BMJ Open is committed to open peer review. As part of this commitment we make the peer review history of every article we publish publicly available.

When an article is published we post the peer reviewers’ comments and the authors’ responses online. We also post the versions of the paper that were used during peer review. These are the versions that the peer review comments apply to.

The versions of the paper that follow are the versions that were submitted during the peer review process. They are not the versions of record or the final published versions. They should not be cited or distributed as the published version of this manuscript.

BMJ Open is an open access journal and the full, final, typeset and author-corrected version of record of the manuscript is available on our site with no access controls, subscription charges or pay-per-view fees (http://bmjopen.bmj.com).

If you have any questions on BMJ Open’s open peer review process please email info.bmjopen@bmj.com
### Psychometric properties of self-reported financial toxicity measures in cancer survivors: A systematic review

| Journal          | BMJ Open                   |
|------------------|----------------------------|
| Manuscript ID    | bmjopen-2021-057215        |
| Article Type     | Original research          |
| Date Submitted by the Author | 13-Sep-2021               |
| Complete List of Authors | Zhu, Zheng; Fudan University School of Nursing, Xing, Weijie; Fudan University School of Nursing, Wen, Huan; Fudan University, Sun, Yanling; Fudan University, So, WK; The Chinese University of Hong Kong, Lizarondo, Lucylynn; University of Adelaide, The Joanna Briggs Institute, Peng, Jian; Fudan University School of Nursing, Hu, Yan; Fudan University, School of Nursing |
| Keywords         | ONCOLOGY, HEALTH ECONOMICS, Health economics < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Quality in health care < HEALTH SERVICES ADMINISTRATION & MANAGEMENT |

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml
I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd (“BMJ”) its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge (“APC”) for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author’s Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.
Psychometric properties of self-reported financial toxicity measures in cancer survivors: A systematic review

Zheng Zhu¹,², Weijie Xing¹,², Huan Wen³, Yanling Sun³, Winnie K. W. So⁴, Lizarondo Lucylynn⁵; Jian Peng¹; Yan Hu¹,²

Author affiliations:

1. School of Nursing, Fudan University, Shanghai, China (Zheng Zhu, Weijie Xing, Jian Peng, Yan Hu);
2. Fudan University Centre for Evidence-based Nursing: A Joanna Briggs Institute Centre of Excellence, Fudan University, Shanghai, China (Zheng Zhu, Weijie Xing, Yan Hu);
3. School of Public Health, Fudan University, Shanghai, China (Huan Wen, Yanling Sun);
4. The Nethersole School of Nursing, The Chinese University of Hong Kong, Hong Kong, China (Winnie K. W. So);
5. Joanna Briggs Institute, University of Adelaide, Adelaide, South Australia, Australia (Lizarondo Lucylynn).

Correspondence:

Weijie Xing, RN, PhD, School of Nursing, Fudan University; Fudan University Centre for Evidence-Based Nursing: A Joanna Briggs Institute Centre of Excellence, 305 Fenglin Rd., Shanghai, China (xingweijie@fudan.edu.cn)
Word count: 2920

Number of tables: 4

Number of figures: 1

Number of references: 57
Psychometric properties of self-reported financial toxicity measures in cancer survivors: A systematic review

Abstract

Objective: The aim of this systematic review was to summarize the psychometric properties of patient-reported outcome measures (PROMs) measuring financial toxicity (FT) in cancer survivors.

Design: Comprehensive searches were performed in seven databases. Quality appraisal, data extraction, data synthesis, and assessment of the certainty of evidence were conducted according to the guidance of the Consensus-Based Standards for the Selection of Health Measurement Instruments (COSMIN) methodology.

Results: A total of 9399 articles were identified via databases. Finally, a total of 24 articles (21 studies, 20 PROMs) were eligible for inclusion in this study. The findings highlighted that the COST had an adequate development process and showed better psychometric properties than other PROMs, especially in internal consistency, reliability, and hypothesis testing.

Conclusions: The COST could be recommended as the most suitable measures for use in research and clinical practice. The FIT and the IOC-CS financial problems domain also had adequate psychometric properties. We suggest that PROMs should
be selected only after careful consideration of the local socioeconomic context. Future studies are warranted to develop various FT PROMs based on different social and cultural backgrounds and to clarify the theoretical grounds for assessing FT.

Keywords: cancer, financial toxicity, measurement, PROM, systematic review
57  **Strength and limitation of this study**

58  ● This is the first systematic review that comprehensively summarized the

59  psychometric properties of 20 PROMs evaluating financial toxicity in cancer

60  survivors.

61  ● The results may provide quantitative evidence for researchers and healthcare

62  professionals to choose PROMs measuring cancer survivors’ financial toxicity in

63  future scientific research and clinical practice.

64  ● This review only included studies that aimed to evaluate the measurement

65  properties of financial toxicity PROMs.
Introduction

The rising cost associated with advancements in cancer treatment and lengthening of cancer survivorship poses a significant challenge to survivors, caregivers, and public healthcare systems.\textsuperscript{1,2} Total global spending on cancer medications grows at a compound annual growth rate of 6.5\% from US$ 96 billion in 2013 to US$ 173 billion in 2020, which is nearly twice the rate of global gross domestic product (GDP) growth.\textsuperscript{3-5} The majority of cancer survivors in middle-income and low-income countries/regions depend on out-of-pocket payments, which may lead to global inequalities in healthcare expenditures and financial insecurity for vulnerable groups.\textsuperscript{6,7}

The term “financial toxicity (FT)” has been described as the economic effect of cancer treatment in the age of precision medicine.\textsuperscript{2,8,9} Witte and colleagues described FT as “the patient-reported outcome (PRO) of perceived subjective financial distress resulting from objective financial burden”.\textsuperscript{10} This concept covers both the objective financial burden and the subjective financial distress that cancer survivors face as a result of high out-of-pocket medical expenses. Regarding the terminology, “financial toxicity”, “financial burden” and “financial distress” are often used interchangeably in research and share a similar definition.\textsuperscript{10,11} In this review, the authors agreed to consistently use the term “financial toxicity.” Financial toxicity is usually measured by patient-reported outcome measures (PROMs); choosing a PROM with high validity and reliability is a prerequisite for robust results.
There are a few cancer-specific and generic FT PROMs that have been reported and used in different contexts. As one of the recent cancer-specific FT PROMs, the Comprehensive Score for Financial Toxicity (COST) is the most commonly used measure for assessing FT. In addition to COST, other cancer-specific measures have been widely used, including the Breast Cancer Finances Survey Inventory (BCFS), Socioeconomic Wellbeing Scale (SWBS), and InCharge Financial Distress/Financial Wellbeing Scale (InCharge). Additionally, validated subscales, such as the Social Difficulties Inventory Cancer Care Outcomes (SDI) and the Cancer Care Outcomes Research and Surveillance Consortium patient survey (CanCORS), were also used to evaluate FT. However, existing PROMs vary significantly in their state of development and degree of validation, and many PROMs have not been psychometrically tested.

A preliminary literature search was conducted in PubMed, PsycINFO (EBSCO), Cochrane Library (Wiley) and JBI (Ovid), which revealed that there exist some reviews regarding measures of FT. Witte and colleagues summarized the content of 352 items from 34 studies measuring FT in cancer survivors. However, this review did not report the psychometric properties of the included PROMs, and most of the included PROMs were not validated through a scientific process, which made it difficult for readers to choose the best measure from existing PROMs to evaluate the level of FT. Salman and colleagues conducted a COSMIN systematic review and
found 8 PROMs and 2 caregiver-reported measures for assessing financial burden in adolescents and young adults.\textsuperscript{18} However, this review focused only on PROMs assessing FT in adolescents and young adults with cancer. The psychological properties of FT measures in adult cancer survivors are still unknown.

The reproducibility, reliability, and accuracy of PROMs are the fundamental premise for achieving robust results.\textsuperscript{22,23} Therefore, it is necessary to summarize the psychometric properties of existing PROMs for future research. However, this information is still lacking. The aim of this systematic review was to summarize the psychometric properties of PROMs for measuring FT in cancer survivors. The review was conducted according to the guidance of the Consensus-Based Standards for the Selection of Health Measurement Instruments (COSMIN) methodology and the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement.\textsuperscript{19,20} The protocol of this review was published in \textit{BMJ OPEN} in 2020.\textsuperscript{21} The registration number of protocol in PROSPERO was CRD42021254721.

\textbf{Methods}

\textbf{Search strategy}

First, we conducted a limited search via PubMed to capture keywords from which to develop search strategies for each database. Subsequently, all identified search strategies across databases were performed in PubMed, MEDLINE (Ovid), Embase (Ovid), CINAHL (EBSCO), Web of Science, ProQuest Dissertations and
Theses, and Cochrane Library (Wiley). In PubMed/Medline, we searched papers in English using MeSH terms ([cancer OR neoplasms] AND [“cancer survivors” OR patient OR survivors] AND “cost of illness”) combined with (cancer OR [patient* OR survivor*] AND [cost OR bill* OR expense OR productivity loss OR “out-of-pocket” OR “economic burden” OR “financial toxicity” OR “financial hardship” OR “financial burden”]). The COSMIN measurement properties filter and exclusion filter were also utilized in the search box. The search strategies for each database are presented in Appendix 1. Finally, the references of all included studies were manually reviewed to supplement the database search.

Inclusion and exclusion criteria

The inclusion criteria were as follows: 1) studies that reported any PROM for measuring FT in cancer survivors who were ≥18 yrs old; 2) studies that evaluated at least one measurement property; and 3) studies published in English. The exclusion criteria were as follows: 1) studies that were not validation studies and used a PROM only as an outcome measurement; 2) studies that used a PROM as a comparator for another instrument; 3) studies that did not provide empirical data; additionally, 4) if a measure was a quality of life PROM and had a domain that assessed FT, we included only the original version of the PROM.

Study screening and selection

We imported all identified citations by search strategies into Endnote X8
(Clarivate Analytics, PA, USA). After duplicates were removed, two reviewers (ZZ & WX) independently screened all titles, abstracts, and full texts (ZZ & WX) based on the established inclusion and exclusion criteria. Any disagreements were resolved by a third reviewer (YH).

**Quality appraisal**

Two reviewers (WH & YS) assessed the methodological quality of the PROM of the included studies by using the COSMIN Risk of Bias Checklist. The checklist consisted of 10 domains (38 items), including PROM development, content validity, structural validity, internal consistency, cross-cultural validity, reliability, measurement error, criterion validity, hypothesis testing, and responsiveness. Each measurement property was rated as “very good”, “adequate”, “doubtful” or “inadequate quality”. The methodological quality of a single study is rated based on the worst score counts method. Any discrepancies were resolved by a third reviewer (ZZ).

**Data extraction**

Two reviewers (ZZ & WX) independently extracted data from the included papers, including authors, year of publication, PROM, country/language, study design, target population, sample size, domains, number of items, total score range, and main findings. The main findings regarding psychometric properties including content validity, structural validity, internal consistency, cross-cultural validity,
reliability, measurement error, criterion validity, hypothesis testing, and responsiveness, were also extracted. Any discrepancies were resolved through discussion between the two reviewers.

**Data synthesis**

We used the COSMIN criteria to summarize and evaluate the psychometric properties of each study regarding structural validity, internal consistency, reliability, measurement error, hypothesis testing for construct validity, cross-cultural validity/measurement invariance, criterion validity, and responsiveness. Each measurement property from each study was rated as sufficient (+), insufficient (−), or indeterminate (?). If the ratings of one psychometric property per study were all sufficient or insufficient, the results were pooled, and the overall rating was rated as sufficient or insufficient. If the ratings were inconsistent, explanations of inconsistency were explored (e.g., different languages). If the explanation was reasonable, we provided ratings in subgroups. If the explanation was not reasonable, the overall rating of this measurement property was rated as inconsistent (±). If there was no information supporting the rating, the overall rating was indeterminate (?).

**Assessing certainty of evidence**

We used a modified Grading of Recommendations Assessment, Development, and Evaluation (GRADE) system to assess the certainty of evidence. Each piece of evidence was graded for risk of bias, inconsistency, imprecision, and indirectness.
Four reviewers (ZZ, WJ, HW, and YS) independently assessed the grade. Any discrepancies were resolved by discussion.

Results

Literature search

Figure 1 shows the process of literature screening and selection. A total of 9399 articles were identified via databases. Six articles were found by additional supplementary searches. After duplications were removed, a total of 9405 articles were retained, 9348 articles were deleted after reading the title and abstract, and 58 were deleted after full-text reading. Finally, a total of 24 articles (21 studies, 20 PROMs) were eligible for inclusion in this study.\textsuperscript{12,14,16,17,22-41}

Study description

Table 1 shows the characteristics of the included studies. All included studies were published from 1988 to 2021. Seven studies were conducted in the United States,\textsuperscript{12,17,24,27,37,41} four in the United Kingdom,\textsuperscript{16,29,35,38} two in Canada,\textsuperscript{31,36} and two in China (mainland and Hong Kong).\textsuperscript{25,39} One study was conducted in 12 counties in Europe and North America.\textsuperscript{22,23} A total of 12412 participants were included, ranging from 7\textsuperscript{36} to 5901\textsuperscript{41} per study. The majority of studies assessed FT in multiple types of cancer. Only three studies focused on a single type of cancer, namely, lung, colorectal, or head and neck cancer.\textsuperscript{22,23,31,37}
Among the 20 PROMs, 9 were FT-related domains of quality of life PROMs, and 11 were independent PROMs focusing on FT. All PROMs were validated in cancer survivors. Thirteen PROMs were in English, and two were in Chinese. Other languages included French, Portuguese, Italian, and Persian. The number of items evaluating FT ranged from 1 to 23. Cancer Care Outcomes Research and Surveillance (CanCORS) did not report the items or total score range of the FT-related domain. The French version of the Patient Self-Administered Financial Effects questionnaire (P-SAFE) did not report the total score range of the whole PROM.

