



NOVA

University of Newcastle Research Online

nova.newcastle.edu.au

Forshaw, Kristy L.; Carey, Mariko L.; Hall, Alix E.; Boyes, Allison W.; Sanson-Fisher, Rob. 'Preparing patients for medical interventions: a systematic review of the psychometric qualities of published instruments'. Published in *Patient Education and Counseling* Vol. 99, Issue 6, p. 960-973 (2016)

Available from: <http://dx.doi.org/10.1016/j.pec.2015.12.008>

© 2016. This manuscript version is made available under the CC-BY-NC-ND 4.0 license
<http://creativecommons.org/licenses/by-nc-nd/4.0/>

Accessed from: <http://hdl.handle.net/1959.13/1321524>

Preparing patients for medical interventions: a systematic review of the psychometric qualities of published instruments

Authors:

Kristy L Forshaw^{a*}

Mariko L Carey^a

Alix E Hall^a

Allison W Boyes^a

Rob Sanson-Fisher^a

Affiliations:

^aPriority Research Centre for Health Behaviour, Health Behaviour Research Group, University of Newcastle & Hunter Medical Research Institute. Callaghan, Australia.

Email addresses:

Ms Kristy Forshaw: kristy.forshaw@newcastle.edu.au

Dr Mariko Carey: mariko.carey@newcastle.edu.au

Dr Alix Hall: alix.hall@newcastle.edu.au

Dr Allison Boyes: allison.boyes@newcastle.edu.au

L/Prof Rob Sanson-Fisher: rob.sanson-fisher@newcastle.edu.au

*** Corresponding author at:**

Kristy Forshaw

Health Behaviour Research Group, School of Medicine & Public Health, University of Newcastle, HMRI Level 4, CALLAGHAN NSW 2308 AUSTRALIA

Telephone: +61 2 4042 0682 Fax: +61 2 4042 0044

ABSTRACT

Objective: Preparing patients for medical interventions improves patient outcomes and is an ethical and legal imperative. This review examines the characteristics and psychometric properties of published instruments which assess patients' preparation for medical interventions.

Methods: Medline, CINAHL, EMBASE and PsycINFO electronic databases were searched from the date of their inception to November 2015. Data-based publications describing the development or validation of a self-report instrument designed to assess the quality of adult patients' perceived preparation for a medical intervention were included.

Results: Nine publications described the development or validation of seven instruments which met inclusion criteria. The psychometric qualities of the instruments varied. None met all of the accepted criteria for psychometric rigour. Although the Satisfaction with Cancer Information Profile met the highest number (n=5) of the defined psychometric properties, the study sample size was less than 100. Overall, content validity of the included instruments was the most frequently assessed criteria.

Conclusion: Few instruments have been specifically developed to assess patients' self-reported preparation for medical interventions. Of the available instruments, none demonstrated adequate rigour across essential psychometric properties.

Practice Implications: The need to develop instruments examining patient preparation for medical interventions is apparent given the limitations of the instruments reviewed.

Keywords: patient preparation, medical intervention, psychometric evaluation, systematic review

1. INTRODUCTION

Medical interventions are common and associated with a high degree of burden

'Medical intervention' is a broad term used to refer to all forms of diagnostic procedures, tests, and treatment [1]. Most people, over their lifetime, will experience a number of medical interventions. In Australia alone, about 11.9 million procedures were reported for admitted hospital patients in 2012-13 [2]. Whilst in the USA, 51.4 million procedures were reported for inpatients from non-Federal short-stay hospitals in 2010 [3]. Despite being common, medical interventions are often associated with a range of adverse physical and psychosocial impacts. Many patients experience fear and anxiety [4-8]; while during and following the intervention many patients experience an array of side effects, such as fatigue [9], pain [10] and distress [11].

Preparing patients for medical intervention improves outcomes

Informing patients about the nature, benefits, risks, alternatives and consequences of the intervention [12] is an ethical and legal requirement [13]. This process is commonly undertaken verbally by the healthcare provider, and may be supplemented with written, video or web-based information [14]. It is recommended that the following content is covered during patient preparation: procedural information (e.g. the sequence of events, the equipment used); sensory information (e.g. sensations that may be felt by the patient before, during or after intervention) [15]; behavioural instruction (e.g. the patients' expected role and what they can do to facilitate the medical intervention or their recovery) [16]; and psychosocial aspects; for example, relaxation training (e.g. breathing exercises or hypnosis); cognitive coping strategies (e.g. coping statements); and emotion-focussed interventions (e.g. discussion of the patient's emotions) [16, 17]. Systematic reviews indicate that preparing patients for medical interventions is related to improved physical and psychological outcomes, and increased patient satisfaction and knowledge [18-22]. A meta-analysis revealed that procedural information and behavioural instruction were the most effective preparatory approaches for surgery in relation to outcomes including pain, negative affect, length of stay, behavioural recovery, clinical recovery, physiological indices and satisfaction [23]. Despite the

existence of such evidence-based recommendations, some patients report suboptimal preparation for medical interventions [24, 25].

Patient self-report is the most appropriate approach for assessing patients' preparation for medical interventions

Assessing the quality of patient preparation allows for monitoring and improving the delivery of healthcare [26, 27]. A variety of approaches including patient feedback surveys [28], audits of medical records administrative systems, and self-report by health care providers have been used to assess quality of care [29]. There are limitations to each of these approaches. For example, aspects of patient preparation are often poorly recorded in the medical record [30], and provider self-report may not align well with patients' perceptions of care [31]. Given the subjective nature of "feeling prepared", patient self-report is the most appropriate method for assessing a patient's level of preparation for a medical intervention. Patients' views about whether care was received and helpful, is an essential component of assessing quality of care [32].

Psychometrically robust instruments assessing patients' preparation for medical interventions are needed

A standardised instrument must be valid, reliable, brief, clinically relevant and easy to administer and interpret [33]. To adequately assess the quality of patient preparation, a comprehensive assessment would cover all aspects of care in relation to preparing patient for medical interventions as recommended in the literature and current guidelines [1, 34]. This would include: patient's involvement in decision-making, provision of risk information and acquiring informed consent, provision of sensory and procedural information, behavioural instruction, and psychosocial aspects [15-17]. A systematic review of the standardised instruments designed to specifically assess patients' level of preparation for medical interventions will help to identify the most rigorous and appropriate instrument to assess the quality of healthcare in this area. A number of published reviews have examined instruments that assess one or two concepts that could be considered relevant to patient preparation, for example, measures of shared decision-making and satisfaction

[26, 35-37]. However, no review has critically examined comprehensive instruments that cover the preparatory content areas (procedural, sensory, behavioural, psychosocial and risk information) recommended for patients.

This systematic review aims to critically examine the characteristics and psychometric properties of published instruments which specifically assess patients' self-reported preparation for medical interventions.

2. METHODS

Electronic database search: Medline, CINAHL, EMBASE and PsycINFO databases were searched to identify publications which describe the development of instruments for assessing patients' perceptions of their preparation for medical interventions. Searches were limited to English language publications, published from the date of inception of each database (Medline 1946, CINAHL 1961, Embase 1980 and PsycINFO 1967) until the 10 November 2015.

The database search was performed using a combination of subject headings and text search terms for informed consent; decision making; quality of care; patient perceptions; education; preparation; psychometrics; and interventions. See appendix A for the complete Medline search strategy. Subject headings were modified according to the specifications of the other databases. Full details of the search strategy for each database are available. Reviews of relevant literature and the reference lists of retrieved articles were also searched.

Inclusion and exclusion criteria

Studies that examined patients' experience, perception, or satisfaction with preparation for medical interventions were included. Preparation was defined as the education, instruction or support provided to a patient relating to their preparation for the period before, during or after the medical intervention. Instruments were included if they were developed to examine the quality of the

following preparation methods: the provision of procedural information; the provision of sensory information; behavioural instruction; and psychosocial aspects (e.g. relaxation training; cognitive coping strategies; emotion-focussed interventions).

