFORT score to determine the level of sedation in paediatric intensive care unit patients: a prospective study. *Critical Care (London, England)*, *9*(1), R9-17. http://doi.org/10.1186/cc2977

Tristão, R. M., Garcia, N. V. M., de Jesus, J. A. L., & Tomaz, C. (2013). COMFORT behaviour scale and skin conductance activity: What are they really measuring? *Acta Paediatrica*, *102*(9), 402–406. http://doi.org/10.1111/apa.12325

Tschiedel, E., Müller, O., Schara, U., Felderhoff-Müser, U., & Dohna-Schwake, C. (2015). Sedation monitoring during open muscle biopsy in children by Comfort Score and Bispectral Index - A prospective analysis. *Paediatric Anaesthesia*, *25*(3), 265–271. http://doi.org/10.1111/pan.12547

Valkenburg, A. J. (2012). *The COMFORT-behavior scale is useful to assess pain and distress in 0-to 3-year-old children with Down syndrome. Thesis*.

van Dijk, M., de Boer, J. B., Koot, H. M., Tibboel, D., Passchier, J., & Duivenvoorden, H. J. (2000). The reliability and validity of the COMFORT scale as a postoperative pain instrument in 0 to 3-year-old infants. *Pain*, *84*(2–3), 367–377. http://doi.org/10.1016/S0304-3959(99)00239-0

van Dijk, M., Roofthooft, D. W., Anand, K. J., Guldemond, F., de Graaf, J., Simons, S., … Tibboel, D. (2009). Taking up the challenge of measuring prolonged pain in (premature) neonates. *Clinical Journal of Pain*, *25*(7), 607–616. http://doi.org/10.1097/AJP.0b013e3181a5b52a

Volicer, L., Hurley, a C., & Blasi, Z. V. (2001). Scales for evaluation of End-of-Life Care in Dementia. *Alzheimer Disease and Associated Disorders*, *15*(4), 194–200. http://doi.org/10.1097/00002093-200110000-00005

Wielenga, J. M., de Vos, R., de Leeuw, R., & de Haan, R. J. (2004). COMFORT Scale: a reliable and valid method to measure the amount of stress of ventilated preterm infants. *Neonatal Network*, *23*(2), 39–44. http://doi.org/10.1891/0730-0832.23.2.39

World Health Organization. (2000). Strategies for assisting health workers to modify and improve skills: developing quality health care-a process of change. *Issues in Health Services*.

World Health Organization. (2004). Standards for Health Promotion in Hospitals.

World Health Organization. (2013). Health 2020: A European Policy Framework and Strategy for the 21st Century.