Quality assessment

Methodological quality assessment. Table 2 shows the methodological quality of the 21 included studies by using the COSMIN checklist. In the PROM development domain, only three studies were rated as doubt, and the others were rated as inadequate. Two studies reported adequate information in testing the relevance, comprehensiveness, and comprehensibility of PROMs. Among all studies, the most reported domain was internal consistency, except three studies. Limited information could be retrieved on cross-cultural validity (three studies), criterion validity (six studies), reliability (nine studies), and responsiveness (two studies). No data were identified on measurement error.
**Measurement property assessment.** Table 3 shows the quality of the psychometric properties retrieved from 20 PROMs. Only the Persian version of the COST-v2 was rated as “+” in structural validity. There were 16 PROMs rated as “+” in internal consistency. Six PROMs were rated as “+” in reliability. Limited information was retrieved on cross-cultural validity (two PROMs), criterion validity (six PROMs), and responsiveness (two PROMs). No PROMs reported data on measurement error.

**Certainty of evidence**

Table 4 shows the certainty of evidence for each measurement property. Among all included PROMs, COST showed the best psychometric properties compared to other measures. COST and its seven versions were rated as high evidence in structural validity, internal consistency, hypothesis testing, and criterion validity. The Financial Index of Toxicity (FIT) and Impact of Cancer – Childhood Survivors IOC-CS financial problems domain reported data on five properties and were rated as very low evidence to high evidence. The European Organisation for Research and Treatment of Cancer (EORTC) Core Quality of Life Questionnaire (QLQ-C30) financial impact domain and CanCORS insurance coverage and income domain did not provide any qualified data on all properties.

**Discussion**
This systematic review identified 20 PROMs and domains of PROMs evaluating FT in cancer survivors, including the COST (original, Brazilian, India, Italian, Persian, Simplified Chinese, Traditional Chinese version), FIT, personal financial burden (PFB), P-SAFE, SWBS, EORTC QLQ-C30 financial impact domain, Quality of Life in Adult Cancer Survivors (QLACS) financial problems domain, Chronic Cancer Experiences Questionnaire (CCEQ) financial advice domain, CanCORS insurance coverage and income domain, Patient Roles and Responsibilities Scale (PRRS) financial well-being domain, SDI-21 providing for the family domain, SDI-16 money matters domain, IOC-CS financial problems domain, and Cancer Problems in Living Scale (CPILS) employment/financial domain. Overall, the COST had a complete development process compared to other PROMs and showed the best psychometric properties, especially in terms of internal consistency, reliability, and hypothesis testing. To the best of our knowledge, this is the first systematic review that has summarized the psychometric properties of FT PROMs in cancer survivors and reported the certainty of evidence for each property of PROMs. The results may provide quantitative evidence for researchers and healthcare professionals to choose PROMs measuring cancer survivors’ FT in future scientific research and clinical practice.

The results highlighted that the COST (of which we studied both version 1 and version 2) had better psychometric properties than other specific and generic PROMs in terms of internal consistency, reliability, and hypothesis testing. The COST could
be recommended as a suitable PROM in research and clinical practice. Other
systematic reviews have also suggested that the COST is a promising measure from a
content perspective\textsuperscript{10,11}. From a psychometric standpoint, there are a few issues that
one must face when evaluating financial toxicity in cancer survivors using the COST.
First, caution should be taken when using the COST in different socioeconomic
conditions outside the United States. In some countries in Europe or Asia, the
majority of medical expenses are covered by social health insurance, and direct
out-of-pocket payments are replaced by prepayment from health insurance
contributions.\textsuperscript{42,43} In addition, social security systems can also benefit cancer
survivors who are not able to work.\textsuperscript{44} These two socioeconomic factors may affect
cancer survivors’ understanding regarding some items related to medical spends and
indirect cost. However, few COST validation studies have considered socioeconomic
issues, adapted the measure in a local context, or provided data on cross-cultural
validity. It is recommended that future COST validation studies recruit cancer
survivors across multiple social and cultural backgrounds to assess cross-cultural
measurement invariance.

Second, the original construct and item generation for the COST were based on a
literature search; thus, the theoretical grounds for the measure are unclear, and the
instrument may not capture detailed information related to the construct. Theoretical
frameworks and conceptual models are crucial for self-reported measures to capture
subtle changes in constructs.\textsuperscript{45} Although FT is a relatively new concept, certain
models can guide item generation in the development of future FT PROMs.

Tucker-Seeley and colleagues developed a conceptual model of FT and emphasized three components of financial burden, namely, the material, psychosocial, and behavioral domains. Head developed SWBS based on James Coleman’s Theory of Social Class; this scale contains 17 items across 3 domains: including human capital, material capital, and social capital. Witte and colleagues’ systematic review analyzed 352 different questions regarding financial spending and found six domains (financial spending, financial resources, psychosocial affect, support seeking, coping care, and coping lifestyle) that can represent reactions to subjective financial distress. Other theories and models, including the Wreckers theory of financial distress, ecological theory, and the functionalist tradition, have also been widely used in cancer survivors. With the increasing number of theoretical studies related to FT, the theoretical grounds for future PROMs need to be clarified.

In addition to the COST, two other PROMs, namely the FIT and the IOC-CS financial problems domain, also provided adequate data on psychometric properties. FIT is relatively new and has fewer items than other included measures. This measure was developed by Hueniken and colleagues and has been validated only in survivors with head and neck cancer. Head and neck cancer, especially laryngeal and hypopharyngeal cancer, has particularly large impacts on survivors’ daily function (e.g., speech and eating) after treatment and affects survivors’ ability to return to work. Only 32% to 59% of head and neck cancer survivors return to work after...
This form of cancer also has short- and long-term financial consequences for caregivers and their families. Therefore, future studies should be aware that the FIT may not be directly applicable to other cancer populations.

Regarding PROM development, we found that all original versions of FT PROMs were established in developed countries. Previous studies have reported that FT is closely related to broad social determinants of economic circumstances, such as healthcare policy and the level of regional economic development. Additionally, cultural factors (e.g., a cultural emphasis on saving and a cultural imperative to have a large family) also affect cancer survivors’ perceived financial security and economic burden. Therefore, further studies are warranted to develop various FT PROMs based on different social and cultural backgrounds.

Limitations

We acknowledge that there are some limitations to this study. First, this review included only studies that aimed to evaluate the measurement properties of FT PROMs. Many studies that aimed to explore the level of FT in cancer survivors also reported the reliability and validity of PROMs. Therefore, the PROMs we summarized in this systematic review had higher psychometric quality than other measures that we did not list in this review. Second, we included only studies published in English. Therefore, studies published in other languages were not included, which may affect the conclusion of this review. Third, we included only the
original version of the FT domain from PROMs assessing quality of life in cancer survivors, such as the EORTC QLQ-C30 and the QLACS. Over 20 language versions of these PROMs do not provide sufficient details on the FT domain individually.

Conclusion

This systematic review summarized the psychometric properties of 20 PROMs evaluating FT in cancer survivors. The findings highlighted that the COST had an adequate PROM development process and showed the best psychometric properties among all examined PROMs, especially in internal consistency, reliability, and hypothesis testing; thus, we recommend the COST as the most suitable measure for use in research and clinical practice. The FIT and the IOC-CS financial problems domain also had adequate psychometric properties. We suggest that PROMs should be selected only after careful consideration of the local socioeconomic context. Future studies are warranted to develop various FT PROMs based on different social and cultural backgrounds and a clear theoretical basis for assessing FT.

Acknowledgements

Contributors

ZZ and WX designed the systematic review, conducted data searching, extraction and analysis, assessing the certainty of evidence, and wrote the draft of the manuscript. HW and YS conducted quality appraisal and assessing the certainty of evidence. WS, LL, JP, and YH provided critical comments. All authors approved the
final version of the manuscript.

Funding

This work was supported by the Shanghai Pujiang Program (Grant Number: 2019PJC017), Shanghai Soft Science Key Program (Grant Number: 20692104800), National Natural Science Foundation of China (Grant Number: 72004034), China Medical Board Open Competition Program (Grant Number: 20-371).

Competing interests

None.

Patient and public involvement

Patients and the public were not involved in the design, or conduct, or reporting, or dissemination plans of our research.

Patient consent for publication

Not applicable.

Ethics approval

This systematic review did not require ethical approval.

Data available statement
No additional data are available.

Reference

1. Bosland MC. Global health and cancer. *Lancet*. 2019;393(10175):983.

2. De Souza JA, Wong YN. Financial distress in cancer patients. *J Med Person*. 2013;11:73-77.

3. World cancer report [Internet]. [cited 2021 July 31]. 2020. Available from: https://publications.iarc.fr/586.

4. Global Oncology Trends 2021[Internet]. [cited 2021 July 31]. 2021. Available from: https://www.iqvia.com/insights/the-iqvia-institute/reports/global-oncology-trends-2021

5. Mariotto AB, Yabroff KR, Shao Y, et al. Projections of the cost of cancer care in the United States: 2010-2020. *J Natl Cancer Inst*. 2011;103:117-128.

6. ACTION Study Group. Financial catastrophe, treatment discontinuation and death associated with surgically operable cancer in South-East Asia: Results from the ACTION Study. *Surgery*. 2015;157:971-982.

7. Mackintosh M, Channon A, Karan A, et al. What is the private sector? Understanding private provision in the health systems of low-income and middle-income countries. *Lancet*. 2016;388:596-605.

8. Yousuf Zafar S. Financial toxicity of cancer care: It’s time to intervene. *J Natl Cancer Inst*. 2015;108:djv370.
9. Carrera PM, Kantarjian HM, Blinder VS. The financial burden and distress of patients with cancer: understanding and stepping-up action on the financial toxicity of cancer treatment. *CA Cancer J Clin.* 2018;68:153-165

10. Witte J, Mehlis K, Surmann B, et al. Methods for measuring financial toxicity after cancer diagnosis and treatment: a systematic review and its implications. *Ann Oncol.* 2019;30:1061-1070.

11. Gordon LG, Merollini KM, Lowe A, Chan RJ. A systematic review of financial toxicity among cancer survivors: we can’t pay the co-pay. *Patient.* 2017;10:295-309.

12. De Souza JA, Yap BJ, Hlubocky FJ et al. The development of a financial toxicity patient-reported outcome in cancer: the COST measure. *Cancer.* 2014;120:3245-3253.

13. Given BA, Given CW, Stommel M. Family and out-of-pocket costs for women with breast cancer. *Cancer Pract.* 1994;2:187-193.

14. Head BA, Faul AC. Development and validation of a scale to measure socioeconomic well-being in persons with cancer. *J Support Oncol.* 2008;6(4):183-192.

15. Prawitz AD, Garman ET, Sorhaindo B et al. InCharge financial distress/ financial well-being scale: development, administration, and score interpretation. *Financial Counsel Plann.* 2006;17:34–50.

16. Wright P, Smith AB, Keding A, Velikova G. The social difficulties inventory (SDI): development of subscales and scoring guidance for staff. *Psychooncology.*
17. Malin JL, Ko C, Ayanian JZ et al. Understanding cancer patients’ experience and outcomes: development and pilot study of the Cancer Care Outcomes Research and Surveillance patient survey. *Support Care Cancer*. 2006;14:837-848.

18. Salsman JM, Danhauer SC, Moore JB, et al. Systematic review of financial burden assessment in cancer: Evaluation of measures and utility among adolescents and young adults and caregivers. *Cancer*. 2021;127(11):1739-1748.

19. Prinsen CAC, Mokkink LB, Bouter LM, Alonso J, Patrick DL, de Vet HCW, Terwee CB. COSMIN guideline for systematic reviews of patient-reported outcome measures. *Qual Life Res*. 2018;27(5):1147-1157.

20. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*. 2021;372:n71.

21. Zhu Z, Xing W, Lizarondo L, Peng J, Hu Y, So WK. Psychometric properties of self-reported financial toxicity measures in cancer survivors: a systematic review protocol using COSMIN methodology. *BMJ open*. 2020;10(5):e036365.

22. Aaronson NK, Bullinger M, Ahmedzai S. A modular approach to quality-of-life assessment in cancer clinical trials. InCancer Clinical Trials 1988 (pp. 231-249). Springer, Berlin, Heidelberg.

23. Aaronson NK, Ahmedzai S, Bergman B, et al. The European Organization for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. *J Natl Cancer Inst*. 1993;85(5):365-376.
24. Avis NE, Smith KW, McGraw S, Smith RG, Petronis VM, Carver CS. Assessing quality of life in adult cancer survivors (QLACS). *Qual Life Res.* 2005;14(4):1007-1023.

25. Chan DN, Choi KC, Ng MS, et al. Translation and validation of the Traditional Chinese version of the COmprehensive Score for financial Toxicity-Functional Assessment of Chronic Illness Therapy (Version 2). *Health Qual Life Outcomes.* 2021;19(1):17.