Inclusion criteria: Papers were included that: 1) described the development and/ or validation of a self-report instrument specifically designed to assess the quality of patients' preparation for a medical intervention; 2) were designed for adults aged 18 years or older; and 3) assessed preparation from the patient's viewpoint.

Exclusion criteria: Papers were excluded if they: 1) reported on instruments not designed specifically to assess patient preparation; 2) were published in a language other than English; or 3) were non data-based (e.g. conference abstracts, erratums, reviews or commentaries).

Literature search

The titles of all articles identified were assessed for relevance by one reviewer (KF). The abstracts of articles that met the initial screening for relevance were then assessed for relevance and eligibility by the same reviewer (KF). A second reviewer (AH) cross-checked a random selection of 20% of the titles and abstracts to confirm their inclusion and exclusion status. Any discrepancies were resolved through discussion, and the decisions applied systematically to all identified articles. Each of the remaining full text articles were assessed for eligibility by two reviewers (KF and AH). Any discrepancies regarding eligibility were resolved by discussion. A search of all databases using the name of each identified instrument was then performed to ensure that all publications relating to the development or validation of the instrument were included. The authors of each instrument were also contacted and requested to provide any unpublished data they had available on the development of the instrument or its psychometric properties.

Data extraction

The general characteristics and psychometric properties of each instrument were extracted and coded by two authors (KF and AH). Any discrepancies were resolved by discussion.

General characteristics of instruments:

The following details were extracted to examine the comprehensiveness of the instrument: a) number of items; b) domains; c) response scale; d) whether the tool was designed to assess quality of preparation pre- or post- intervention and e) population included, sample size and response rate. We also assessed the content areas covered by the included instruments. Guideline recommendations [1, 34] and available international evidence [15, 16, 23, 38] were used to develop a checklist of content areas that should be covered when preparing patients for a medical intervention. The areas included in the checklist were: procedural information, sensory information, behavioural instruction, psychosocial aspects and risk information. Two authors (KF and AH) used this checklist to independently assess the comprehensiveness of each instrument. The results were compared and any discrepancies regarding coding were resolved by discussion.

Psychometric properties

Measurement properties of the instruments were evaluated using standard criteria for scale development as outlined in Table 1 [39]. Two commonly used statistical approaches for developing scales are Classical Test Theory (CTT) and Item Response Theory (IRT) [33, 40]. As none of the instruments included in this review used IRT analyses, the methods and results of this review are based on CTT.

3. RESULTS

A total of 2,767 potentially relevant publications were identified (Figure 1). Of these, 677 were duplicates and 1,940 were excluded based on either title or abstract review. Of 150 full-text publications reviewed, nine publications were identified describing the development or validation of

seven instruments. The seven instruments identified were: 1) Preoperative Teaching Interview Guide (PTIG) [41]; 2) RT Concerns Scale and RT Information Needs Scale [42]; 3) The Wisdom Tooth Knowledge Scale [43]; 4) Patient Satisfaction with Cancer Treatment Education (PS-CaTE) [44]; 5) The Satisfaction with Cancer Information Profile (SCIP) [45-47]; 6) Cancer Treatment Survey (CaTS) [48]; and 7) Quality Of Care Through the patients' Eyes (QUOTE^{chemo}) [49].

Table 2 provides an overview of the instruments. Five of the instruments were developed to examine cancer-related medical interventions [42, 44, 45, 48, 49], one examined day-surgery interventions [41] and one examined knowledge of wisdom tooth removal [43]. Data collection time points were reported for all instruments. One assessed patient's preparation both pre and post-intervention [45], one pre-intervention [48] and four post-intervention [41, 42, 44, 49]. The remaining study assessed knowledge before and after provision of an information leaflet in relation to wisdom tooth removal [43]. There was significant variability in sample size with sizes ranging from 31 [42] to 345 [49]. In addition, response rates for only three instruments were clearly reported [42, 45, 49]. As outlined in Table 3, one instrument [48], the Cancer Treatment Survey, assessed components of all of the content areas generally recommended for preparing patients for a medical intervention. The content areas covered by the Wisdom Tooth Knowledge Scale [43] could not be assessed as a list of items is not available. Across all of the included instruments, the most frequently assessed content areas were psychosocial aspects (6 instruments), procedural information (6 instruments), and behavioural instruction (6 instruments).

Psychometric properties

Reliability

Table 4 shows reliability indices for each of the seven instruments.

Internal consistency

Internal consistency was assessed for three instruments at the total scale level [41, 42, 44]. Adequate internal consistency was reported for all, with Cronbach alpha values ranging between 0.81 and 0.95 [41, 42, 44]. At the sub-scale or domain level, Cronbach alpha values were reported for six instruments [41, 42, 44, 45, 48, 49]. Adequate values were reported for four of these instruments, with values ranging from 0.72 to 0.92 [42, 44, 45, 49]. Both of the subscales of the CaTS [48] were just outside the upper limit of adequate Cronbach alpha values, at 0.96 and 0.97 respectively. Of the five information domains of the Preoperative Teaching Interview Guide [41], only one Cronbach's alpha value was adequate in the Preoperative Teaching Received subscale, and three were adequate for the Preoperative Teaching Valued sub-scale.

Item-to-total correlations were reported for only three instruments [42, 47, 49]. Values were acceptable for all items for both the Quote^{chemo} and SCIP instruments; ranging from 0.35 to 0.79 for the Quote^{chemo} and from 0.39-0.73 and 0.52-0.74 for both subscales of the SCIP, respectively [47]. In relation to the RT Concerns and RT Information Needs Scales, values ranged from 0.30 to 0.79 and were adequate for the RT Concerns Scale items. However, values for the RT Information Needs Scale items ranged from 0.06-0.74 and were adequate for 18 of the 22 items. One additional study reported inter-item correlations but not item-to-total correlations [48].

Test-retest

Test-retest reliability was reported for only two instruments [42, 43]. In relation to the total instrument, the Wisdom Tooth Knowledge Scale demonstrated acceptable reliability with a Pearson's Correlation Coefficient of 0.73 [43]. However, inadequate coefficients were reported for the RT Concerns and RT Information Needs Scales [42]; with mean ICCs of 0.60 and 0.55 respectively. A more powerful measure of test-retest reliability is provided when the completed individual items are examined at time one and time two. The only instrument for which this task was undertaken was the RT Concerns and RT Information Needs Scales; and inadequate coefficients were also reported.

Two of the nine items of the RT Concerns Scale and 6 of the 24 items of the RT Information Needs Scale achieved 0.70 or greater.

Validity

Table 5 shows the validity indices for each of the seven instruments.

Face and Content Validity

Face or content validity was examined for all instruments. *Face validity* involving patients was adequately assessed for five of the instruments [42, 44, 45, 48, 49]. However, few details were provided for the SCIP scale [45]. *Content validity* was adequately assessed by experts including healthcare professionals for six instruments.

Construct Validity

Aspects of construct validity were assessed by *factor analysis* for three instruments [47-49], with the final factor structures reporting a good fit to the data for all three instruments. Table 5 presents the details of the specific factor analyses performed and the associated results for the SCIP[47], CaTS[48] and QUOTE^{chemo}[49] instruments.

Convergent validity: Two instruments, the SCIP[45] and the QUOTE^{chemo}[49] assessed convergent validity. For the SCIP, the relationship between satisfaction with information and a scale measuring pre-treatment illness beliefs was assessed. However, adequate correlation coefficients were not demonstrated. Convergent validity for the QUOTE^{chemo}[49] instrument was reported with three other scales examining information preferences, monitoring, intrusive and avoidant thinking. However, whilst the authors reported satisfactory correlations, these were also less than the 0.40 criteria and thus convergent validity was deemed not demonstrated.