Figure 1. PRISMA Flow Diagram

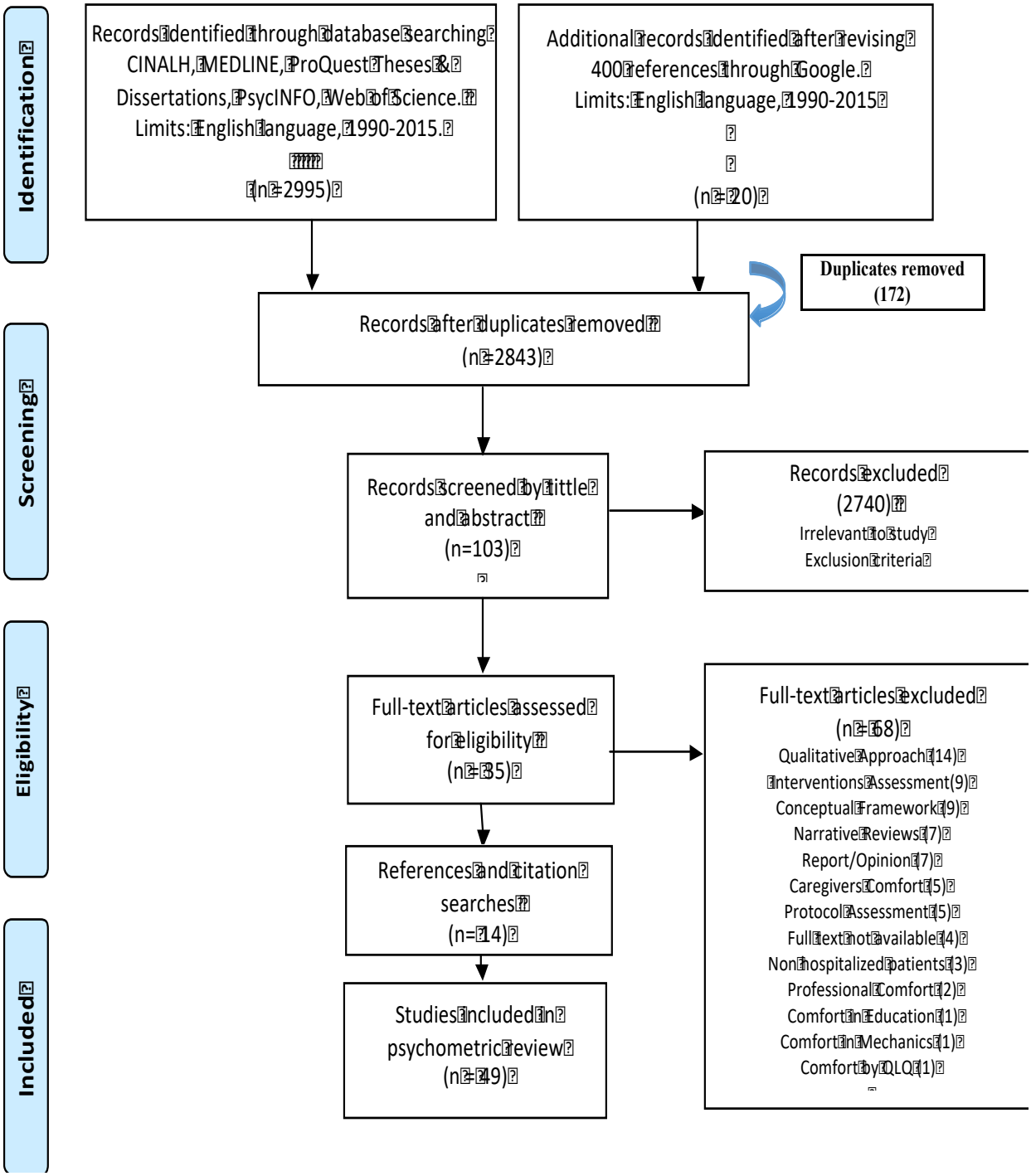

Table 1. Studies overview

Authors	Year	Questionnaire
Comfort of paediatric patients		
Ambuel et al.	1992	Comfort Scale (CS)
Blauer	1996	Comfort Scale (CS)
Bruno de Carvalho et al.	1999	Comfort Scale (CS)
Van Djik et al.	2000	Comfort Scale (CS)
Courtman et al.	2003	Comfort Scale (CS)
Wielenga et al.	2004	Comfort Scale (CS)
Ista et al.	2005	Comfort Scale (CS)
Triltsch et al.	2005	Comfort Scale (CS)
Bear & Ward-Smith	2006	Comfort Scale (CS)
Froom et al.	2008	Comfort Scale (CS)
Lamas et al.	2008	Comfort Scale (CS)
Franck et al.	2011	Comfort Scale (CS)
Cury et al.	2013	Comfort Scale (CS)
Nievas et al.	2014	Comfort Scale (CS)
Tschiedel et al.	2015	Comfort Scale (CS)
Lee & Young	2005	Modified Comfort Scale
Caljouw et al.	2007	Adapted Comfort Scale
Gjerstad et al.	2008	Modified Comfort Scale
Van Djik et al.	2009	Confort Scale Neo
Johansson & Kokinsky	2009	Confort Behavioural Scale (CBS)
de Jong et al.	2010	Confort Behavioural Scale(CBS)
Amigoni et al.	2012	Confort Behavioural Scale (CBS)
Bai et al.	2012	Confort Behavioural Scale (CBS)
Boerlage et al.	2012	Confort Behavioural Scale (CBS)
de Jong et al.	2012	Confort Behavioural Scale (CBS)
Valkenburg	2012	Confort Behavioural Scale (CBS)
da Costa Silva et al.	2013	Confort Behavioural Scale (CBS)
Tristão et al.	2013	Confort Behavioural Scale (CBS)