26. Dar MA, Chauhan R, Sharma KK, Trivedi V, Dhingra S, Murti K. Assessing the reliability and validity of comprehensive score for financial toxicity (COST) among radiation oncology patients in India: a cross-sectional pilot study. *Ecanermedicscience.* 2021;15:1219.

27. De Souza JA, Yap BJ, Wroblewski K, et al. Measuring financial toxicity as a clinically relevant patient-reported outcome: the validation of the COmprehensive Score for financial Toxicity (COST). *Cancer.* 2017;123(3):476-484.

28. Durber K, Halkett GK, McMullen M, Nowak AK. Measuring financial toxicity in Australian cancer patients–Validation of the COmprehensive Score for financial Toxicity (FACT COST) measuring financial toxicity in Australian cancer patients. *Asia Pac J Clin Oncol.* 2021;17(4):377-387.

29. Harley C, Pini S, Kenyon L, Daffu-O'Reilly A, Velikova G. Evaluating the experiences and support needs of people living with chronic cancer: development and initial validation of the Chronic Cancer Experiences Questionnaire (CCEQ). *BMJ Support Palliat Care.* 2019;9(1):e15.
30. Head BA. Development and validation of a scale to measure health-related socioeconomic well-being in persons with a cancer diagnosis. Electronic Theses and Dissertations. Paper 594. https://doi.org/10.18297/etd/594

31. Hueniken K, Douglas CM, Jethwa AR, et al. Measuring financial toxicity incurred after treatment of head and neck cancer: Development and validation of the Financial Index of Toxicity questionnaire. *Cancer*. 2020;126(17):4042-4050.

32. de Alcantara Nogueira L, Koller FJ, Marcondes L, et al. Validation of the comprehensive score for financial toxicity for Brazilian culture. *Ecancermedicalscience*. 2020;14:1158.

33. Ripamonti CI, Chiesi F, Di Pede P, et al. The validation of the Italian version of the COmprehensive Score for financial Toxicity (COST). *Supportive Care Cancer*. 2020;28(9):4477-4485.

34. Sharif SP, Sim OF, Moshtagh M, Naghavi N, Nia HS. Psychometric evaluation of the COmprehensive Score for financial Toxicity scale among Iranian cancer patients. *Research Square*. 2020. https://doi.org/10.21203/rs.2.22672/v1

35. Shilling V, Starkings R, Jenkins V, Cella D, Fallowfield L. Development and validation of the patient roles and responsibilities scale in cancer patients. *Qual Life Res*. 2018;27(11):2923-2934.

36. Tremblay D, Poder TG, Vasiliadis HM, Touati N, Fortin B, Lévesque L, Longo C. Translation and Cultural Adaptation of the Patient Self-Administered Financial Effects (P-SAFE) Questionnaire to Assess the Financial Burden of Cancer in French-Speaking Patients. *Healthcare (Basel)*. 2020;8(4):366.
37. Veenstra CM, Regenbogen SE, Hawley ST, et al. A composite measure of personal financial burden among patients with stage III colorectal cancer. *Med Care*. 2014;52(11):957-962.

38. Wright EP, Kiely M, Johnston C, Smith AB, Cull A, Selby PJ. Development and evaluation of an instrument to assess social difficulties in routine oncology practice. *Qual Life Res*. 2005;14(2):373-386.

39. Yu HH, Yu ZF, Li H, Zhao H, Sun JM, Liu YY. The COPRehensive Score for financial Toxicity in China: Validation and Responsiveness. *J Pain Symptom Manage*. 2021;61(6):1297-1304.e1.

40. Zebrack BJ, Donohue JE, Gurney JG, Chesler MA, Bhatia S, Landier W. Psychometric evaluation of the impact of cancer (IOC-CS) scale for young adult survivors of childhood cancer. *Qual Life Res*. 2010;19(2):207-218.

41. Zhao L, Portier K, Stein K, Baker F, Smith T. Exploratory factor analysis of the Cancer Problems in Living Scale: a report from the American Cancer Society's Studies of Cancer Survivors. *J Pain Symptom Manage*. 2009;37(4):676-686.

42. De Lorenzo F, Apostolidis K, Florindi F, Makaroff LE. Improving European policy to support cancer survivors. *J Cancer Policy*. 2018;15:72-75.

43. Bhoo-Pathy N, Ng CW, Lim GC, et al. Financial toxicity after cancer in a setting with universal health coverage: A call for urgent action. *J Oncol Pract*. 2019;15(6):e537-e546.

44. Philip CC, Mathew A, John MJ. Cancer care: Challenges in the developing world. *Cancer Res Stat Treat*. 2018;1(1):58.
45. Prinsen CAC, Mokkink LB, Bouter LM, et al. COSMIN guideline for systematic reviews of patient-reported outcome measures. *Qual Life Res*. 2018;27(5):1147-1157.

46. Tucker-Seeley RD, Thorpe Jr RJ. Material–psychosocial–behavioral aspects of financial hardship: a conceptual model for cancer prevention. *Gerontologist*. 2019;59(Suppl 1):S88-S93.

47. Coleman JS. Foundations of social theory. Harvard university press; 1994.

48. A ‘Wreckers Theory’ of Financial Distress. [Internet]. [cited 2021 July 31]. 2005. Available from: https://www.econstor.eu/bitstream/10419/19625/1/200540dkp.pdf

49. Bergeron K, Abdi S, DeCorby K, Mensah G, Rempel B, Manson H. Theories, models and frameworks used in capacity building interventions relevant to public health: a systematic review. *BMC Public Health*. 2017;17(1):914.

50. Acker J, Baca-Zinn M, Collins PH, et al. Social class and stratification: Classic statements and theoretical debates. Rowman & Littlefield Publishers; 2006.

51. Miller A. Returning to work after head and neck cancer. *Curr Opin Otolaryngol Head Neck Surg*. 2020;28(3):155-160.

52. Koch R, Wittekindt C, Altendorf–Hofmann A, Singer S, Guntinas–Lichius O. Employment pathways and work-related issues in head and neck cancer survivors. *Head Neck*. 2015;37(4):585-593.

53. Giuliani M, Papadakos J, Broadhurst M, et al. The prevalence and determinants of return to work in head and neck cancer survivors. *Support Care Cancer*. 
551 2019;27(2):539-546.

552 54. Dewa CS, Trojanowski L, Tamminga SJ, Ringash J, McQuestion M, Hoch JS. Work-related experiences of head and neck cancer survivors: an exploratory and descriptive qualitative study. *Disabil Rehabil.* 2018;40(11):1252-1258.

555 55. Zhu Z, Xing W, Zhang X, Hu Y, So WK. Cancer survivors' experiences with financial toxicity: A systematic review and meta-synthesis of qualitative studies. *Psychooncology.* 2020;29(6):945-959.

558 56. Chebli P, Lemus J, Avila C, et al. Multilevel determinants of financial toxicity in breast cancer care: perspectives of healthcare professionals and Latina survivors. *Support Care Cancer.* 2020;28(7):3179-3188.

561 57. Kong YC, Wong LP, Ng CW, et al. Understanding the financial needs following diagnosis of breast cancer in a setting with universal health coverage. *Oncologist.* 2020;25(6):497-504.
| Author (year), country | PROM | Country | Language(s) of PROM | Study design | Target population | Sample size | Measurement domain | Number of items | Total score range |
|------------------------|------|---------|---------------------|--------------|-------------------|-------------|-------------------|-----------------|------------------|
| Aaronson et al., 1988 | EORTC QLQ-C30 Financial impact domain | 12 countries in Europe and North America | 9 languages | Development study and validation study | Age: 63±10 yrs Male: 76% Cancer type: Lung cancer | 305 | 1 domain regarding perceived financial impact | 1 | 1-4 |
| Aaronson et al., 1993 | | | | | | | | | |
| Avis et al., 2005 | QLACS financial problems domain | US | English | Development and validation study | Age: 71.4±11.5 yrs Male: 42% Cancer type: Breast, bladder, head and neck, gynecologic, prostate, colorectal cancer | 242 | 1 domain regarding financial problems | 4 | 4-28 |
| Chan et al., 2021 | COST-v2, Traditional Chinese version | Hong Kong | Traditional Chinese | Validation study | Age: 59.9±11.1 yrs Male: 35.3% Cancer type: Breast, gynecological, head and neck, gastric and colorectal, genitourinary, lung, hematological, skin, bone and soft tissue, brain and central nervous system cancer and others | 640 | No subdomain | 12 | 0-44 |
| Dar et al., 2021 | COST-v1, India version | India | Hindi or English | Validation study | Age: 49.5±16.8 yrs Male: 82.8% Cancer type: Tongue, gingival buccal sulcus, buccal mucosa, supraglottic larynx, hypopharynx, parotid and others | 29 | No subdomain | 11 | 0-44 |
| de Alcantara Nogueira et al., 2020 | COST-v1, Brazilian version | Brazil | Brazilian Portuguese | Validation study | Mean age: 56 yrs Male: 40.5% Cancer type: Not specific | 126 | No subdomain | 11 | 0-44 |
| Reference                        | Initiative | Country | Language | Study Type                       | Participants | Domain(s) | Age | Cancer Type                                                                |
|--------------------------------|------------|---------|----------|----------------------------------|--------------|-----------|-----|-----------------------------------------------------------------------------|
| De Souza et al., 2014          | COST-v1    | US      | English  | Development and validation study | 233          | No subdomain | 58.4±11.5 yrs, Male: 41.6% Cancer type: Not specified (diagnosis of AJCC stage IV cancer) | 0-68 |
| De Souza et al., 2017          | COST-v1    | US      | English  | Validation study                 | 257          | No subdomain | ≤50 yrs 23%; 51-64 yrs 30%; ≥65 yrs 48% Male: 46% Cancer type: Thoracic, breast, sarcoma, skin, central nervous system, gynecological, head and neck, colorectal, upper gastrointestinal, urological, and miscellaneous cancer | 0-44 |
| Durber et al., 2021            | COST-v1    | Australia | English  | Validation study                 | 103          | 1 domain regarding financial advice | 65 (41-90) yrs, Male: 48.5% Cancer type: Breast, colorectal/gastrointestinal, gynecological, prostate, and renal cancer | 5-25 |
| Harley et al. 2019             | CCEQ financial advice domain | UK | English  | Development and validation study | 313          | 2 domains: material and social capital. | 67 (41-88) yrs, Male: 50.0% Cancer type: Breast, colorectal/gastrointestinal, gynecological, prostate, and renal cancer | 0-68 |
| Head, 2008                      | SWBS       | US      | English  | Development and validation study | 266          | Two domains: material and social capital. | 59.6±12.7 yrs, Male: 35.7% Cancer type: Breast, melanoma, head and neck, prostate, rectum/anus, colon, endometrium, lung/tracheal/bronchus, and | 0-68 |
| Head & Faul, 2008               | SWBS       | US      | English  | Development and validation study |              |           |                               |

**Notes:**
- **SWBS:** Strengths and Weaknesses of Socioeconomic Status scale
- **CCEQ:** Comprehensive Care Experience Questionnaire
- **COST-v1:** Cancer of Stratum-V1 scale
- **AJCC:** American Joint Committee on Cancer staging system
| Authors               | Study Name | Country   | Language | Study Type          | Age, yrs              | Male, %  | Cancer Type                                                                 | Subdomain | NR   | 0-100 |
|----------------------|------------|-----------|----------|---------------------|-----------------------|----------|-------------------------------------------------------------------------------|-----------|------|-------|
| Hueniken et al., 2020| FIT        | Canada    | English  | Development and validation study | 61.6 (25.5-88.5) | 77.2%    | Oropharynx, oral cavity, larynx, nasopharynx, hypopharynx cancers, and others | No subdomain | 9    | 0-100 |
| Malin et al., 2006   | CanCORS    | US        | English  | Development and validation study | 66±11                 | 51%      | Lung and colorectal cancer                                                   | 1 domain regarding insurance coverage and income | NR  | NR    |
| Ripamonti et al., 2020| COST-v2, Italian version | Italy    | Italian  | Validation study      | 61.5±12.7                   | 52.5%    | Breast, lung, colon, gastric, hepatocellular, endometrial, prostate, sarcoma, bladder, head and neck, Hodgkin lymphoma, non-Hodgkin lymphoma, leukemia, myeloma, and others | No subdomain | 11   | 0-44  |
| Sharif et al., 2020  | COST-v2, Persian version | Iran     | Persian  | Validation study      | 50.0±14.3                | 51.0%    | Not specific                                                                  | No subdomain | 11   | 0-44  |
| Shilling et al., 2018| PRRS       | UK        | English  | Development and validation study | ≤50 yrs 25%; 51-65 yrs 41%; ≥66 yrs 34% | 23%      | Breast, gynecological, lung, and melanoma cancers                            | 1 domain regarding financial wellbeing | 6    | 0-24  |
| Tremblay et al., 2020| P-SAFE, French Version | Canada   | French   | Cross-adaption study   | 50-59 yrs 57%; 60-69 yrs 29%; ≥70 yrs 34% | 14%      | Colorectal, lung, breast, and prostate cancer                                | NR         | 23   | NR    |
| Veenstra et al., 2014| PFB        | US        | English  | Validation study      | <50 yrs 17%; 50-64 yrs 37%; 65-74 yrs 23%; >75 yrs 24% | 24%      | Not specific                                                                  | No subdomain | 7    | 0-7   |
| Researcher et al., Year | SDI Code | Country | Language | Study Type | Age | Male (%) | Cancer Type | Sample Size | Subdomain | Score |
|------------------------|----------|---------|----------|------------|-----|----------|-------------|-------------|-----------|-------|
| Wright et al., 2005    | SDI-21 providing for the family domain | UK | English | Development and validation study | 53.8±14.1 yrs | 53% | Stage III colorectal cancer | 271 | 1 domain regarding providing for the family | 0-20 |
| Wright et al., 2011    | SDI-16 money matters domain | UK | English | Development and validation study | 56 (18-88) yrs for men; 56 (21-88) yrs for women | 48% | Stage III colorectal cancer | 652 | 1 domain regarding money matters | 0-20 |
| Yu et al., 2021        | COST-v1, Simplified Chinese version | Mainland China | Chinese | Validation study | 57±9.2 yrs | 45.7% | Lung, stomach, colorectal, and breast cancer | 440 | No subdomain | 0-44 |
| Zebrack et al., 2010   | IOC-CS financial problems domain | US | English | Validation study | 26.7±5.3 yrs | 48.0% | Lung, stomach, colorectal, and breast cancer | 519 | 1 domain regarding financial problems | 1-15 |
| Zhao et al., 2009      | CPILS employment/financial domain | US | English | Validation study | ≤55 yrs 48.8%; >55 51.2% | 41.6% | Lung, stomach, colorectal, and breast cancer | 5901 | 1 domain regarding employment/finances | 0-12 |
### Table 2 Methodological quality assessment of the measures