Divergent validity: Divergent validity was established for only one instrument, with correlations for both subscales of the CaTS being assessed with the Hospital Anxiety and Depression Scale (HADS) [48]. A significant, weak association was reported between the sensory/psychological subscale and both the HADS total scale and the anxiety subscale; 0.24 and 0.26 respectively. A significant, weak association of 0.15 between the procedural concerns subscale and the HADS anxiety subscale was

also reported. However, the association between this subscale and the HADS total scale was very weak and not significant. The results indicate that both subscales of the CaTS were related to, but also distinguishable from anxiety and emotional distress.

Known groups or discriminate validity was assessed for three of the seven instruments [43, 45, 48]. All three instruments reported statistically significant differences amongst the groups. Results demonstrated that dental students possessed a significantly greater knowledge level in relation to wisdom tooth removal than sixth-formers and relatives [43]; and younger participants had significantly greater chemotherapy related procedural concerns [48]. Discriminative validity also confirmed that patient characteristics such as levels of optimism did not influence satisfaction with illness and treatment information in head and neck cancer patients [45]. The ability to discriminate between different levels of individual satisfaction was reported for the SCIP instrument [47].

Criterion Validity

Concurrent validity: The concurrent validity of one instrument assessing satisfaction with cancer information was assessed [45], but it was labelled as “concurrent (convergent) validity”, grouping concurrent validity and convergent validity as one. As there is no gold standard instrument measuring satisfaction with information, for this review we deemed that concurrent validity was not appropriately assessed nor met for this instrument. However, methods were consistent with convergent validity, and for this review were assessed as such.

Predictive validity: Predictive validity was examined for one instrument. In relation to the SCIP, it was reported that satisfaction with the content and amount of information was significantly associated with quality of life and depression scores one month after treatment, with those less satisfied more likely post-treatment to have lower quality of life scores and high levels of depression [45]. In addition, lower levels of satisfaction with information before treatment was found to be predictive of higher levels of depression and lower health-related quality of life scores 6-8 months after treatment [46].

Cross-cultural validity

One instrument, the PS-CaTE [44], was translated and adapted from Canadian English to German following a forward/backward method, with similar reliability indices to the Canadian version reported. However, the authors recommended that further research should be conducted to verify the instrument's validity.

Responsiveness

The ability to detect change over time was assessed for only one instrument, the SCIP [45]. This was assessed using only one possible measure of responsiveness, via the ability of the instrument to detect a change in satisfaction after treatment by examining both individual items and the subscales. McNamar and Wilcoxon signed-ranks tests were used to examine the change in scores. The results indicate that the instrument is responsive to change over time based on the criteria of responsiveness used, with both the ability to detect change in specific areas important to patients and the direction of the change demonstrated. However, whilst scores on subscale 2 (satisfaction with form and timing of information) were significantly lower post-treatment; no significant post-treatment changes in subscale 1 (satisfaction with content and amount of information) items were reported.

The assessment of floor or ceiling effects for the total score or subscale scores was reported for one instrument [41]. The reported percentage of highest ratings given by subscale of the PTIG [41] provided an indicator that ceiling effects were present for three of the PTIG information subscales for teaching received. Ceiling effects for six of the seven items of the Satisfaction with the Form and Timing of the Information received subscale of the SCIP were also reported [47]. The scores on the Satisfaction with Content and Amount of Information subscale of the SCIP were negatively skewed [45], another indicator of possible ceiling effects. Although not specifically assessed, there were suggestions that possible ceiling or floor effects existed in relation to three other instruments [43, 44, 49]. In two instances there was data illustrating a high percentage (more than 15%) of respondents selecting the highest possible response option on a number of items in these instruments [44, 49]. For the third instrument, the authors made a comment in the discussion that

more than half of the items were answered correctly by a large proportion of respondents [43].
However, the assessment of ceiling or floor effects for the total score was not specifically reported
for any of these instruments. No indicators of floor and ceiling effects were reported for two
instruments [42, 48].

Acceptability

Response rates: As shown in Table 4, aspects of acceptability were reported for five of the seven instruments [41, 42, 44, 45, 49]. However, response rates were clearly reported for only three instruments [42, 45, 49]. Values were acceptable for the three instruments and ranged from 59-86%. The percentage of missing items was reported for only one of the instruments [45]. Reported values were acceptable, with individual item response rates ranging from 79-95% and 91-98% for subscales 1 and 2 of the SCIP, respectively.

Reading level was not reported for any of the instruments. However, this lack of assessment was justified for one of the instruments as the questionnaire was read to each participant, thus reading level was not applicable [41].

Time to complete: The time to complete the questionnaire was reported for one instrument only, and was acceptable at 15-20 minutes [41].

Feasibility

Feasibility was examined for only one of the seven instruments[41], reported as the time for a health professional to administer the instrument via interview. The cost, time to administer, score or interpret was not reported for any of the other instruments.

Table 6 provides a summary of the psychometric properties reported for each instrument.

4. DISCUSSION AND CONCLUSION

4.1 Discussion

A large number of other instruments have been developed which assess some aspects of preparation

As shown in figure 1, 79 articles were excluded from this review as they were not specifically developed to examine preparation for a medical intervention. However, each of these articles and their associated instruments did assess a concept that could be considered relevant to patient preparation, such as information provision in relation to perioperative period, patient satisfaction with care, or supportive care needs. However, these instruments also assessed other areas not relevant to patient preparation, and thus did not meet the eligibility criteria of this review. For example, one instrument which measured satisfaction in the perioperative period surrounding general anaesthesia [50] included relevant aspects such as satisfaction with information provision; but also included other dimensions not relevant to preparation, such as attention, privacy, discomfort, and waiting [50].

Comprehensiveness of available instruments

Despite the ethical and clinical imperative to adequately prepare patients for medical interventions, this review identified only seven instruments developed to assess patients' self-reported preparation for medical interventions. Only one of the instruments, the Cancer Treatment Survey [48], covered all of the content areas generally recommended for preparing patients for medical interventions. However, it should be noted that there are no clear internationally recognised criteria for preparing patients. The author-developed content criteria used in this review were based on existing Australian guidelines [1, 34] and international evidence. However, this research was mainly conducted in the 1980s with surgery patients [15, 16, 23, 38]. Content coding was difficult for some instruments due to vague wording and overlap of some items. However, coding agreement between two authors was achieved by in-depth discussion. Psychosocial aspects were comprehensively covered by all instruments. All other content areas were comprehensively covered by most instruments with the exception of sensory information. Only two out of the instruments assessed

this area comprehensively. The lack of instruments addressing the provision of sensory information is a particular concern, given that systematic reviews have shown that this component of preparation assists with reduction of anxiety, distress and pain [21-23]. These findings highlight the need to develop instruments which measure the breadth of commonly recommended preparatory content.

The data collection time-points also potentially limit the interpretation of some of the findings in relation to patients' preparation. Whilst reported for all studies, the level of detail provided varied, with only three of the six instruments that assessed post-intervention specifying how soon after the intervention the assessment occurred [41, 43, 45]. Recall bias may therefore be an issue, particularly in relation to the RT Concerns and RT Information Needs Scales, due to the fact that participants' may have completed this instrument up to 18 months after finishing treatment [42]. Two instruments were validated for use both pre and post intervention [43, 45]. One additional instrument assessed pre-chemotherapy [48], and developed an instrument useful for the assessment of concerns in patients scheduled to receive their first course of treatment. However, as it was developed for the assessment of the adequacy of preparation pre-intervention only, it is not a suitable instrument for use to examine patient's perceptions of the adequacy of their preparation once they have actually had treatment.