Boerlage et al.	2014	Confort Behavioural Scale (CBS)
Andersen et al.	2015	Confort Behavioural Scale (CBS)
Moriber	2009	Pediatric PerioperativeComfort Instrument (PPCI)
Comfort of adult patients		
Jensen et al.	1991	Pain Discomfort Scale (PDS)
Kolcaba	1992	General Comfort Questionnaire (GCQ)
Yen	1994	Psychosocial Comfort
Betemps	1999	Psychiatric Discomfort Scale (PDS)
Durnell et al.	2003	Childbirth Comfort Questionnaire (CCQ)
Alves-Apóstolo et al.	2007	Psychiatric In-Patients Comfort Scale (PICS)
Kalfon et al.	2010	Pediatric Perioperative Comfort Instrument (IPREA)
Ashkenazy& DeKeyser-Ganz	2011	Comfort Scale (CS)
Comfort of elderly patients with dementia		
Hurley et al.	1992	Discomfort Scale-Dementia of Alzheimer Type (DS-DAT)
Morrison et al.	1998	Discomfort Scale (DS)
Novak et al.	2001	Hospice Comfort Questionnaire (HCQ)
Volicer et al.	2001	Comfort Assessment Dying with Dementia (CAD-EOLD)
Kiely et al.	2006	Comfort Assessment Dying with Dementia (CAD-EOLD)
Stevenson et al.	2006	Discomfort Behaviour Scale (DBS)
Dello Russo et al.	2008	Discomfort Scale-Dementia of Alzheimer Type in Italian (DS-DAT)
Tanatwanit et al.	2011	Hospice Comfort Questionnaire in Thai (HQC)
Cohen- Mansfield et al.	2013	Source of Discomfort Scale (SODS)
Oliveira et al.	2016	End of Life Comfort Planning Questionnaire

Note: For further details and instruments characteristics, see the Supplementary File 6.