| Author (year) | PROM Development | Content Validity | Structural Validity | Internal Consistency | Cross-cultural Validity | Reliability | Measurement Error | Criterion Validity | Hypothesis Testing | Responsiveness |
|---------------|------------------|------------------|--------------------|----------------------|------------------------|-------------|------------------|-------------------|-------------------|-----------------|
| Aaronson et al., 1988 | EORTC QLQ-C30 | Inadequate | R: NR | NR | NR | NR | NR | NR | NR | NR |
| Aaronson et al., 1993 | Financial impact domain | Inadequate | R: NR | C1: NR | C2: NR | NR | NA | NR | NR | NR |
| Avis et al., 2005 | QLACS financial problems domain | Inadequate | R: NR | C1: NR | C2: NR | Adequate | Very good | NR | NR | Adequate |
| Chan et al., 2021 | COST-v2, Traditional Chinese version | Doubtful | R: Doubtful | C1: NR | C2: Doubtful | Very good | Very good | NR | Doubtful | NR |
| Dar et al., 2021 | COST-v1, India version | Inadequate | R: NR | C1: NR | C2: NR | Inadequate | Very good | NR | NR | NR |
| de Alcantara Nogueira et al., 2020 | COST-v1, Brazilian version | Inadequate | R: NR | C1: NR | C2: NR | Very good | Very good | Inadequate | NR | NR |
| De Souza et al., 2014 | COST-v1 | Doubtful | R: Adequate | C1: Adequate | C2: Adequate | NR | Very good | NR | Adequate | NR |
| De Souza et al., 2017 | COST-v1, Australia | Inadequate | R: NR | C1: NR | NR | Very good | NR | Adequate | NR | Very good |
| Durber et al., 2021 | COST-v1, Australia | Inadequate | R: NR | C1: NR | NR | Very good | NR | Adequate | NR | Very good |
| Source                        | Version                                | R: Adequate | C1: Adequate | C2: Adequate | Adequate | Very good | NR | NR | NR | NR | Very good | NR |
|-------------------------------|----------------------------------------|-------------|--------------|--------------|----------|-----------|----|----|----|----|-----------|----|
| Harley et al., 2019          | CCEQ financial advice domain           | Doubtful    | Adequate     | Adequate     | Adequate | Very good | NR | NR | NR | NR | Very good | NR |
| Head, 2008                   | SWBS                                   | Inadequate  |              |              | R: NR    | C1: Doubt | Very good | Very good | NR | NR | NR | NR | Very good | NR |
| Head & Faul, 2008            | SWBS                                   | Inadequate  |              |              | R: NR    | C1: NR    | Adequate  | Very good | Doubtful | NR | NR | NR | Very good | Very good |
| Hueniken et al., 2020        | FIT                                    | Inadequate  |              |              |          |           |     |     |     |     |           |    |
| Malin et al., 2006           | CanCORS insurance coverage and income  | Inadequate  |              |              |          |           |     |     |     |     |           |    |
| Ripamonti et al., 2020       | COST-v2, Italian version               | Inadequate  | R: NR        | C1: NR       | Inadequate | Very good | NR | Inadequate | NR | Very good | Very good | NR |
| Sharif et al., 2020          | COST-v2, Persian version               | Inadequate  | R: NR        | C1: NR       | Very good | Very good | NR | NR | NR | NR | Inadequate | NR |
| Shilling et al., 2018        | PRRS financial wellbeing domain        | Inadequate  | R: NR        | C1: NR       | Adequate  | Very good | NR | Inadequate | NR | Inadequate | NR | NR |
| Tremblay et al., 2020        | P-SAFE, French version                 | Inadequate  |              |              |          | NR        | NR | Inadequate | NR | NR | NR | NR | NR |
| Veenstra et al., 2014        | PFB                                    | Inadequate  | R: NR        | C1: NR       | Very good | Inadequate | NR | NR | NR | NR | NR | NR |
| Wright et al., 2005          | SDI-21 providing for                    | Inadequate  | R: Doubt     | C1: NR       | Very good | Very good | NR | Inadequate | NR | Inadequate | NR | NR |
| Study Reference | Domain Details | C2: NR | R: Doubt | C1: NR | C2: NR | Very good | Very good | NR | Inadequate | NR | Inadequate | NR | NR | NR | Doubt | Doubt |
|----------------|----------------|--------|----------|--------|--------|-----------|-----------|----|------------|----|------------|----|----|----|--------|--------|
| Wright et al., 2011 | SDI-16 money matters in the family domain |        |          |        |        |           |           |    |            |    |            |    |    |    |        |        |
| Yu et al., 2021    | COST-v1, Simplified Chinese version | Inadequate | R: NR    | C1: NR | C2: NR | Very good | Very good | NR | Very good  | NR | NR         | Doubt | Doubt | |
| Zebrack et al., 2010 | IOC-CS financial problems domain | Inadequate | R: NR    | C1: NR | C2: Doubt | Adequate | Very good | NR | Adequate  | NR | Inadequate | Very good | NR | |
| Zhao et al., 2009  | CPILS employment/financial domain | Inadequate | R: NR    | C1: NR | C2: NR | Very good | Very good | NR | NR        | NR | NR         | Very good | NR | |

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml
| PROM                                      | Author (year)                           | Structural validity | Internal consistency | Reliability | Measurement error | Hypothesis testing | Cross-cultural validity | Criterion validity | Responsiveness |
|-------------------------------------------|-----------------------------------------|---------------------|----------------------|-------------|-------------------|--------------------|------------------------|-------------------|------------------|
| CanCORS insurance coverage and income     | Malin et al., 2006                      | NR                  | NR                   | NR          | NR                | NR                 | NR                     | NR                | NR               |
| domain                                    |                                         |                     |                      |             |                   |                    |                        |                   |                  |
| CCEQ financial advice domain              | Harley et al., 2019                     | NR                  | +                    | NR          | NR                | +                  | NR                     | NR                | NR               |
| domain                                    |                                         |                     |                      |             |                   |                    |                        |                   |                  |
| COST-v1                                   | De Souza et al., 2014; De Souza et al., 2017 | NR                  | +                    | +           | NR                | +                  | NR                     | NR                | NR               |
| COST-v1, Australia version                | Durber et al., 2021                     | NR                  | +                    | +           | NR                | +                  | NR                     | NR                | NR               |
| COST-v1, Brazilian version                | de Alcantara Nogueira et al., 2020      | -                   | +                    | NR          | NR                | NR                 | +                      | NR                | NR               |
| COST-v1, India version                    | Dar et al., 2021                        | -                   | ?                    | NR          | NR                | NR                 | NR                     | NR                | NR               |
| COST-v1, Simplified Chinese version       | Yu et al., 2021                         | -                   | +                    | +           | NR                | +                  | NR                     | NR                | ?                |
| COST-v2, Italian version                  | Ripamonti et al., 2020                  | ?                   | ?                    | -           | NR                | ?                  | NR                     | ?                 | NR               |
| COST-v2, Persian version                  | Sharif et al., 2020                     | +                   | +                    | NR          | NR                | ?                  | NR                     | NR                | NR               |
| COST-v2, Traditional Chinese version      | Chan et al., 2021                       | -                   | +                    | +           | NR                | +                  | NR                     | NR                | NR               |
| CPILS employment/financial domain         | Zhao et al., 2009                       | ?                   | +                    | NR          | NR                | ?                  | NR                     | NR                | NR               |
| EORTC QLQ-C30 financial impact domain     | Aaronson et al., 1988; Aaronson et al., 1993 | NR                  | NA                   | NR          | NR                | NR                 | NR                     | NR                | NR               |
| FIT                                       | Hueniken et al., 2020                   | ?                   | +                    | +           | NR                | +                  | NR                     | NR                | +                |
| Domain                                      | Source                                                                 | IOC-CS | PFB | PRRS financial wellbeing | P-SAFE, French Version | QLACS financial problems domain | SDI-16 money matters domain | SDI-21 providing for the family domain | SWBS |
|--------------------------------------------|------------------------------------------------------------------------|-------|-----|--------------------------|-------------------------|-------------------------------|-----------------------------|----------------------------------------|-------|
| IOC-CS financial problems domain          | Zebrack et al., 2010                                                   | ?     | -   | -                        | NR                      | ?                             | NR                          | -                                      | NR    |
| PFB                                        | Veenstra et al., 2014                                                  | ?     | +   | NR                       | NR                      | NR                            | NR                          | NR                                     | NR    |
| PRRS financial wellbeing domain            | Shilling et al., 2018                                                  | ?     | +   | +                        | NR                      | NR                            | NR                          | +                                      | NR    |
| P-SAFE, French Version                     | Tremblay et al., 2020                                                 | NR    | NR  | NR                       | NR                      | NR                            | NR                          | +                                      | NR    |
| QLACS financial problems domain            | Avis et al., 2005                                                     | -     | +   | NR                       | NR                      | +                             | NR                          | -                                      | NR    |
| SDI-16 money matters domain                | Wright et al., 2011                                                   | ?     | +   | -                        | NR                      | NR                            | NR                          | ?                                      | NR    |
| SDI-21 providing for the family domain     | Wright et al., 2005                                                   | ?     | +   | -                        | NR                      | NR                            | NR                          | ?                                      | NR    |
| SWBS                                       | Head, 2008; Head & Faul, 2008                                          | ?     | +   | NR                       | NR                      | +                             | NR                          | NR                                     | NR    |
## Table 4 Certainty of evidence of measurement properties

| PROM                                                                 | Author (year)                      | Structural validity | Internal consistency | Reliability | Measurement error | Hypothesis testing | Cross-cultural validity | Criterion validity | Responsiveness |
|----------------------------------------------------------------------|------------------------------------|---------------------|----------------------|-------------|-------------------|--------------------|------------------------|-------------------|-----------------|
| CanCORS insurance coverage and income domain                        | Malin et al., 2006                | -                   | -                    | -           | -                 | -                  | -                      | -                | -               |
| CCEQ financial advice domain                                         | Harley et al., 2019               | Moderate            | High                 | -           | -                 | High               | -                      | -                | -               |
| COST-v1                                                              | De Souza et al., 2014; De Souza et al., 2017 | -       | High                 | Moderate    | -                 | High               | -                      | -                | -               |
| COST-v1, Australia version                                           | Durber et al., 2021               | -                   | High                 | Moderate    | -                 | High               | -                      | -                | -               |
| COST-v1, Brazilian version                                           | de Alcantara Nogueira et al., 2020 | High               | High                 | -           | -                 | Very low           | -                      | -                | -               |
| COST-v1, India version                                               | Dar et al., 2021                  | Very Low            | Low                  | -           | -                 | -                  | -                      | -                | -               |
| COST-v1, Simplified Chinese version                                  | Yu et al., 2021                   | High                | High                 | High        | Low               | -                  | -                      | Low              | -               |
| COST-v2, Italian version                                             | Ripamonti et al., 2020            | Very low            | High                 | Very low    | High              | -                  | High                   | -                | -               |
| COST-v2, Persian version                                             | Sharif et al., 2020               | High                | High                 | -           | -                 | Very low           | -                      | -                | -               |
| COST-v2, Traditional Chinese version                                 | Chan et al., 2021                 | High                | High                 | Low         | High              | -                  | -                      | -                | -               |
| CPILS employment/financial domain                                    | Zhao et al., 2009                 | High                | High                 | -           | -                 | High               | -                      | -                | -               |
| EORTC QLQ-C30 financial impact domain                                | Aaronson et al., 1988; Aaronson et al., 1993 | -       | -                    | -           | -                 | -                  | -                      | -                | -               |
| FIT                                                                  | Hueniken et al., 2020             | Low                 | Moderate             | Very Low    | -                 | Moderate           | -                      | -                | Moderate        |
| Domain                                      | Study/Version                          | 2010   | 2014   | 2018   | 2020   | 2005   | 2011   | 2008; 2008 |
|--------------------------------------------|----------------------------------------|--------|--------|--------|--------|--------|--------|-----------|
| IOC-CS financial problems domain           | Zebrack et al., 2010                   | Moderate | High   | Moderate | -     | High   | -      | Very low |
| PFB                                        | Veenstra et al., 2014                  | Moderate | Very low | -     | -     | -     | -     | -         |
| PRRS financial wellbeing domain            | Shilling et al., 2018                  | Moderate | High   | Very low | -     | -     | Very low |           |
| P-SAFE, French Version                     | Tremblay et al., 2020                  | -      | -      | -      | -      | -      | Very low | -         |
| QLACS financial problems domain            | Avis et al., 2005                      | Moderate | High   | -      | Low   | -      | Moderate | -         |
| SDI-16 money matters domain                | Wright et al., 2011                    | High   | High   | Very low | -     | -     | Very low | -         |
| SDI-21 providing for the family domain     | Wright et al., 2005                    | High   | High   | Very low | -     | -     | Very low | -         |
| SWBS                                       | Head, 2008; Head & Faul, 2008          | High   | High   | -      | High   | -      | -      | -         |
Figure 1 PRISMA Flow Chat of Selection Process