Limitations with regard to generalizability of available instruments

Much of the research has focussed on patients undergoing cancer-related interventions, with over 70% of the reviewed instruments (five out of seven) developed for the assessment of cancer-related medical interventions (surgery, chemotherapy or radiotherapy). The majority of the cancer-related instruments are limited in their generalizability as most were validated with samples comprising solely or mainly of breast cancer patients [42, 44, 48].

Of the two instruments that did not examine a cancer-related medical intervention, one assessed preoperative teaching in relation to day surgery; the other knowledge of wisdom tooth removal. However, generalisability of the latter instrument is extremely poor due to the fact that, despite being designed to assess the patient perspective, the instrument was only validated with a student and relative sample who may or may not have undergone this intervention previously. Thus, whether this instrument is relevant to and accurately reflects aspects of preparation important to patients undergoing wisdom tooth removal is unknown.

Overall, the generalisability of the reviewed instruments is likely limited by small sample sizes with most less than 200, which may lead to erroneous assessment of the instrument's psychometric properties [51]. One of the instruments was developed and tested among a sample size over 300 [49]. Whilst there is no consensus defining an appropriate sample size [51], Comrey and Lee [52] have suggested that for the purposes of factor analysis sample sizes less than 100 are poor, 300 as good, 500 as very good, and greater than 1000 as being excellent. For the three instruments for which a factor analysis was performed, only one had a sample size greater than 300 [49]. Further limiting the findings about the properties is the lack of clearly reported response rates for four of the studies and the lack of reporting of the inclusion or exclusion criteria for three studies. Without these details, it is difficult to confidently confirm the accuracy of the instrument due to the inability to assess for potential response bias, and determine if the study sample adequately covered the target population.

The psychometric quality of the currently available instruments is generally poor

Whilst most of the instruments had established internal consistency, test-retest reliability was poorly examined, with only two of the seven studies examining this form of reliability. Consequently, the stability of five of the instruments over time is not known. Further, only one of the seven instruments [42] assessed both internal consistency and test-retest reliability. Adding to the issue of

detecting changes over time, only one study examined the responsiveness (or sensitivity) to change over time, with mixed results [45].

In relation to the psychometric properties for validity, content validity was the most frequently assessed, with all instruments assessing this property. However, face validity was less widely confirmed, with only five of the instruments involving patients [42, 44, 45, 48, 49]. Without the involvement of patients, it remains unclear whether the instrument is appropriate and relevant to the intended users. Only three of the instruments performed a factor analysis to establish the construct validity of the scale. As there is no current 'gold standard' instrument measuring patient preparedness, it was not possible for the current instruments to assess concurrent validity, a sub-type of criterion validity. Whilst predictive validity was assessed for only one of the instruments, this is not surprising as it is a difficult property to assess, given that it necessitates a longer study and there are logistical problems with the method [39].

Given the poor assessment and reporting of acceptability and feasibility for most of the instruments contained in this review, any burdens that these instruments place on either the patient or health care provider are poorly understood. For instance, it is unknown whether the interview mode of delivery used by the PTIG [41] affects its widespread use given the increased time required of healthcare professionals to administer it, compared to a pen-and-paper patient self-report instrument. For such measures to be routinely used in the health care setting, their feasibility and acceptability in real-world practice must be determined.

Based on the reviewed psychometric properties for each of the seven instruments, there is no one instrument that meets all of the necessary criteria for validity and reliability. In particular, test-retest reliability, construct validity, acceptability and feasibility were poorly assessed. Further testing of each instrument is needed to establish the psychometric qualities of the seven instruments included

in this review. However, it appears that the SCIP [45-47], CaTS [48] and QUOTE^{chemo} [49] instruments are the most psychometrically valid and reliable instruments currently available for the assessment of patient preparation for medical interventions. However, further testing of each instrument is recommended to further establish their psychometric qualities and feasibility for use in the real-world setting.

Limitations and strengths

The literature search did not include grey literature, thus it is possible that some relevant instruments may have been missed. A further issue is the lack of a standardized taxonomy related to preparing patients for medical interventions, with no standard method for the indexing papers in the area and no clear internationally recognised criteria for preparing patients. This may have led to some relevant papers being overlooked. However, the comprehensive search strategy utilized reduced the likelihood of missing any relevant instruments. For this same reason, the evaluation of the comprehensiveness of each instrument in relation to preparatory content should be considered exploratory and somewhat subjective. Future research should strive to develop a standardised criteria for what warrants high quality preparation for patients. There were a large number of instruments that whilst not designed specifically to assess patient preparation, did assess some concepts that could be considered relevant to patient preparation (e.g. supportive care needs). Thus while only seven instruments were identified that specifically assessed patient preparation, there are likely to be copious others that assess at least some content areas relevant to patient preparation. However, such instruments are not specifically focused on patient preparation for a medical intervention, and thus do not provide a comprehensive assessment of this construct. The psychometric assessment of each instrument was based on the details provided in the published literature. It is possible that further details of additional psychometric properties may exist for each instrument. An effort was made to contact the authors of each instrument to obtain details of any unpublished data. However, only two responses were received and no additional data was reported nor provided.

4.2 Conclusion

There are few instruments that have been specifically developed to address patients' preparation for medical interventions. Promising instruments have been identified, primarily in relation to cancer-related interventions. However, further testing of psychometric properties of each instrument is recommended. Future research with larger sample sizes is also suggested.

4.3 Practice Implications

The need for the development of instruments examining patient preparation for medical interventions is apparent given the limitations of the instruments reviewed. An instrument that covers all of the consistently recommended preparatory content for medical interventions and addresses the above limitations is warranted.[53] Any new instrument developed should aim to meet the psychometric criteria as outlined in the review.

ACKNOWLEDGEMENTS

This research was supported by a Strategic Research Partnership Grant (CSR 11-02) from the Cancer Council NSW to the Newcastle Cancer Control Collaborative (NEW-3C), a Cancer Institute New South Wales Evidence to Practice Grant, and infrastructure funding from the University of Newcastle and Hunter Medical Research Institute (HMRI). Ms Kristy Forshaw is supported by a University of Newcastle Postgraduate Research Scholarship. Dr Mariko Carey is supported by a National Health and Medical Research Council Translating Research into Practice (TRIP) Fellowship (APP1073031). Dr Allison Boyes is supported by Early Career Fellowships from the National Health and Medical Research Council (APP1073317) and Cancer Institute NSW (13/ECF/1-37).

CONFLICTS OF INTEREST: None

AUTHORS' CONTRIBUTIONS

KF, MC and RSF conceived of and designed the study. KF and AH undertook data extraction. KF, MC, AH and AB assisted in the analysis and interpretation of the data. All authors contributed to drafting of the manuscript and read and approved the final manuscript.

REFERENCES

- [1] NHMRC. Communicating with Patients: Advice for Medical Practitioners. Canberra: NHMRC; 2004.
- [2] Australian Institute of Health and Welfare. Australia's hospitals 2012-13 at a glance. Health services series no. 55. Cat. no. HSE 146. Canberra: AIHW; 2014.
- [3] CDC/NCHS. National Hospital Discharge Survey. Number of all-listed procedures for discharges from short-stay hospitals, by procedure category and age: United States, 2010. Hyattsville, MD: National Center for Health Statistics; 2010.
- [4] Munn Z, Jordan Z. The patient experience of high technology medical imaging: A systematic review of the qualitative evidence. *Radiography*. 2011;17:323-31.
- [5] Worster B, Holmes S. The preoperative experience of patients undergoing surgery for colorectal cancer: a phenomenological study. *Eur J Oncol Nurs*. 2008;12:418-24.
- [6] Breen S, Baravelli CM, Schofield PS, Jefford M, Yates P, Aranda SK. Is symptom burden a predictor of anxiety and depression in patients with cancer about to commence chemotherapy? *Med J Aust*. 2009;190:S99-S104.
- [7] Hess CB, Chen AM. Measuring psychosocial functioning in the radiation oncology clinic: a systematic review. *Psychooncology*. 2014;23:841-54.
- [8] Guo Z, Tang HY, Li H, Tan SK, Feng KH, Huang YC, et al. The benefits of psychosocial interventions for cancer patients undergoing radiotherapy. *Health Qual Life Outcomes*. 2013;11:121.
- [9] Servaes P, Verhagen C, Bleijenberg G. Fatigue in cancer patients during and after treatment: prevalence, correlates and interventions. *Eur J Cancer*. 2002;38:27-43.
- [10] Shang AB, Gan TJ. Optimising postoperative pain management in the ambulatory patient. *Drugs*. 2003;63:855-67.
- [11] Cremeans-Smith JK, Greene K, Delahanty DL. Symptoms of postsurgical distress following total knee replacement and their relationship to recovery outcomes. *J Psychosom Res*. 2011;71:55-7.