Table 2. Methodological quality

Author	Year	Questionn.	Internal consistency		Reproducibility			Content validity		Structural validity	Construct validity		Cross-cultural	Resp.		Floor-Ceiling	Interp.
					Ag.	Reliability											
Comfort of paediatric patients																	
Ambuel	1992	CS	**	+	0	**	?	****	?	***	**	?			0	+	?
Blauer	1996	CS		0	0		0		0			0		**	?	0	+
de Carvalho	1999	CS		0	0		0		0		*	?			0	0	+
Van Djik	2000	CS	****	+	0	***	+		0	****	**	?			0	0	+
Courtman	2003	CS		0	0		0		0		**	?			0	0	+
Wielenga	2004	CS		0	0	**	+		0		**	?			0	0	0
Ista	2005	CS	***	+	0	***	+		0		**	?			0	0	+
Triltsch	2005	CS		0	0		0		0		**	?			0	0	+
Bear	2006	CS	***	+	0	**	?		0			0			0	0	?
Froom	2008	CS		0	0		0		0		*	?			0	0	+
Lamas	2008	CS		0	0		0		0		***	+			0	0	0
Franck	2011	CS		0	0	***	+		0	***		0		**	?	0	+
Cury	2013	CS		0	0		0		0		*	?		*	?	0	+
Nievas	2014	CS		0	0		0		0		*	?			0	0	?
Tschiedel	2015	CS		0	0		0		0		*	?			0	0	+
Lee	2005	Modified CS	*	+	0	*	?		0		**	0			0	0	?
Caljouw	2007	Adapted CS	***	+	0	***	+		0		***	+		***	+	0	?
Gjerstad	2008	Modified CS		0	0		0		0		*	?			0	0	+
Van Djik	2009	CNeo	***	+	0	***	+		0		**	?		**	?	0	+
Johansson	2009	CBS		0	0	**	+		0		**	?			0	0	+
de Jong	2010	CBS	***	+	0	**	+		0		**	?		**	+	0	?
Amigoni	2012	CBS		0	0	**	-		0		**	?		**	?	0	+
Bai	2012	CBS		0	0		0		0		**	?			0	0	?
Boerlage	2012	CBS		0	0	**	?		0			0			0	0	0
de Jong	2012	CBS		0	0		0		0		**	?			0	0	?
Valkenburg	2012	CBS	***	+	0		0		0	***	**	?			0	+	?
da Costa	2013	CBS		0	0	*	+		0		*	?			0	0	?
Tristão	2013	CBS		0	0		0		0		**	+		**	+	0	?
Boerlage	2014	CBS		0	0	**	+		0			0		****	+	0	+
Andersen	2015	CBS		0	0	**	+		0		**	+			0	-	+
Moriber	2009	PPCI	***	+	0		0	****	+		***	+			0	0	+
Comfort of adults patients																	
Jensen	1991	PDS	***	+	0	**	?	****	+	***	**	?			0	+	?
Kolcaba	1992	GCQ	***	+	0		0	****	?	***		0			0	0	0
Yen	1994	Psychosocial Comfort	***	+	0		0	****	+	***		0			0	0	+
Betemps	1999	PDS	***	+	0		0	***	?	***	**	?		**	?	0	0
Durnell	2003	CCQ	***	+	0		0	****	+	***	***	+			?	+	+
Alves-Apóstolo	2007	PICS	***	+	0		0	****	+	***	**	?			0	0	0
Kalfon	2010	IPREA	***	+	0	***	+	****	+	***		0			0	-	+
Ashkenazy	2011	CS	***	-	0	***	-		0		**	?		*	?	0	+
Comfort of elderly patients with dementia																	
Hurley	1992	DS-DAT	**	?	0	**	?	****	+			0		**	?	0	?
Morrison	1998	Discomfort Scale		0	0	**	?	****	+		**	?			0	0	+
Novak	2001	HCQ	**	+	0	***	+	****	+		**	?			0	0	+
Volicer	2001	CAD-EOLD	***	+	0		0	****	+	***	**	?			0	0	?
Kiely	2006	CAD-EOLD	***	+	0		0		0		**	?			0	0	?
Stevenson	2006	DBS	***	+	0		0	****	+	***		0			0	0	+
Dello Russo	2008	DS-DAT	***	+	0	***	+	****	+			0	*		0	0	+
Tanatwanit	2011	HQC	***	+	0		0	****	+		**	?	*		0	+	+
Cohen-Mansfield	2013	SODS		0	0	**	?	****	+		***	+			0	0	?
Oliveira	2016	End of Life	***	+	0		0	****	+	**	**	?	*		0	0	0

Note: COSMIN Ratings: **** Excellent; ***Good; ** Fair; * Poor.
Quality Criteria for Measureme
nt Properties Ratings: **+** A positive rating indicates strong properties according to quality criteria using design and method; **?** An intermediate rating indicates some but not all aspects of psychometric are positive, or doubtful design or method: **-** A negative rating indicates psychometric properties do not meet criteria despite adequate design and method; **0** No information provided.
Abbreviations: Questionn.= Questionnaire; Ag.= Agreement; Resp. =Responsiveness; Interp.= Interpretability.