Appendix 1 Search Strategy and Results
PubMed (N=2981); MEDLINE (OVID) (N=26); Embase (OVID) (N=1721); Web of Science (N=425); CINAHL (EBSCO) (N=774); PsycINFO (N=33); ProQuest (N=3); Cochrane Library (N=4304)

Other resources (N=6)

Total number after deduplication (N=9405) → Records excluded (N=9348)

Full-text papers assessed for eligibility (N=57) → Full text excluded (N=33)
- Phenomena of interest (n=4)
- Population group (n=15)
- Study design (n=9)
- Abstract (n=5)

Studies included in the review (N=24)

Identification
Screening
Eligibility
Included

Other resources (N=6)
### Appendix 1 Search strategy and results

**Search strategy for PubMed**

Search time: 2021-5-3 14:32 (UTC+8)

| Search | Query                                                                 | Items found |
|--------|------------------------------------------------------------------------|-------------|
| #1     | Cancer[Title/Abstract] OR neoplasms[MeSH]                             | 3971704     |
| #2     | Patient'[Title/Abstract] OR survivor'[Title/Abstract] OR patients[MeSH] OR “cancer survivors”[MeSH] OR survivors[MeSH] | 2564363     |
| #3     | Cost[Title/Abstract] OR bill'[Title/Abstract] OR expense[Title/Abstract] OR “productivity loss”[Title/Abstract] OR “out-of-pocket”[Title/Abstract] OR “economic burden”[Title/Abstract] OR “financial toxicity”[Title/Abstract] OR “financial hardship”[Title/Abstract] OR “financial effect”[Title/Abstract] OR “financial stress”[Title/Abstract] OR “economic burden”[Title/Abstract] OR “economic hardship”[Title/Abstract] OR “co-payment”[Title/Abstract] OR “cost of illness”[MeSH] | 530485      |
| #4     | Scale'[Title/Abstract] OR “patient reported outcome measure”'[Title/Abstract] OR PROM? [Title/Abstract] OR measure* [Title/Abstract] OR “Patient Reported Outcome Measures*[MeSH] OR “Surveys and Questionnaires*[MeSH] | 4782348     |
| #5     | (instrumentation[sh] OR methods[sh] OR “Validation Studies”[pt] OR “Comparative Study”[pt] OR “psychometrics*[MeSH] OR psychometr*[tiab] OR clinimetr*[tw] OR clinometr*[tw] OR “outcome assessment (health care)”[MeSH] OR “outcome assessment”[tiab] OR “outcome measure”[tw] OR “observer variation”[MeSH] OR “observer variation”[tiab] OR “Health Status Indicators*[Mesh] OR “reproducibility of results*[MeSH] OR reproduci*[tiab] OR “discriminant analysis*[MeSH] OR reliab*[tiab] OR un reliab*[tiab] OR valid*[tiab] OR “coefficient of variation*[tiab] OR coefficient[tiab] OR homogeneity[tiab] OR homogeneous[tiab] OR “internal consistency*[tiab] OR (cronbach*[tiab] AND (alpha[tiab] OR alpha*[tiab])) OR (item[tiab] AND (correlation*[tiab] OR selection*[tiab] OR reduction*[tiab])) OR agreement[tw] OR precision[tw] OR imprecision[tw] OR “precise values”[tw] OR test-retest[tiab] OR (test[tiab] AND retest[tiab]) OR (reliab*[tiab] OR (test[tiab] OR retest[tiab])) OR stability[tiab] OR interrater[tiab] OR inter-rater[tiab] OR intrarater[tiab] OR intra-rater[tiab] OR interrater[tiab] OR inter-rater[tiab] OR intertester[tiab] OR intra-tester[tiab] OR intra-rater[tiab] OR intratester[tiab] OR (test[tiab] OR tests[tiab]) OR generaliza*[tiab] OR (intraclass[tiab] AND correlation*[tiab]) OR discriminative[tiab] OR “known group”[tiab] OR “factor analysis*[tiab] OR “factor analyses*[tiab] OR “factor structure*[tiab] OR “factor structures*[tiab] OR dimension*[tiab] OR subscale*[tiab] OR (multitrait*[tiab] AND scaling[tiab] AND (analysis[tiab] OR analyses[tiab])) OR “item discriminant*[tiab] OR “interscale correlation*[tiab] OR error[tiab] OR errors[tiab] OR “individual variability*[tiab] | 9660087     |
Search strategy for MEDLINE (Ovid)
Search time: 2021-5-3 14:46 (UTC+8)

| Search | Query                                                                 | Items found |
|--------|-----------------------------------------------------------------------|-------------|
| #1     | cancer.ab. or cancer.ti. or neoplasms.hw.                            | 3022755     |
| #2     | patient?.ab. or patient?.ti. or survivor?.ab. or survivor?.ti. or patients.hw. or cancer survivors.hw. or survivors.hw. | 6174913     |
| #3     | Cost.ab. or Cost.ti. or bill?.ab. or bill?.ti. or expense.ab. or expense.ti. or productivity loss.ab. or productivity loss.ti. or out-of-pocket.ab. or out-of-pocket.ti. or economic burden.ab. or economic burden.ti. or financial toxicity.ab. or financial toxicity.ti. or financial hardship.ab. or financial hardship.ti. or financial burden.ab. or financial burden.ti. or financial effect.ab. or financial effect.ti. or financial stress.ab. or financial stress.ti. or economic burden.ab. or economic burden.ti. or economic hardship.ab. or economic hardship.ti. or co-payment.ab. or co-payment.ti. or (cost of illness).hw. | 419745     |
| #4     | Scale?.ab. or Scale?.ti. or patient reported outcome measur*.ab. or patient reported outcome measur*.ti. or PROM?.ab. or PROM?.ti. or measure*.ab. or measure*.ti. or Patient Reported Outcome Measures*.ab. or Patient Reported Outcome Measures*.ti. or (Surveys and Questionnaires).hw. | 3676481     |
| #5     | (instrumentation or methods).fs.                                     | 4290317     |
| #6     | (Validation Studies or Comparative Study).pt.                        | 1888426     |
|   | exp Psychometrics/                                      | 78410     |
|---|--------------------------------------------------------|-----------|
|   | psychometr*.ti,ab.                                     | 41677     |
|   | (clinimetr* or clinometr*).tw.                         | 1010      |
|   | exp "Outcome Assessment Health Care"/                  | 1188149   |
|   | outcome assessment.ti,ab.                              | 3533      |
|   | exp Observer Variation/                                | 212018    |
|   | observer variation.ti,ab.                              | 43395     |
|   | exp Health Status Indicators/                          | 319012    |
|   | exp Reproducibility of Results/                        | 416843    |
|   | reproducib*.ti,ab.                                     | 143145    |
|   | exp Discriminant Analysis/                             | 11117     |
|   | (reliab* or unrelia*b* or valid* or coefficient or homogeneity or homogeneous or internal consistency).ti,ab. | 1217341   |
|   | (cronbach* and (alpha or alphas)).ti,ab.               | 20027     |
|   | (item and (correlation* or selection* or reduction*)).ti,ab. | 19993     |
|   | (agreement or precision or "precise values" or test-retest).ti,ab. | 339147    |
|   | (test and retest).ti,ab.                               | 24523     |
|   | (reliab* and (test or retest)).ti,ab.                  | 80342     |
|   | (stability or interrater or inter-rater or intrarater or intra-rater or intertaster or inter-tester or intratester or intra-tester or interobserver or inter-observer or intraobserver or intraobserver or intertechnician or inter-technician or intratechnician or intra-technician or interexaminer or inter-examiner or intralexaminer or inter-lexaminer or interassay or inter-assay or intraassay or intra-assay or intra-individual or inter-individual or intra-individual or interparticipant or inter-participant or intraparticipant or intra-participant or kappa or kappa's or kappas or repeatab*).ti,ab. | 507384    |
|   | (((repliab* or repeated) and (measure or measures or findings or result or results or test or tests)).ti,ab. | 179804    |
|   | (generaliza* or generalisa* or concordance).ti,ab.     | 78096     |
|   | (intraclass and correlation*).ti,ab.                   | 22856     |
|   | (discriminative or known group or factor analysis or factor analyses or dimension* or subscale*).ti,ab. | 506663    |
|   | (multitrait and scaling and (analysis or analyses)).ti,ab. | 132       |
|   | (item discriminant or interscale correlation* or error or errors or individual variability).ti,ab. | 252802    |
|   | (variability and (analysis or values)).ti,ab.          | 90102     |
|   | (uncertainty and (measurement or measuring)).ti,ab.     | 5260      |
|   | (standard error of measurement or sensitiv* or responsive*).ti,ab. | 1429839   |
|   | (((minimal or minimally or clinical or clinically) and (important or significant or detectable) and (change or difference)).ti,ab. | 214078    |
|   | (small* and (real or detectable) and (change or difference)).ti,ab. | 6561      |
|   | (meaningful change or ceiling effect or floor effect or Item response model or IRT or Rasch or Differential item functioning or DIF or computer adaptive testing or item bank or cross-cultural equivalence).ti,ab. | 11890     |
### Search strategy for EMBASE (Ovid)

**Search time:** 2021-5-3 16:48 (UTC+8)

| Search | Query                                                                                           | Items found |
|--------|-------------------------------------------------------------------------------------------------|-------------|
| #1     | cancer.ab. or cancer.ti. or neoplasms.hw.                                                      | 2586704     |
| #2     | patient?.ab. or patient?.ti. or survivor?.ab. or survivor?.ti. or patients.hw. or cancer survivors.hw. or survivors.hw. | 10409882    |
| #3     | Cost.ab. or Cost.ti. or bill?.ab. or bill?.ti. or expense.ab. or expense.ti. or productivity loss.ab. or productivity loss.ti. or out-of-pocket.ab. or out-of-pocket.ti. or economic burden.ab. or economic burden.ti. or financial toxicity.ab. or financial toxicity.ti. or financial hardship.ab. or financial hardship.ti. or financial burden.ab. or financial burden.ti. or financial effect.ab. or financial effect.ti. or financial stress.ab. or financial stress.ti. or economic burden.ab. or economic burden.ti. or economic hardship.ab. or economic hardship.ti. or co-payment.ab. or co-payment.ti. or (cost of illness).hw. | 702534     |
| #4     | Scale?.ab. or Scale?.ti. or patient reported outcome measur*.ab. or patient reported outcome measur*.ti. or PROM?.ab. or PROM?.ti. or measure*.ab. or measure*.ti. or Patient Reported Outcome Measures*.ab. or Patient Reported Outcome Measures*.ti. or (Surveys and Questionnaires).hw. | 5336565    |
| #5     | exp Intermethod comparison/ OR exp data collection method/ OR exp validation study/ OR exp feasibility study/ OR exp pilot study/ OR exp psychometry/ OR exp reproducibility/ OR reproducib*:ab,ti OR audit:ab,ti OR psychometr*:ab,ti OR clinimetr*:ab,ti OR clinometr*:ab,ti OR exp observer variation/ OR observer variation:ab,ti OR exp discriminant analysis/ OR exp validity/ OR reliab*:ab,ti OR valid*:ab,ti OR coefficient:ab,ti OR internal consistency:ab,ti OR (cronbach*:ab,ti AND (alpha:ab,ti OR alphas:ab,ti)) OR item correlation:ab,ti OR item item correlations:ab,ti OR item selection:ab,ti OR item item selection:ab,ti OR item item reduction:ab,ti OR item item reductions:ab,ti OR agreement:ab,ti OR precision:ab,ti OR imprecision:ab,ti OR precise values:ab,ti OR test-retest:ab,ti OR (test:ab,ti AND retest:ab,ti) OR (reliab*:ab,ti AND (test:ab,ti OR retest:ab,ti)) OR stability:ab,ti OR interrater:ab,ti OR inter-rater:ab,ti OR intra-rater:ab,ti OR interrater:ab,ti OR intratester:ab,ti OR inter-tester:ab,ti OR intratester:ab,ti OR inter-individual:ab,ti OR intra-individual:ab,ti OR intraindividual:ab,ti OR intra-individual:ab,ti OR inter-participant:ab,ti OR inter-participant:ab,ti OR intra-participant:ab,ti OR kappa:ab,ti OR kappa:ab,ti OR coefficient of variation:ab,ti OR | 2110942   |