- [12] Leclercq WK, Keulers BJ, Scheltinga MR, Spauwen PH, van der Wilt GJ. A review of surgical informed consent: past, present, and future. A quest to help patients make better decisions. *World J Surg.* 2010;34:1406-15.
- [13] National Quality Forum (NQF). Improving Patient Safety Through Informed Consent for Patients with Limited Health Literacy. An Implementation Report. In: Wu HW, Nishimi RY, Page-Lopez CM, Kizer KW, editors. Washington, DC: NQF; 2005.
- [14] Kinnersley P, Phillips K, Savage K, Kelly MJ, Farrell E, Morgan B, et al. Interventions to promote informed consent for patients undergoing surgical and other invasive healthcare procedures. *Cochrane Database Syst Rev.* 2013;7:CD009445.
- [15] Ludwick-Rosenthal R, Neufeld RW. Stress management during noxious medical procedures: an evaluative review of outcome studies. *Psychol Bull.* 1988;104:326-42.
- [16] Powell R, Bruce J, Johnston M, Vögele C, Scott N, Shehmar M, et al. Psychological preparation and postoperative outcomes for adults undergoing surgery under general anaesthesia. CD008646. *Cochrane Database Syst Rev.* 2010.
- [17] Cochran TM. Psychological preparation of patients for surgical procedures. *Patient Educ Couns.* 1984;5:153-8.
- [18] National Breast Cancer Centre and National Cancer Control Initiative. Clinical practice guidelines for the psychosocial care of adults with cancer. Camperdown, NSW: National Breast Cancer Centre; 2003.
- [19] Devine EC, Westlake SK. The effects of psychoeducational care provided to adults with cancer: meta-analysis of 116 studies. *Oncol Nurs Forum.* 1995;22:1369-81.
- [20] Waller A, Forshaw K, Bryant J, Mair S. Interventions for preparing patients for chemotherapy and radiotherapy: a systematic review. *Support Care Cancer.* 2014;22:2297-308.
- [21] Suls J, Wan CK. Effects of sensory and procedural information on coping with stressful medical procedures and pain: a meta-analysis. *J Consult Clin Psychol.* 1989;57:372-9.

- [22] Hathaway D. Effect of preoperative instruction on postoperative outcomes: a meta-analysis. *Nursing research*. 1986;35:269-75.
- [23] Johnston M, Vögele C. Benefits of psychological preparation for surgery: A meta-analysis. *Ann Behav Med*. 1993;15:245-56.
- [24] Coulter A, Collins A. Making shared decision making a reality: no decision about me, without me. Available from: www.kingsfund.org.uk. The King's Fund; 2011.
- [25] Bureau of Health Information. Adult Admitted Patient Survey Results for all inpatients. Jan-June 2013. 2014.
- [26] Crow R, Gage H, Hampson S, Hart J, Kimber A, Storey L, et al. The measurement of satisfaction with healthcare: implications for practice from a systematic review of the literature. *Health Technol Assess*. 2002;6:1-244.
- [27] Garratt AM, Bjaertnes OA, Krogstad U, Gulbrandsen P. The OutPatient Experiences Questionnaire (OPEQ): data quality, reliability, and validity in patients attending 52 Norwegian hospitals. *Qual Saf Health Care*. 2005;14:433-7.
- [28] Australian Commission on Safety and Quality in Health Care. Patient-centred care: Improving quality and safety through partnerships with patients and consumers. Sydney: ACSQHC; 2011.
- [29] Carey M, Paul C, Sanson-Fisher R, Buchan H. Data systems for assessing quality of cancer care: Are building blocks in place for person-centered care? *Int J Pers Cent Med*. 2012;2.
- [30] Holmboe E. Practice Audit, Medical Record Review, and Chart-Stimulated Recall. *Practical Guide to the Evaluation of Clinical Competence*. Pennsylvania: Mosby; 2008.
- [31] Mohammed K, Nolan MB, Rajjo T, Shah ND, Prokop LJ, Varkey P, et al. Creating a Patient-Centered Health Care Delivery System: A Systematic Review of Health Care Quality From the Patient Perspective. *Am J Med Qual*. 2014.
- [32] Coulter A. Can patients assess the quality of health care? *BMJ*. 2006;333:1-2.
- [33] Keszei AP, Novak M, Streiner DL. Introduction to health measurement scales. *J Psychosom Res*. 2010;68:319-23.

- [34] Schofield MJ, Sanson-Fisher R. How to prepare cancer patients for potentially threatening medical procedures: consensus guidelines. NSW Cancer Council Cancer Education Research Program. J Cancer Educ. 1996;11:153-8.
- [35] NHS. Measuring Shared Decision Making. A review of research evidence. A report for the Shared Decision Making programme. In partnership with Capita Group Plc. Available from: http://www.rightcare.nhs.uk/wp-content/uploads/2012/12/Measuring_Shared_Decision_Making_Dec12.pdf
- [36] Scholl I, Koelewijn-van Loon M, Sepucha K, Elwyn G, Legare F, Harter M, et al. Measurement of shared decision making - a review of instruments. Zeitschrift fur Evidenz, Fortbildung und Qualitat im Gesundheitswesen. 2011;105:313-24.
- [37] Elwyn G, Edwards A, Mowle S, Wensing M, Wilkinson C, Kinnersley P, et al. Measuring the involvement of patients in shared decision-making: a systematic review of instruments. Patient Educ Couns. 2001;43:5-22.
- [38] Schultheis K, Peterson L, Selby V. Preparation for stressful medical procedures and person × treatment interactions. Clin Psychol Rev. 1987;7:329-52.
- [39] McDowell I. Measuring Health: A Guide to Rating Scales and Questionnaires. 3rd ed. New York, NY: Oxford University Press; 2006.
- [40] Hambleton RK, Jones RW. Comparison of Classical Test Theory and Item Response Theory and Their Applications to Test Development. J Edu Meas. 1993;12:38-47.
- [41] Bernier MJ, Sanares DC, Owen SV, Newhouse PL. Preoperative teaching received and valued in a day surgery setting. AORN journal. 2003;77:563-82.
- [42] Halkett GK, Kristjanson LJ. Validity and reliability testing of two instruments to measure breast cancer patients' concerns and information needs relating to radiation therapy. Radiat Oncol. 2007;2:43.
- [43] Humphris GM, O'Neill P, Field EA. Knowledge of wisdom tooth removal: influence of an information leaflet and validation of a questionnaire. Br J Oral Maxillofac Surg. 1993;31:355-9.