Table 3. Utility matrix

Author	Year	Questionn.	Cost efficiency				Acceptability			Educational Impact		
			CostEf1	CostEf2	CostEf3	CostEf4	Accept1	Accept2	Accept3	EdImp1	EdImp2	EdImp3
Comfort of paediatric patients												
Ambuel	1992	CS	***	****	**	***	na	nr	****	****	****	****
Blauer	1996	CS	***	****	**	***	na	nr	****	****	****	****
de Carvalho	1999	CS	****	****	**	****	na	nr	****	****	****	***
Van Djik	2000	CS	*	****	**	*	na	****	****	****	****	****
Courtman	2003	CS	***	****	**	***	na	nr	****	****	****	****
Wielenga	2004	CS	****	****	**	****	na	nr	****	****	****	****
Ista	2005	CS	**	****	**	**	na	nr	****	****	****	****
Triltsch	2005	CS	***	****	**	***	na	nr	****	****	****	****
Bear	2006	CS	**	****	**	**	na	nr	****	****	****	****
Froom	2008	CS	****	****	**	****	na	nr	****	****	****	****
Lamas	2008	CS	**	****	**	**	na	nr	****	****	****	****
Franck	2011	CS	**	****	**	**	na	nr	****	****	****	***
Cury	2013	CS	****	****	**	****	na	nr	****	****	****	****
Nievas	2014	CS	****	****	**	****	na	nr	****	****	****	***
Tschiedel	2015	CS	***	****	**	***	na	nr	****	****	****	****
Lee	2005	Modified CS	****	****	**	****	na	nr	****	****	****	****
Caljouw	2007	Adapted CS	**	****	**	**	na	nr	****	****	****	****
Gjerstad	2008	Modified CS	****	****	**	****	na	nr	****	****	****	****
Van Djik	2009	CNeo	*	****	**	*	na	nr	****	****	****	****
Johansson	2009	CBS	***	****	**	***	na	nr	****	****	****	***
de Jong	2010	CBS	*	****	**	*	na	nr	****	****	****	****
Amigoni	2012	CBS	***	****	**	***	na	nr	****	****	****	****
Bai	2012	CBS	*	****	**	*	na	nr	****	****	****	***
Boerlage	2012	CBS	**	****	**	**	na	nr	****	****	****	***
de Jong	2012	CBS	*	****	**	*	na	nr	****	****	****	****
Valkenburg	2012	CBS	**	****	**	**	na	nr	****	****	****	****
da Costa	2013	CBS	***	****	**	***	na	nr	****	****	****	****
Tristão	2013	CBS	***	****	**	***	na	nr	****	****	****	****
Boerlage	2014	CBS	*	****	**	*	na	***	****	****	****	****
Andersen	2015	CBS	***	****	**	***	na	nr	****	****	****	****
Moriber	2009	PPCI	*	nr	***	*	na	nr	****	****	****	****
Comfort of adults patients												
Jensen	1991	PDS	**	****	***	**	**	**	****	****	****	****
Kolcaba	1992	GCQ	*	nr	***	*	*	nr	****	****	****	****
Yen	1994	Psychosocial Comfort	*	nr	***	*	*	nr	****	****	****	****
Betemps	1999	PDS	*	nr	***	*	**	nr	****	****	****	***
Durnell	2003	CCQ	**	****	***	**	**	nr	****	****	****	****
Alves Apóstolo	2007	PICS	*	nr	***	*	****	nr	****	****	****	****
Ashkenazy	2011	CS	**	****	**	**	na	nr	****	****	****	***
Comfort of elderly patients with dementia												
Hurley	1992	DS-DAT	**	****	**	**	na	nr	****	****	****	****
Morrison	1998	Discomfort Scale	*	nr	***	*	**	nr	****	****	****	****
Novak	2001	HCQ	*	****	***	*	**	***	****	****	****	****
Volicer	2001	CAD-EOLD	**	nr	***	**	**	nr	****	****	****	****
Kiely	2006	CAD-EOLD	*	nr	***	*	**	nr	****	****	****	****
Stevenson	2006	DBS	*	nr	**	*	na	****	****	****	****	****
Dello Russo	2008	DS-DAT	**	****	**	**	na	nr	****	****	****	***
Tanatwanit	2011	HQC	*	**	***	*	**	nr	****	****	****	****
Cohen-Mansfield	2013	SODS	*	nr	**	*	na	nr	****	****	****	****
Oliveira	2016	End of Life	*	****	***	*	****	nr	****	****	****	****

Note: Ratings: * poor, **fair, ***good, ****excellent, **nr** not reported, **na** no applicable. Utility aspects: CostEf1=number of observations needed to reach the required level of reliability; CostEf2= time to complete the questionnaire; CostEf3=administrative costs; CostEf4=cost to complete a reliable sample;Accept1=subjects understanding of the instrument; Accept2=assessments not completed; Accept3=instrument tested in an appropriate context; EdImp1=evidence of instrument purpose is achieved;EdImp2=scoring system stated and/or available in an easy format; EdImp3=feedback from the results can be used for action.

Abbreviations: Questionn.= Questionnaire; CostEf= Cost Efficiency; Accept=Acceptability; EdImp= Educational Impact.