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml
Search strategy for CINAHL (EBSCO)
Search time: 2021-5-3 17:10 (UTC+8)

| Search | Query | Items found |
|--------|-------|-------------|
| #1     | TI cancer OR AB cancer OR MH neoplasms | 438994 |
| #2     | AB patient? OR TI patient? OR AB survivor? OR TI survivor? OR MH patients OR MH “cancer survivors” OR MH survivors | 1637009 |
| #3     | AB Cost OR TI Cost OR AB bill? OR TI bill? OR AB expense OR TI expense OR AB “productivity loss” OR TI “productivity loss” OR AB out-of-pocket OR TI out-of-pocket OR AB “economic burden” OR TI “economic burden” OR AB “financial toxicity” OR TI “financial toxicity” OR AB “financial hardship” OR TI “financial hardship” OR AB “financial burden” OR TI “financial burden” OR AB “financial effect” OR TI “financial effect” OR AB “financial stress” OR TI “financial stress” OR AB “economic burden” OR TI “economic burden” OR AB “economic hardship” OR TI “economic hardship” OR AB co-payment OR TI co-payment OR MH “cost of illness” | 186320 |
| #4     | AB Scale? OR TI Scale? OR AB “patient reported outcome measur***” OR TI “patient reported outcome measur***” OR AB PROM? OR TI PROM? OR AB measure* OR TI measure* OR AB “Patient Reported Outcome Measures***” OR MH “Patient Reported Outcome Measures***” OR MH “Surveys and Questionnaires” | 862405 |
| #5     | (MH “Psychometrics”) or ( TI psychometr* or AB psychometr* ) or ( TI clinimetr* or AB | 649386 |
For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

| Search Strategy for PsycINFO (EBSCO) |
|--------------------------------------|
| Search time: 2021-5-3 17:32 (UTC+8) |

| Search | Query | Items found |
|--------|-------|-------------|
| #1     | TI cancer OR AB cancer OR MH neoplasms | 1172        |
| #2     | AB patient? OR TI patient? OR AB survivor? OR TI survivor? OR MH patients OR MH “cancer survivors” OR MH survivors | 16703       |
| #3     | AB Cost OR TI Cost OR AB bill? OR TI bill? OR AB expense OR TI expense OR AB “productivity loss” OR TI “productivity loss” OR AB out-of-pocket OR TI out-of-pocket OR AB “economic burden” OR TI “economic burden” OR AB “financial toxicity” OR TI “financial toxicity” OR AB “financial hardship” OR TI “financial hardship” OR AB “financial burden” OR TI “financial burden” OR AB “financial effect” OR TI “financial effect” OR AB “financial | 3812        |
**Search strategy for Web of Science**

Search time: 2021-5-3 17:57 (UTC+8)

| Search | Query | Items found |
|--------|-------|-------------|
| #1     | TI=cancer OR TS=(cancer OR neoplasms) | 2596936 |
| #2     | TI=(patient? OR survivor?) OR TS=(patients OR "cancer survivors" OR survivors) | 6121549 |
| #3     | TI=(Cost OR bill? OR expense OR "productivity loss" OR out-of-pocket OR "economic burden" OR "financial toxicity" OR "financial hardship" OR "financial burden" OR "financial effect" OR "financial stress" OR "economic burden" OR "economic hardship") OR TS="cost of illness" | 211496 |
| #4     | TI=(Scale? OR "patient reported outcome measur*" OR PROM? OR measure*) OR TS="Patient Reported Outcome Measures*" OR "Surveys and Questionnaires" | 788584 |
| #6     | #4 AND #3 AND #2 AND #1 | 425 |

**Search strategy for ProQuest Dissertations and Theses**

Search time: 2021-5-3 18:41 (UTC+8)

| Search | Query | Items found |
|--------|-------|-------------|
| #1     | ti(cancer OR neoplasms) OR su(cancer OR neoplasms) | 41848 |
| #2     | ti(patient? OR survivor? OR "cancer survivors") OR su(patient? OR survivor? OR cancer survivors) | 44297 |
| #3     | Title : Cost OR bill? OR expense OR "productivity loss" OR out-of-pocket OR "economic burden" OR "financial toxicity" OR "financial hardship" OR "financial burden" OR "financial effect" OR "financial stress" OR Summary : Cost OR bill? OR expense OR "productivity loss" OR out-of-pocket OR "economic burden" OR "financial toxicity" OR "financial hardship" OR "financial burden" OR "financial effect" OR "financial stress" OR Title : "economic burden" OR "economic hardship" OR "cost of illness" OR Summary : "economic burden" OR "economic hardship" OR "cost of illness" | 32997 |
| #4     | All : Scale? OR "patient reported outcome measure" OR PROM? OR measure* | 2738959 |
| #6     | #4 AND #3 AND #2 AND #1 | 3 |

**Search strategy for Cochrane Library (Wiley)**

Search time: 2020-5-18 13:57 (UTC+8)

| Search | Query | Items found |
|--------|-------|-------------|
| #1     | #1 | 47588 |
| #1 | (Cancer):ti,ab,kw | 164548 |
|----|------------------|--------|
| #2 | MeSH descriptor: [Neoplasms] explode all trees | 81536 |
| #3 | (patient? OR survivor? OR "cancer survivors"):ti,ab,kw | 1016991 |
| #4 | MeSH descriptor: [Patients] explode all trees | 2715 |
| #5 | MeSH descriptor: [Cancer Survivors] explode all trees | 255 |
| #6 | MeSH descriptor: [Survivors] explode all trees | 1409 |
| #7 | (Cost OR bill? OR expense OR "productivity loss" OR out-of-pocket OR "economic burden" OR "financial toxicity" OR "financial hardship" OR "financial burden" OR "financial effect" OR "financial stress" OR "economic burden" OR "economic hardship" OR "cost of illness"):ti,ab,kw | 62550 |
| #8 | MeSH descriptor: [Cost of Illness] explode all trees | 802 |
| #9 | (Scale? OR "patient reported outcome measure" OR PROM? OR measure*):ti,ab,kw | 560355 |
| #10 | MeSH descriptor: [Patient Reported Outcome Measures] explode all trees | 706 |
| #11 | MeSH descriptor: [Surveys and Questionnaires] explode all trees | 55522 |
| #12 | ("Validation Studies" OR "Comparative Study"):pt | 167084 |
| #13 | MeSH descriptor: [Psychometrics] explode all trees | 2846 |
| #14 | (psychometr*):ti,ab,kw | 6346 |
| #15 | #1 OR #2 | 194269 |
| #16 | #3 OR #4 OR #5 OR #6 | 924360 |
| #17 | #7 OR #8 OR #9 OR #10 OR #11 | 621163 |
| #18 | #12 OR #13 OR #14 | 172446 |
| #19 | #15 AND #16 AND #17 AND #18 | 4304 |
### PRISMA 2020 Checklist

| Section and Topic | Item # | Checklist item                                                                                                                                                                                                 | Location where item is reported |
|-------------------|--------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------|
| **TITLE**         | 1      | Identify the report as a systematic review.                                                                                                                                                                  | Title                           |
| **ABSTRACT**      | 2      | See the PRISMA 2020 for Abstracts checklist.                                                                                                                                                                 | P1-2                            |
| **INTRODUCTION**  | 3      | Describe the rationale for the review in the context of existing knowledge.                                                                                                                                   | P3-5                            |
| **OBJECTIVES**    | 4      | Provide an explicit statement of the objective(s) or question(s) the review addresses.                                                                                                                       | P5                              |
| **METHODS**       | 5      | Specify the inclusion and exclusion criteria for the review and how studies were grouped for the syntheses.                                                                                                  | P6                              |
|                   | 6      | Specify all databases, registers, websites, organisations, reference lists and other sources searched or consulted to identify studies. Specify the date when each source was last searched or consulted. | P5-6                            |
|                   | 7      | Present the full search strategies for all databases, registers and websites, including any filters and limits used.                                                                                         | Appendix 1                     |
|                   | 8      | Specify the methods used to decide whether a study met the inclusion criteria of the review, including how many reviewers screened each record and each report retrieved, whether they worked independently, and if applicable, details of automation tools used in the process. | P6-7                            |
|                   | 9      | Specify the methods used to collect data from reports, including how many reviewers collected data from each report, whether they worked independently, any processes for obtaining or confirming data from study investigators, and if applicable, details of automation tools used in the process. | P6-7                            |
|                   | 10a    | List and define all outcomes for which data were sought. Specify whether all results that were compatible with each outcome domain in each study were sought (e.g. for all measures, time points, analyses), and if not, the methods used to decide which results to collect. | P7-8                            |
|                   | 10b    | List and define all other variables for which data were sought (e.g. participant and intervention characteristics, funding sources). Describe any assumptions made about any missing or unclear information. | P7-8                            |
|                   | 11     | Specify the methods used to assess risk of bias in the included studies, including details of the tool(s) used, how many reviewers assessed each study and whether they worked independently, and if applicable, details of automation tools used in the process. | P7                              |
| Study risk of bias assessment | 12     | Specify for each outcome the effect measure(s) (e.g. risk ratio, mean difference) used in the synthesis or presentation of results.                                                                          | NA                              |
| **SYNTHESIS METHODS** | 13a    | Describe the processes used to decide which studies were eligible for each synthesis (e.g. tabulating the study intervention characteristics and comparing against the planned groups for each synthesis (item #5)). | P8                              |
|                   | 13b    | Describe any methods required to prepare the data for presentation or synthesis, such as handling of missing summary statistics, or data conversions.                                                           | P8                              |
|                   | 13c    | Describe any methods used to tabulate or visually display results of individual studies and syntheses.                                                                                                       | P8                              |
|                   | 13d    | Describe any methods used to synthesize results and provide a rationale for the choice(s). If meta-analysis was performed, describe the model(s), method(s) to identify the presence and extent of statistical heterogeneity, and software package(s) used. | P8                              |
|                   | 13e    | Describe any methods used to explore possible causes of heterogeneity among study results (e.g. subgroup analysis, meta-regression).                                                                    | NA                              |
|                   | 13f    | Describe any sensitivity analyses conducted to assess robustness of the synthesized results.                                                                                                                    | NA                              |
| **REPORTING BIAS ASSESSMENT** | 14     | Describe any methods used to assess risk of bias due to missing results in a synthesis (arising from reporting biases).                                                                                   | NA                              |
| **CERTAINTY ASSESSMENT** | 15     | Describe any methods used to assess certainty (or confidence) in the body of evidence for an outcome.                                                                                                | P8-9                            |
## PRISMA 2020 Checklist

| Section and Topic | Item # | Checklist item                                                                                                                                                                                                                                                                                                                                 | Location where item is reported |
|-------------------|--------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------|
| **RESULTS**       |        |                                                                                                                                                                                                                                                                                                                                             |                                 |
| Study selection   | 16a    | Describe the results of the search and selection process, from the number of records identified in the search to the number of studies included in the review, ideally using a flow diagram.                                                                                                         | Figure 1                        |
|                   | 16b    | Cite studies that might appear to meet the inclusion criteria, but which were excluded, and explain why they were excluded.                                                                                                                                                                                                                  | P9                              |
| Study characteristics | 17    | Cite each included study and present its characteristics.                                                                                                                                                                                                                                                                                    | Table 1                         |
| Risk of bias in studies | 18    | Present assessments of risk of bias for each included study.                                                                                                                                                                                                                                                                               | Appendix 3                      |
| Results of individual studies | 19    | For all outcomes, present, for each study: (a) summary statistics for each group (where appropriate) and (b) an effect estimate and its precision (e.g. confidence/credible interval), ideally using structured tables or plots.                                                    | Appendix 2 & Table 1            |
| Results of syntheses | 20a   | For each synthesis, briefly summarise the characteristics and risk of bias among contributing studies.                                                                                                                                                                          | Appendix 2                      |
|                   | 20b    | Present results of all statistical syntheses conducted. If meta-analysis was done, present for each the summary estimate and its precision (e.g. confidence/credible interval) and measures of statistical heterogeneity. If comparing groups, describe the direction of the effect. | NA                              |
|                   | 20c    | Present results of all investigations of possible causes of heterogeneity among study results.                                                                                                                                                                                   | NA                              |
|                   | 20d    | Present results of all sensitivity analyses conducted to assess the robustness of the synthesized results.                                                                                                                                                                    | Table 2                         |
| Reporting biases  | 21     | Present assessments of risk of bias due to missing results (arising from reporting biases) for each synthesis assessed.                                                                                                                                                             | NA                              |
| Certainty of evidence | 22   | Present assessments of certainty (or confidence) in the body of evidence for each outcome assessed.                                                                                                                                                                               | Table 2                         |
| **DISCUSSION**    |        |                                                                                                                                                                                                                                                                                                                                             |                                 |
| Discussion        | 23a    | Provide a general interpretation of the results in the context of other evidence.                                                                                                                                                                                                                                                        | P12                             |
|                   | 23b    | Discuss any limitations of the evidence included in the review.                                                                                                                                                                                                                 | P15-16                          |
|                   | 23c    | Discuss any limitations of the review processes used.                                                                                                                                                                                                                                                                                      | P15-16                          |
|                   | 23d    | Discuss implications of the results for practice, policy, and future research.                                                                                                                                                                                                   | P16                             |
| **OTHER INFORMATION** |      |                                                                                                                                                                                                                                                                                                                                             |                                 |
| Registration and protocol | 24a  | Provide registration information for the review, including register name and registration number, or state that the review was not registered.                                                                                                                                                                                                 | P5                              |
|                   | 24b    | Indicate where the review protocol can be accessed, or state that a protocol was not prepared.                                                                                                                                                                                                                                           | P5                              |
|                   | 24c    | Describe and explain any amendments to information provided at registration or in the protocol.                                                                                                                                                                               | No amendments                   |
| Support           | 25     | Describe sources of financial or non-financial support for the review, and the role of the funders or sponsors in the review.                                                                                                                                                      | Title page                      |
| Competing interests | 26    | Declare any competing interests of review authors.                                                                                                                                                                                                                             | Title page                      |
| Availability of data, code and other materials | 27    | Report which of the following are publicly available and where they can be found: template data collection forms; data extracted from included studies; data used for all analyses; analytic code; any other materials used in the review.                               | NA                              |
# Psychometric properties of self-reported financial toxicity measures in cancer survivors: A systematic review