- [44] Liekweg A, Eckhardt M, Taylor SC, Erdfelder E, Jaehde U. Psychometric assessment and application of a questionnaire measuring patient: satisfaction with information on cancer treatment. *Pharm World Sci.* 2005;27:96-103.
- [45] Llewellyn CD, Horne R, McGurk M, Weinman J. Development and preliminary validation of a new measure to assess satisfaction with information among head and neck cancer patients: the satisfaction with cancer information profile (SCIP). *Head Neck.* 2006;28:540-8.
- [46] Llewellyn CD, McGurk M, Weinman J. How satisfied are head and neck cancer (HNC) patients with the information they receive pre-treatment? Results from the satisfaction with cancer information profile (SCIP). *Oral Oncol.* 2006;42:726-34.
- [47] Hankins MC, Llewellyn CD. Is the Satisfaction with Cancer Information Profile (SCIP) valid for tailoring information for patients with head and neck cancer? *BMC Cancer.* 2008;8:164.
- [48] Schofield P, Gough K, Ugalde A, Carey M, Aranda S, Sanson-Fisher R. Cancer Treatment Survey (CaTS): development and validation of a new instrument to measure patients' preparation for chemotherapy and radiotherapy. *Psychooncology.* 2012;21:307-15.
- [49] van Weert JC, Jansen J, de Bruijn GJ, Noordman J, van Dulmen S, Bensing JM. QUOTEchemo: a patient-centred instrument to measure quality of communication preceding chemotherapy treatment through the patient's eyes. *Eur J Cancer.* 2009;45:2967-76.
- [50] Auquier P, Pernoud N, Bruder N, Simeoni MC, Auffray JP, Colavolpe C, et al. Development and validation of a perioperative satisfaction questionnaire. *Anesthesiology.* 2005;102:1116-23.
- [51] Anthoine E, Moret L, Regnault A, Sébille V, Hardouin J-B. Sample size used to validate a scale: a review of publications on newly-developed patient reported outcomes measures. *Health Quality Life Outcomes.* 2014;12:1-10.
- [52] Comrey A, Lee H. *A First Course in Factor Analysis.* 2nd ed. Hillsdale, NJ: Lawrence Erlbaum Associates; 1992.

- [53] Booij JC, Zegers M, Evers PM, Hendriks M, Delnoij DM, Rademakers JJ. Improving cancer patient care: development of a generic cancer consumer quality index questionnaire for cancer patients. *BMC Cancer*. 2013;13:203.
- [54] Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med*. 2009;6:e1000097.
- [55] DeVellis R. *Scale Development: Theory and Application*. 2nd ed. Thousand Oakes, CA: Sage; 2003.
- [56] Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol*. 2007;60:34-42.
- [57] Streiner D.L, G.R. N. *Health Measurement Scales: A Practical Guide to their Development and Use* (3rd edn). New York: Oxford University Press; 2003.
- [58] Lohr KN, Aaronson NK, Alonso J, Audrey Burnam M, Patrick DL, Perrin EB, et al. Evaluating quality-of-life and health status instruments: development of scientific review criteria. *Clinical Therapeutics*. 1996;18:979-92.
- [59] Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33:159-74.
- [60] Anastasi AU, S. *Psychological Testing*. Upper Saddle River, NJ: Prentice Hall; 1997.
- [61] Cohen J. *Statistical Power Analysis for the Behavioural Sciences*. Hillsdale, NJ: Erlbaum; 1988.
- [62] Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol 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.
- [63] Pearce NJM, Sanson-Fisher R, Campbell HS. Measuring quality of life in cancer survivors: a methodological review of existing scales. *Psychooncology*. 2008;17:629-40.
- [64] Brazier J, Deverill M. A checklist for judging preference-based measures of health related quality of life: learning from psychometrics. *Health Econ*. 1999;8:41-51.

LEGENDS

Tables and Figures

Table 1. Psychometric properties assessed in this review

Table 2. General characteristics of instruments measuring preparation for medical interventions included in this review

Table 3. Comprehensiveness of each instrument in relation to preparation content

Table 4. Reliability, acceptability and feasibility of each instrument

Table 5. Validity of each instrument

Table 6. Summary of the psychometric properties of each instrument

Figure 1. Flowchart of the inclusion and exclusion process based on Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) guidelines [54]

Supplementary material: APPENDIX A: Search terms from Medline (extracted 10/11/15)

Table 1: Psychometric properties assessed in this review

Property	Definition	Criteria for assessing acceptability of reported values
RELIABILITY	Error in measurement [39]	
<i>Internal consistency</i>	Extent to which items on a scale are homogeneous [55]	Cronbach's coefficient alpha values >0.70[39]-0.95 [56] or Kuder-Richardson formula 20 values >0.70 (for dichotomous scales). Item-total correlations between 0.20-0.80 [57].
<i>Test-retest reliability</i>	Reproducibility, or an instrument's stability over time [58]	Kappa coefficient >0.60 (for nominal or ordinal scales) [59]; interclass correlation coefficient (ICC) or Pearson correlation coefficient >0.70 [39]. Time between administrations: whilst 7-14 days usually considered appropriate [56], there is no gold standard. Use of a different time period was considered acceptable if described and justified [56].
VALIDITY	Extent to which an instrument measures what it intends to measure [39]	
<i>Face validity</i>	Subjective assessment of whether a scale appears to measure what it was intended to measure [60]	Target population included in the development of the instrument [56]. Face and content validity considered adequate if both experts and intended users agree that the instrument appears to measure what it was designed to [60].
<i>Content validity</i>	Extent to which items are representative of issue being measured [39, 60]	
Construct validity	The way in which the internal structure of the scale is related to other conceptual constructs [39]	
<i>Factor analysis</i>	Measurement of the underlying theoretical construct.	Use of an appropriate test (e.g. Exploratory or Confirmatory Factor Analysis); and accepted criteria utilised (e.g. Eigenvalues >1 for Exploratory Factor Analysis).
<i>Convergent validity</i>	A positive correlation with another scale measuring the same concept [39]	Pearson correlation coefficients (r) >0.40.
<i>Divergent validity</i>	No correlation with another scale measuring a different concept [39]	Pearson correlation coefficients (r) <0.30 [61].
<i>Known groups</i>	Ability of the scale to differentiate among two or more known groups [55]	Statistically significant differences among different groups [58].
Criterion validity	How well a scale agrees with some criterion or existing 'gold standard' instrument [39, 55]	
<i>Concurrent validity</i>	Ability of the scale to obtain equivalent scores to an alternative 'gold standard' scale measuring the same issue [39]	Rationale for choice of the 'gold standard' instrument appropriate and the correlation between the new scale and the gold standard ≥ 0.70 [56].
<i>Predictive validity</i>	Ability of the scale to predict scores on a future outcome [39]	Reported accuracy with which the scale predicted some future event [39].
<i>Cross-cultural validity</i>	Degree to which a translated or culturally-adapted version displays similar performance to the original instrument [62]	Similar reported indices of reliability and validity between both versions [58].
Responsiveness:	Ability of the scale to detect clinically important changes over time [58]	<u>A justified approach with reported acceptable values. For example effect sizes [39], change scores, sensitivity and specificity [63], or receiver and operator characteristic (ROC) curves[56]; reported ability to distinguish clinically</u>

		<u>important change from measurement error. Evidence of floor and ceiling effects <15%[56] for (total) score.</u>
Acceptability:	Extent to which an instrument is acceptable to those who complete it.	Time to complete ≤20 minutes[63]; evidence of reading age and low levels of non-responders and missing items.
Feasibility:	Relates to the level of burden for those administering the instrument [58]	<u>Instrument freely available (i.e. no cost associated with use), time to administer [64], score and interpret reported.</u>

Table 2: General characteristics of instruments measuring preparation for medical interventions included in this review

Instrument	Items/ Domains	Description of domains	Response scale	Time point/s	Sample
PTIG [41]	26 items, 2 scales, 5 domains	<ul style="list-style-type: none"> - Situational/ procedural (11 items) - Sensation/discomfort (4 items) - Patient role (4 items) - Psychosocial support (2 items) - Skills training (1 item) 	<p><i>Teaching Received Scale:</i> 3-level rating scale from 0=no to 1=yes.</p> <p><i>Teaching Valued Scale:</i> 5-point Likert-type scale from 0=not at all important to 4=very important.</p>	Post-surgery (first post-operative visit)	USA patients undergoing same day surgery at a university-based, academic medical centre. <i>Sample size:</i> 116 (Phase 1: 50; Phase 2: 66)
RT Concerns Scale and RT Information Needs Scale [42]	31 items, 2 scales, 4 subscales	<p><i>RT Concerns Scale:</i> Concerns about radiotherapy (9 items)</p> <p><i>RT Information Needs Scale</i> (22 items):</p> <ul style="list-style-type: none"> - Information about radiation therapy (3 items) - Information relating to planning treatment (4 items) - Information relating to the first day of treatment (10 items) - Effect treatment will have on day to day living during treatment (5 items) 	<p><i>RT Concerns Scale:</i> 9-point Likert-type scale from 0=Not concerned to 9=Very concerned.</p> <p><i>RT Information Needs Scale:</i> 9-point Likert-type scale from 1=least important to 9=most important. Additional rating of whether need for this information had been met (met, partially met, unmet).</p>	Post-radiotherapy (within 18 months of completion)	Australian breast cancer patients who had completed radiation, recruited from one hospital and via local media advertisements. <i>Sample size:</i> 31
The Wisdom Tooth Knowledge Scale [43]	33 Items	<p>Topic areas: problems caused by wisdom teeth, signs of infection;</p> <ul style="list-style-type: none"> - Reasons for wisdom tooth removal - Types of anaesthesia - Surgical procedures (incl. complications) - Postoperative advice and recommendations 	Dichotomous forced-choice (True or false)	Pre- and post-leaflet provision (no procedure received)	UK dental students in fourth year of training (n=55); sixth-formers visiting the dental school for an open day (n=50); relatives of patients attending the dental hospital clinic (n=51) <i>Sample size:</i> 156
PS-CaTE [44]	14 items, 4 subscales	<ul style="list-style-type: none"> - Satisfaction with general treatment information (4 items) - Satisfaction regarding side effects (4 items) - Information regarding vitamins, herbal products and complementary therapies (3 items) - Satisfaction with information sources and way information provided (3 items) 	5-point Likert scale from 1=strongly disagree to 5=strongly agree.	Post-commencing treatment (timeframe not clear)	German cancer patients from cooperating hospitals, oncology practices, pharmacies and self-aid groups. Therapy setting: inpatient treatment (n=29); outpatient clinic (n=94); primary care (n=80); unknown (n=29) <i>Sample size:</i> Pre-test: 47; Main Survey: 232

Instrument	Items/ Domains	Description of domains	Response scale	Time point/s	Sample
SCIP[45-47]	21 items, 2 subscales	<ul style="list-style-type: none"> - Satisfaction with the amount and content of information (subscale 1) (14 items) - Satisfaction with the form and timing of the information received (subscale 2) (7 items) 	<p><i>Subscale 1:</i> 4-item response scale (too much, about right, too little, none wanted).</p> <p><i>Subscale 2:</i> 5-item response scale from 1=very dissatisfied to 5=very satisfied</p>	Pre-treatment and post-treatment (one month, and 6-8 months later)	UK patients newly diagnosed with head and neck cancer, recruited from four hospitals. Surgery (27%); radiotherapy (26%); surgery and radiotherapy (31%); radiotherapy and chemotherapy (11%); and surgery, radiotherapy and chemotherapy (5%). <i>Sample size:</i> 82
CaTS [48]	25 items, 2 domains	<p>What hospital staff have done to help patients cope better before treatment in relation to:</p> <ul style="list-style-type: none"> - Sensory/psychological concerns (11 items) - Procedural concerns (14 items) 	5-point Likert-type scale from 1=strongly disagree to 5=strongly agree	Pre-chemotherapy (at least 2 days prior)	Australian cancer patients attending targeted outpatient clinics, commencing chemotherapy for lymphoma (n=52), breast (n=83) or colon cancer (n=57). <i>Sample size:</i> 192
QUOTE ^{chemo} [49]	67 items, 2 subscales, 7 domains	<ul style="list-style-type: none"> - Treatment-related information (20 items) - Prognosis information (3 items) - Rehabilitation information (11 items) - Coping information (7 items) - Interpersonal communication (6 items) - Tailored communication (10 items) - Affective communication (10 items) 	<p><i>Importance Scale:</i> 4-point Likert scale from 1=not important to 4=very important).</p> <p><i>Performance Scale:</i> 2-point response category (yes or no)</p>	Post-commencing chemotherapy (no specific timeframe)	The 60 most recent patients who had commenced chemotherapy at 10 hospitals in the Netherlands (identified through hospital medical records) were sent the survey. <i>Sample size:</i> 345

Table 3: Comprehensiveness of each instrument in relation to preparation content*

Instrument	Procedural information	Sensory information	Behavioural instruction	Psychosocial aspects	Risk information
PTIG [41]	++	++	++	++	0
RT Concerns Scale and RT Information Needs Scale [42]	++	0	+	++	++
The Wisdom Tooth Knowledge Scale [43]*	U	U	U	U	U
PS-CaTE [44]	+	0	++	++	++
SCIP [45-47]	++	+	+	++	++
CaTS [48]	++	++	++	++	++
QUOTE ^{chemo} [49]	++	0	++	++	++

*Content was not able to be assessed at the item-level as a list of items was not available to the authors.

Key:

U= Unable to be assessed

0 = Not examined

+ = One item examined content

++ = More than one item examined content

Table 4: Reliability, acceptability and feasibility of each instrument

Instrument	Internal consistency (Cronbach's alpha values unless otherwise stated)	Test-retest reliability	Acceptability	Feasibility
PTIG [41]	<i>Preoperative teaching received scale:</i> Phase 1: 0.84. Phase 2: 0.81; subscales: 0.54-0.72 <i>Preoperative teaching valued scale:</i> Phase 1: 0.88. Phase 2: 0.90; subscales: 0.59-0.84	Not Assessed	255 eligible, 18 refused. 116 participated. Time to complete: 15-20 minutes.	<u>Time for health professional to administer the instrument via interview:</u> 15-20 minutes.
RT Concerns Scale and RT Information Needs Scale[42]	<i>RT Concerns Scale:</i> Time 1: 0.91. Time 2: 0.94. Item-to-total correlations: 0.30-0.79. <i>RT Information Needs Scale:</i> 0.84; subscales: 0.73-0.79 Item-to-total correlations: 0.06-0.74 subscales: 0.30-0.74	Timeframe: Within 10 days <i>RT Concerns Scale:</i> ICC: 0.44-0.73; mean: 0.60. 2/9 items achieved 0.70. <i>RT Information Needs Scale:</i> ICC: 0.18-0.79; mean 0.55. 6/24 items achieved 0.70.	Response rate: 86%	Not Assessed
The Wisdom Tooth Knowledge Scale [43]	Not Assessed	Timeframe: 1 week. Pearson's correlation coefficient=0.73	Not Assessed	Not Assessed
PS-CaTE [44]	<i>Pre-test:</i> 0.92; subscales: 0.52-0.90 <i>Split-half reliability:</i> 0.92; subscales: 0.57-0.89 <i>Main Survey:</i> 0.95; subscales: 0.78-0.90. <i>Split-half reliability:</i> 0.91; subscales: 0.78-0.89.	Not Assessed	Response rate: 65% (for surveys distributed by the self-aid group only)	Not Assessed
SCIP [45-47]	<i>Satisfaction With the Amount and Content of Information subscale:</i> 0.89.[45] KR-20: 0.89[47] Item-to-total correlations: 0.39-0.73[47] <i>Satisfaction With the Form and Timing of the Information Received subscale:</i> 0.87[45] Item-to-total correlations: 0.52-0.74[47]	Not Assessed	Recruitment rate: 76%. Follow-up response rate (one month later): 83%; 6-8 months later: 61%. Ease of completion stated to be satisfactory. Individual item response rates: subscale 1: 79-95% (baseline); 90-93% (one month follow-up). Subscale 2: 91-98% (baseline); 93-97% (one month follow-up).	Not Assessed
CaTS [48]	<i>Sensory/psychological concerns subscale:</i> 0.96. <i>Procedural concerns subscale:</i> 0.97	Not Assessed	Not Assessed	Not Assessed
QUOTE ^{chemo} [49]	<i>Importance Scale:</i> domains: 0.76–0.90. <i>Performance Scale:</i> domains: 0.72–0.92 Item-total correlations: 0.35-0.79	Not Assessed	Response rate: 59.3%, analysis revealed no sig. differences	Not Assessed