| Journal:            | BMJ Open                                           |
|---------------------|---------------------------------------------------|
| Manuscript ID       | bmjopen-2021-057215.R1                            |
| Article Type:       | Original research                                 |
| Date Submitted by the Author: | 30-Mar-2022                                      |
| Complete List of Authors: | Zhu, Zheng; Fudan University School of Nursing, Xing, Weijie; Fudan University School of Nursing, Wen, Huan; Fudan University Sun, Yanling; Fudan University So, WK; The Chinese University of Hong Kong Lizarondo, Lucylynn; University of Adelaide, The Joanna Briggs Institute Peng, Jian; Fudan University School of Nursing Hu, Yan; Fudan University, School of Nursing |
| Primary Subject Heading: | Oncology                                        |
| Secondary Subject Heading: | Nursing, Health economics                          |
| Keywords:           | ONCOLOGY, HEALTH ECONOMICS, Health economics < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Quality in health care < HEALTH SERVICES ADMINISTRATION & MANAGEMENT |
I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd (“BMJ”) its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge (“APC”) for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author’s Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.
Psychometric properties of self-reported financial toxicity measures in cancer survivors: A systematic review

Zheng Zhu1,2, Weijie Xing1,2, Huan Wen3, Yanling Sun3, Winnie K. W. So4, Lizarondo Lucylynn5; Jian Peng1; Yan Hu1,2

Author affiliations:
1. School of Nursing, Fudan University, Shanghai, China (Zheng Zhu, Weijie Xing, Jian Peng, Yan Hu);
2. Fudan University Centre for Evidence-based Nursing: A Joanna Briggs Institute Centre of Excellence, Fudan University, Shanghai, China (Zheng Zhu, Weijie Xing, Yan Hu);
3. School of Public Health, Fudan University, Shanghai, China (Huan Wen, Yanling Sun);
4. The Nethersole School of Nursing, The Chinese University of Hong Kong, Hong Kong, China (Winnie K. W. So);
5. Joanna Briggs Institute, University of Adelaide, Adelaide, South Australia, Australia (Lizarondo Lucylynn).

Correspondence:
Weijie Xing, RN, PhD, School of Nursing, Fudan University; Fudan University Centre for Evidence-Based Nursing: A Joanna Briggs Institute Centre of Excellence, 305 Fenglin Rd., Shanghai, China (xingweijie@fudan.edu.cn)
Word count: 3335

Number of tables: 4

Number of figures: 1

Number of references: 58
Psychometric properties of self-reported financial toxicity measures in cancer survivors: A systematic review

Abstract

Objective: The aim of this systematic review was to summarize the psychometric properties of patient-reported outcome measures (PROMs) measuring financial toxicity (FT) in cancer survivors.

Design: This systematic review was conducted according to the guidance of the Consensus-Based Standards for the Selection of Health Measurement Instruments (COSMIN) methodology.

Data sources: Comprehensive searches were performed in PubMed, MEDLINE, Embase, CINAHL, PsycINFO, Web of Science, ProQuest, and Cochrane Library from database inception to February 2022.

Eligibility criteria for selecting studies: We included studies that reported any PROMs for measuring FT in cancer survivors who were ≥18 years old. FT was defined as perceived subjective financial distress resulting from objective financial burden. Studies that were not validation studies and that used a PROM only as an outcome measurement were excluded.
Data extraction and synthesis: Two reviewers independently extracted data from the included papers. We used the COSMIN criteria to summarize and evaluate the psychometric properties of each study regarding structural validity, internal consistency, reliability, measurement error, hypothesis testing for construct validity, cross-cultural validity/measurement invariance, criterion validity, and responsiveness.

Results: A total of 23 articles (21 PROMs) were eligible for inclusion in this study. The findings highlighted that the Comprehensive Score for Financial Toxicity (COST) had an adequate development process and showed better psychometric properties than other PROMs, especially in internal consistency (Cronbach’s α=0.92), reliability (intraclass correlation coefficient=0.80), and hypothesis testing (r=0.42-0.20).

Conclusions: From a psychometric property perspective, the COST could be recommended as the most suitable worldwide available measure for use in research and clinical practice across different contexts. We suggest that PROMs should be selected only after careful consideration of the local socioeconomic context. Future studies are warranted to develop various FT PROMs based on different social and cultural backgrounds and to clarify the theoretical grounds for assessing FT.

Keywords: cancer, financial toxicity, measurement, PROM, systematic review
Strength and limitation of this study

- This is the first systematic review that comprehensively summarized the psychometric properties of 21 PROMs evaluating financial toxicity in cancer survivors.

- The results may provide quantitative evidence for researchers and healthcare professionals to choose PROMs measuring cancer survivors’ financial toxicity in future scientific research and clinical practice.

- This review only included studies that aimed to evaluate the measurement properties of financial toxicity PROMs.
**Introduction**

The rising cost associated with advancements in cancer treatment and lengthening of cancer survivorship poses a significant challenge to survivors, caregivers, and public healthcare systems.\(^1\)\(^2\) Total global spending on cancer medications grows at a compound annual growth rate of 6.5%, growing from US$ 96 billion in 2013 to US$ 173 billion in 2020, which is nearly twice the rate of global gross domestic product (GDP) growth.\(^3\)\(^-\)\(^5\) The majority of cancer survivors in middle-income and low-income countries/regions depend on out-of-pocket payments, which may lead to global inequalities in healthcare expenditures and financial insecurity for vulnerable groups.\(^6\)\(^7\)

The term “financial toxicity (FT)” has been described as the economic effect of cancer treatment in the age of precision medicine.\(^2\)\(^,\)\(^8\)\(^,\)\(^9\) Witte and colleagues described FT as “the patient-reported outcome (PRO) of perceived subjective financial distress resulting from objective financial burden”.\(^10\) This concept covers both the objective financial burden and the subjective financial distress that cancer survivors face as a result of high out-of-pocket medical expenses. Regarding the terminology, “financial toxicity”, “financial burden” and “financial distress” are often used interchangeably in research and share a similar definition.\(^10\)\(^,\)\(^11\) In this review, the authors agreed to consistently use the term “financial toxicity.” Financial toxicity is usually measured by patient-reported outcome measures (PROMs); choosing a PROM with high validity and reliability is a prerequisite for robust results.
There are a few cancer-specific and generic FT PROMs that have been reported and used in different contexts. As one of the recent cancer-specific FT PROMs, the Comprehensive Score for Financial Toxicity (COST) is the most commonly used measure for assessing FT. In addition to COST, other cancer-specific measures have been widely used, including the Breast Cancer Finances Survey Inventory (BCFS), Socioeconomic Wellbeing Scale (SWBS), and InCharge Financial Distress/Financial Wellbeing Scale (InCharge). Additionally, validated subscales, such as the Social Difficulties Inventory Cancer Care Outcomes (SDI), the Cancer Care Outcomes Research and Surveillance Consortium patient survey (CanCORS), and Italian version of the Edmonton Symptom Assessment System (ESAS)-Total Care (TC), were also used to evaluate FT. However, existing PROMs vary significantly in their state of development and degree of validation, and many PROMs have not been psychometrically tested. A preliminary literature search was conducted in PubMed, PsycINFO (EBSCO), Cochrane Library (Wiley) and JBI (Ovid), which revealed that there exist some reviews regarding measures of FT. Witte and colleagues summarized the content of 352 items from 34 studies measuring FT in cancer survivors. However, this review did not report the psychometric properties of the included PROMs, and most of the included PROMs were not validated through a scientific process, which made it difficult for readers to choose the best measure from existing PROMs to evaluate the
level of FT. Salman and colleagues conducted a systematic review and found 8
PROMs and 2 caregiver-reported measures for assessing financial burden in
adolescents and young adults. However, this review focused only on PROMs
assessing FT in adolescents and young adults with cancer. The psychological
properties of FT measures in adult cancer survivors are still unknown.

The reproducibility, reliability, and accuracy of PROMs are the fundamental
premise for achieving robust results. Therefore, it is necessary to summarize the
psychometric properties of existing PROMs for future research. However, this
information is still lacking. The aim of this systematic review was to summarize the
psychometric properties of PROMs for measuring FT in cancer survivors. The review
was conducted according to the guidance of the Consensus-Based Standards for the
Selection of Health Measurement Instruments (COSMIN) methodology and the
Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA)
statement. The protocol of this review was published in BMJ OPEN in 2020. The
registration number of the protocol in PROSPERO was CRD42021254721.

Methods

Search strategy

First, we conducted a limited search via PubMed to capture keywords from
which to develop search strategies for each database. Subsequently, all identified
search strategies across databases were performed in PubMed/MEDLINE, MEDLINE
For peer review only

(Ovid), Embase (Ovid), CINAHL (EBSCO), PsycINFO (EBSCO), Web of Science, ProQuest Dissertations and Theses, and Cochrane Library (Wiley). The search time frame was set from database inception to February 2022. To include more studies published in 2021 and 2022, the end date of the search was updated to February 2022. In PubMed/Medline, we searched papers in English using MeSH terms ([cancer OR neoplasms] AND [“cancer survivors” OR patient OR survivors] AND “cost of illness”) combined with (cancer OR [patient* OR survivor*] AND [cost OR bill* OR expense OR productivity loss OR “out-of-pocket” OR “economic burden” OR “financial toxicity” OR “financial hardship” OR “financial burden”]). The COSMIN measurement properties filter and exclusion filter were also utilized in the search box. The search strategies for each database are presented in Appendix 1. Finally, the references of all included studies were manually reviewed to supplement the database search.

Inclusion and exclusion criteria

The inclusion criteria were as follows: 1) studies that reported any PROMs for measuring FT in cancer survivors who were ≥18 years old. If the studies reported results in a population combined with both ≥18- and <18-year-old cancer survivors and the majority of survivors were not < 18 years old, the studies were also considered; 2) studies that evaluated at least one measurement property; and 3) studies published in English. The exclusion criteria were as follows: 1) studies that were not validation studies and used a PROM only as an outcome measurement; 2) studies that
used a PROM as a comparator for another instrument; 3) studies that did not provide empirical data; and 4) if a measure was a quality of life PROM and had a domain that assessed FT, we included only the original version of the PROM. If the measure/domain included only one item and reported the measurement property as an independent domain, the measure/domain was also considered.

Study screening and selection

We imported all identified citations by search strategies into Endnote X8 (Clarivate Analytics, PA, USA). After duplicates were removed, two reviewers (ZZ & WX) independently screened all titles, abstracts, and full texts (ZZ & WX) based on the established inclusion and exclusion criteria. Any disagreements were resolved by a third reviewer (YH).

Quality appraisal

Two reviewers (WH & YS) assessed the methodological quality of the PROM of the included studies by using the COSMIN Risk of Bias Checklist (Appendix II). The checklist consisted of 10 domains (116 items), including PROM development, content validity, structural validity, internal consistency, cross-cultural validity, reliability, measurement error, criterion validity, hypothesis testing, and responsiveness. Each measurement property was rated as “very good”, “adequate”, “doubtful” or “inadequate quality”. According to the COSMIN guidelines, the methodological quality of a single study is rated based on the worst score count
method. For example, if the lowest rating is “inadequate” in the PROM development domain, the overall methodological quality of that domain is “inadequate”. The worst score counts method takes into account that inadequate quality items could affect the overall results of the measurement property of each PROM. Any discrepancies were resolved by a third reviewer (ZZ).

**Data extraction**

Two reviewers (ZZ & WX) independently extracted data from the included papers, including authors, year of publication, PROM, country/language, study design, target population, sample size, domains, number of items, total score range, and main findings. The main findings regarding psychometric properties, including content validity, structural validity, internal consistency, cross-cultural validity, reliability, measurement error, criterion validity, hypothesis testing, and responsiveness, were also extracted. Any discrepancies were resolved through discussion between the two reviewers.

**Data synthesis**

We used the COSMIN criteria to summarize and evaluate the psychometric properties of each study regarding structural validity, internal consistency, reliability, measurement error, hypothesis testing for construct validity, cross-cultural validity/measurement invariance, criterion validity, and responsiveness. Each measurement property from each study was rated as sufficient (+), insufficient (−), or
indeterminate (?). The criteria for the measurement property rating can be found in Appendix II. If the ratings of one psychometric property per study were all sufficient (+) or insufficient (-), the results were pooled, and the overall rating was rated as sufficient (+) or insufficient (-). If the ratings were inconsistent, explanations of inconsistency were explored (e.g., different languages). For example, in our review, different language, social, economic, and cultural contexts may contribute to inconsistencies in psychometric properties. Our review team (ZZ, WJ, HW, and YS) discussed the potential explanations of inconsistency. If the review team regarded the explanation as reasonable, we provided ratings (“+”, “-“, and “?”) in subgroups (e.g., language subgroup). If the explanation was not reasonable, the overall rating of this measurement property was rated as inconsistent (±).