Table 5: Validity of each instrument

Instrument	Face/content validity	Construct validity		Criterion Validity
		Factor analysis (<u>type, criteria and goodness of fit if reported</u>)	Other method	Concurrent or predictive
PTIG [41]	Abbreviated version of the PTQ, items included if rated by experts (n=9) as essential information for day surgery patients.	Not Assessed	Not Assessed	Not Assessed
RT Concerns Scale and RT Information Needs Scale [42]	Items generated from literature review, previous research, and interviews with breast cancer patients (n=34) and health professionals (n=14). 6 experts assessed appropriateness and redundancy.	Not Assessed	Not Assessed	Not Assessed
Wisdom Tooth Knowledge Scale [43]	Items generated from information leaflets. Expert panel (n=3) of oral surgeons coded each item as correct or incorrect, only items with complete agreement were included.	Not Assessed	<i>Known groups:</i> Sig. effect for groups ($P<0.001$). Dental students possessed a sig. greater knowledge level than sixth-formers and relatives ($p<0.01$) (ie. non-dental groups).	<u>Not Assessed</u>
PS-CaTE [44]	Canadian PS-CaTE was translated into German using forward/backward translation. Face and content validity of the translated version was assessed by pharmacy and medical practice experts and patients (n=47).	Not Assessed	Not Assessed	Not Assessed
SCIP [45-47]	Items derived from previously published SIMS instrument. Additional items were generated from qualitative interviews with patients.	Subscale 1: PCA with oblique rotation; eigen values >1 , scree plot; 3 factor solution. Factor loadings: 0.21-0.89. Explained variance: Factor 1=40.9%, factor 2=10.6% and factor 3=8.1%. Subscale 2: PCA with oblique rotation and scree	<i>Convergent validity:</i> [45] Satisfaction with content and amount of information with: IPQ-R: Symptoms (illness identity)=-0.32; IPQ-R: Personal control =0.29; IPQ-R: Illness coherence=0.30. Satisfaction with form and timing of information with: IPQ-R: personal control= 0.28. <i>Discriminant validity:</i> [45] Satisfaction with information not affected by levels of optimism ($r= 0.1-0.10$)	<i>Predictive validity (one month post-treatment)</i> [45] Satisfaction with content and amount of information with: Global QoL= 0.38; HADS-D= 0.42; IPQ-R: Consequences= -0.45; Treatment control= 0.40; Symptoms (illness identity) beliefs= -0.30; Satisfaction with form and timing of information with: necessity of

Instrument	Face/content validity	Construct validity		Criterion Validity
		Factor analysis (type, criteria and goodness of fit if reported)	Other method	Concurrent or predictive
		plot; 1 factor solution. Factor loadings: 0.56-0.87. Explained variance: 61%.	<i>Ability to discriminate between different levels of individual satisfaction:[47]</i> Subscale 1: Discrimination: Ferguson's Delta=0.92. Item discriminations: 0.44- 0.89. Item delta coefficients: 0.60-0.98. Subscale 2: Discrimination: Ferguson's Delta=0.90. Item delta coefficients: 0.71-0.93.	treatment (BMQ-specific necessity subscale) = 0.40; IPQ-R: Treatment control= 0.48. <i>Predictive validity (6-8 months post-treatment)[46]:</i> Satisfaction with content and amount of information with: HADS-D= -0.34; HrQoL (MCS) = 0.32. Anxiety and PCS were not associated with satisfaction with information.
CaTS [48]	Items generated by psychosocial/clinical experts (n=3) and literature review. Items pilot tested with cancer patients (n=10).	Non-rotated PFA; parallel analysis to confirm scree test; PFA runs with oblique rotation determined final two-factor composition of scale. Factor loadings: 0.63-0.95. Explained variance: Factor 1=59.1%, factor 2=8.7%.	<i>Divergent validity:</i> Sensory/Psychological subscale with: HADS-T= 0.24 HADS-A=0.26. Procedural concerns subscale with: HADS-T= 0.13 HADS-A= 0.15 <i>Known groups:</i> Younger participants (under 65 years of age) had significantly greater procedural concerns (p=0.001; medium effect).	Not Assessed
QUOTE ^{chemo} [49]	Items generated from literature review and focus group discussions (n=5) with cancer patients. Content validity of the categorisation was tested by coders (n=10).	CFA using SEM NF, CFI, RMSEA and AIC indicated good fit to the data. Factor loadings: 0.43-0.77.	<i>Convergent validity:</i> Importance Scale domains with: Information preferences: ISQ= -0.26 to -0.9. Monitoring style: TMSI=0.09 to 0.26 Intrusion: IES subscale=0.13 to 0.25 Avoidance: IES subscale=0.07 to 0.16	Not Assessed

AIC= Aikake Information Criterion; BMQ= Beliefs about Medicine Questionnaire; CFA=Confirmatory Factor Analysis; CFI= Comparative Fit Index; HADS-A=Hospital Anxiety and Depression Scale (anxiety subscale); HADS-D=HADS (depression subscale); HADS-T=HADS (total scale); HrQoL= Health-related Quality of Life; IPQ-R=Illness Perception Questionnaire-Revised; Impact of Event Scale-Dutch Version (IES); ISQ=Information Satisfaction Questionnaire; MCS=Mental Component Summary; NF=Normed Fit Index; Principal Components Analysis (PCA); PFA= Principal Factors Extraction; PTQ=Providence Portland Medical Center, ORE, Preoperative Teaching Questionnaire; QoL=Quality of Life; RMSEA=Root Mean Square of Error Approximation; SEM=Structural Equation Modelling; SIMS=Satisfaction with Information about Medicines Scale; and TMSI=Threatening Medical Situation Inventory.

Table 6: Summary of the psychometric properties of each instrument*

Instrument	Internal consistency	Test-retest reliability	Face/content validity	Construct validity			Criterion Validity		Acceptability	Feasibility
				Factor analysis	Convergent/divergent	Other methods	Concurrent	Predictive		
PTIG [41]	+	0	+	0	0	0	0	0	+	+
RT Concerns Scale and RT Information Needs Scale [42]	++	+	++	0	0	0	0	0	++	0
The Wisdom Tooth Knowledge Scale [43]	0	++	+	0	0	++	0	0	0	0
PS-CaTE [44]	++	0	++	0	0	0	0	0	+	0
SCIP [45-47]	++	0	+	++	+	++	+	++	++	0
CaTS [48]	+	0	++	++	++	++	0	0	0	0
QUOTE ^{chemo} [49]	++	0	++	++	+	0	0	0	++	0

*Measurement properties for responsiveness and cross-cultural validity not summarised

Key:

0 = Not assessed

+ = Weak evidence (some results did not meet predetermined criteria)

++ = Good evidence (all results met predetermined criteria)