Assessing certainty of evidence

We used a modified Grading of Recommendations Assessment, Development, and Evaluation (GRADE) system to assess the certainty of evidence. Each piece of evidence was graded for risk of bias, inconsistency, imprecision, and indirectness. The instructions for downgrading for risk of bias, inconsistency, imprecision, and indirectness are shown in Appendix II. Four reviewers (ZZ, WJ, HW, and YS) independently assessed the grade. Any discrepancies were resolved by discussion.

Patient and public involvement

No patients or the public were directly involved in the development of the
research question, selection of the outcome measures, design and implementation of the study, or interpretation of the results.

Results

Literature search

Figure 1 shows the process of literature screening and selection. A total of 9399 articles were identified via databases. Six articles were found by additional supplementary searches. After duplications were removed, a total of 11731 articles were retained, 11669 articles were deleted after reading the title and abstract, and 39 were deleted after full-text reading. Finally, a total of 23 articles (21 PROMs) were eligible for inclusion in this study.\textsuperscript{12,14,16,23-42}

Study description

Table 1 shows the characteristics of the included studies. All included studies were published from 2005 to 2022. Eight studies were conducted in the United States,\textsuperscript{12,14,23,27,30,37,39,41} four in the United Kingdom,\textsuperscript{16,29,35,38} two in Canada,\textsuperscript{31,36} and two in China (mainland and Hong Kong),\textsuperscript{25,39} India,\textsuperscript{26,34} and Italy.\textsuperscript{33,42} One study was conducted in 12 countries in Europe and North America.\textsuperscript{22,23} Other studies were conducted in Brazil\textsuperscript{32} and Iran\textsuperscript{34}. A total of 12362 participants were included, ranging from 7\textsuperscript{36} to 5901\textsuperscript{41} per study. The majority of studies assessed FT in multiple types of cancer. Only two studies focused on a single type of cancer, namely, lung, colorectal, or head and neck cancer.\textsuperscript{31,37}
Among the 21 PROMs, 7 were FT-related domains of quality of life PROMs, and 14 were independent PROMs focusing on FT. All PROMs were validated in cancer survivors. Fifteen PROMs were in English,\(^{12,14,16,23,25-31,35,37,38,40,41,42}\) and two were in Chinese.\(^{24,39}\) Other languages included French,\(^{36}\) Portuguese,\(^{32}\) Italian,\(^{33,42}\) Hindi,\(^{25,26}\) and Persian.\(^{34}\) The number of items evaluating FT ranged from 3\(^{40}\) to 23.\(^{36}\) The French version of the Patient Self-Administered Financial Effects Questionnaire (P-SAFE) did not report the total score range of the whole PROM.\(^{36}\)

**Quality assessment**

*Methodological quality assessment.* Table 2 shows the methodological quality of the 23 included studies by using the COSMIN checklist. In the PROM development domain, only one study was rated as adequate,\(^ {42}\) three studies were rated as doubt,\(^ {12,24,27,29}\) and the others were rated as inadequate. Two studies reported adequate information in testing the relevance, comprehensiveness, and comprehensibility of PROMs.\(^ {12,27,29}\) One study reported adequate relevance and comprehensiveness.\(^ {42}\) Among all studies, the most reported domain was internal consistency, except one study.\(^ {36}\) Limited information could be retrieved on cross-cultural validity (three studies),\(^ {31,32,36}\) criterion validity (six studies),\(^ {16,23,33,35,38,40}\) reliability (ten studies),\(^ {12,16,24,27,28,33,35,38,39,40,42}\) and responsiveness (two studies).\(^ {31,39}\) No data were identified as measurement error.
Measurement property assessment. Table 3 shows the quality of the psychometric properties retrieved from 21 PROMs. Only the Persian version of the COST-v2 and Subjective Financial Distress Questionnaire (SFDQ) were rated as “+” in structural validity. There were 17 PROMs rated as “+” in internal consistency. Eight PROMs were rated as “+” in reliability. Ten PROMs were rated as “+” in hypothesis testing. Limited information was retrieved on cross-cultural validity (two PROMs), criterion validity (six PROMs), and responsiveness (two PROMs). No PROMs reported data on measurement error.

Certainty of evidence

Table 4 shows the certainty of evidence for each measurement property. Among all included PROMs, the COST showed the best psychometric properties compared to other measures. The COST and its seven versions were rated as having high evidence of structural validity, internal consistency, hypothesis testing, and criterion validity. The Financial Index of Toxicity (FIT) and Impact of Cancer – Childhood Survivors (IOC-CS) financial problems domain reported data on five properties and were rated on a scale from “very low evidence” to “high evidence.”

Discussion

This systematic review identified 21 PROMs and domains of PROMs evaluating
302 FT in cancer survivors, including the COST (original, Brazilian, India, Italian, Persian, Simplified Chinese, Traditional Chinese version), FIT, Personal Financial Burden (PFB), P-SAFE, SWBS, Quality of Life in Adult Cancer Survivors (QLACS) financial problems domain, Chronic Cancer Experiences Questionnaire (CCEQ) financial advice domain, Patient-Reported Outcome for Fighting Financial Toxicity (PROFFIT), Patient Roles and Responsibilities Scale (PRRS) financial well-being domain, SDI-21 providing for the family domain, SDI-16 money matters domain, Subjective Financial Distress Questionnaire (SFDQ), IOC-CS financial problems domain, and Cancer Problems in Living Scale (CPILS) employment/financial domain.

303 Overall, the COST had a complete development process compared to other PROMs and showed the best psychometric properties, especially in terms of internal consistency, reliability, and hypothesis testing. To the best of our knowledge, this is the first systematic review that has summarized the psychometric properties of FT PROMs in cancer survivors and reported the certainty of evidence for each property of PROMs. The results may provide quantitative evidence for researchers and healthcare professionals to choose PROMs measuring cancer survivors’ FT in future scientific research and clinical practice.

304 The results highlighted that the COST (of which we studied both version 1 and version 2) had better psychometric properties than other specific and generic PROMs in terms of internal consistency, reliability, and hypothesis testing. The COST could be recommended as the most suitable worldwide available measure for use in research.
and clinical practice across different contexts. Other systematic reviews have also
suggested that the COST is a promising measure from a content perspective\textsuperscript{10,11}. From
a psychometric standpoint, there are a few issues that one must face when evaluating
financial toxicity in cancer survivors using the COST. First, caution should be taken
when using the COST in different socioeconomic conditions outside the United
States. In some countries in Europe or Asia, the majority of medical expenses are
covered by social health insurance, and direct out-of-pocket payments are replaced by
prepayment from health insurance contributions.\textsuperscript{43,44} In addition, social security
systems can benefit cancer survivors who are not able to work.\textsuperscript{45} These two
socioeconomic factors may affect cancer survivors’ understanding regarding some
items related to medical spending and indirect cost. However, few COST validation
studies have considered socioeconomic issues, adapted the measure in a local context,
or provided data on cross-cultural validity. It is recommended that future COST
validation studies recruit cancer survivors across multiple social and cultural
backgrounds to assess cross-cultural measurement invariance.

Second, the original construct and item generation for the COST were based on a
literature search; thus, the theoretical grounds for the measure are unclear, and the
instrument may not capture detailed information related to the construct. Theoretical
frameworks and conceptual models are crucial for self-reported measures to capture
subtle changes in constructs.\textsuperscript{46} Although FT is a relatively new concept, certain
models can guide item generation in the development of future FT PROMs.
Tucker-Seeley and colleagues developed a conceptual model of FT and emphasized three components of financial burden, namely, the material, psychosocial, and behavioral domains. Head developed SWBS based on James Coleman’s Theory of Social Class; this scale contains 17 items across 3 domains: human capital, material capital, and social capital. Witte and colleagues’ systematic review analyzed 352 different questions regarding financial spending and found six domains (financial spending, financial resources, psychosocial affect, support seeking, coping care, and coping lifestyle) that can represent reactions to subjective financial distress. Other theories and models, including the Wreckers theory of financial distress, ecological theory, and the functionalist tradition, have also been widely used in cancer survivors. With the increasing number of theoretical studies related to FT, the theoretical grounds for future PROMs need to be clarified.

In addition to the COST, two other PROMs, namely, the FIT and the IOC-CS financial problems domains, also provided adequate data on psychometric properties. The FIT is relatively new and has fewer items than the other included measures. This measure was developed by Hueniken and colleagues and has been validated only in survivors with head and neck cancer. Head and neck cancer, especially laryngeal and hypopharyngeal cancer, has particularly large impacts on survivors’ daily function (e.g., speech and eating) after treatment and affects survivors’ ability to return to work. Only 32% to 59% of head and neck cancer survivors return to work after treatment. This form of cancer also has short- and long-term financial
consequences for caregivers and their families. Therefore, future studies should be aware that the FIT may not be directly applicable to other cancer populations.

Regarding PROM development, we found that only two PROMs, PROFFIT and SFDQ, were not developed in the context of English-speaking developed countries such as the United States, the United Kingdom, and Canada. The socioeconomic contexts and healthcare systems in these countries may be significantly different from those in other parts of the world and ultimately lead to a nuance in the perceived causes and consequences of FT. Previous studies have reported that FT is closely related to broad social determinants of economic circumstances. Factors including healthcare policy, healthcare system, insurance system, specific micro contexts, and the level of regional economic development could not only affect the cancer survivors’ perceived level of FT but also determine the origins of FT. Additionally, cultural factors (e.g., a cultural emphasis on saving and a cultural imperative to have a large family) also affect cancer survivors’ perceived financial security and economic burden.

PROFFIT, which was developed in 2021 in the Italian context, also reported higher quality PROM development and content validity than other PROMs. We would consider it to be a good FT PROM against the COSMIN criteria if more validation studies were conducted to report a greater effect size of the measurement properties. Therefore, we recommend that researchers use context-specific measures
to assess FT in cancer survivors (e.g., using PROFFIT in Italy). Further studies are
warranted to develop various FT PROMs based on different social and cultural
backgrounds. Worldwide measures, such as COST, should be analyzed to determine
the differences between social, cultural, and economic contexts.

Limitations

We acknowledge that there are some limitations to this study. First, this review
included only studies that aimed to evaluate the measurement properties of FT
PROMs. Many studies that aimed to explore the level of FT in cancer survivors also
reported the reliability and validity of PROMs. Therefore, the PROMs we
summarized in this systematic review had higher psychometric quality than other
measures that we did not list in this review. Second, we included only studies
published in English. Therefore, studies published in other languages were not
included, which may affect the conclusion of this review. Third, we included only the
original version of the FT domain from PROMs assessing quality of life in cancer
survivors, such as the EORTC QLQ-C30 and the QLACS. Over 20 language versions
of these PROMs do not provide sufficient details on the FT domain individually.

Conclusion

This systematic review summarized the psychometric properties of 20 PROMs
evaluating FT in cancer survivors. The findings highlighted that, from a psychometric
property perspective, the COST had an adequate PROM development process and
showed the best psychometric properties among all examined PROMs, especially in internal consistency, reliability, and hypothesis testing; thus, we recommend the COST as the most suitable worldwide available measures for use in research and clinical practice across different contexts. The FIT and the IOC-CS financial problems domain also had adequate psychometric properties. We suggest that PROMs should be selected only after careful consideration of the local socioeconomic context. Future studies are warranted to develop various FT PROMs based on different social and cultural backgrounds and a clear theoretical basis for assessing FT.

Acknowledgements

Contributors

ZZ and WX designed the systematic review, conducted data searching, extraction and analysis, assessing the certainty of evidence, and wrote the draft of the manuscript. HW and YS conducted quality appraisal and assessing the certainty of evidence. WS, LL, JP, and YH provided critical comments. All authors approved the final version of the manuscript.

Funding

This work was supported by the Shanghai Pujiang Program (Grant Number: 2019PJC017), Shanghai Soft Science Key Program (Grant Number: 20692104800), National Natural Science Foundation of China (Grant Number: 72004034), China Medical Board Open Competition Program (Grant Number: 20-371).
Competing interests

None.

Patient and public involvement

Patients and the public were not involved in the design, or conduct, or reporting, or dissemination plans of our research.

Patient consent for publication

Not applicable.

Ethics approval

This systematic review did not require ethical approval.

Data available statement

No additional data are available.

Reference

1. Bosland MC. Global health and cancer. Lancet. 2019;393(10175):983.

2. De Souza JA, Wong YN. Financial distress in cancer patients. J Med Person. 2013;11:73-77.

3. World cancer report [Internet]. [cited 2021 July 31]. 2020. Available from:
456  https://publications.iarc.fr/586.

457  4. Global Oncology Trends 2021 [Internet]. [cited 2021 July 31]. 2021. Available
from:

459  https://www.iqvia.com/insights/the-iqvia-institute/reports/global-oncology-trends
-2021

461  5. Mariotto AB, Yabroff KR, Shao Y, et al. Projections of the cost of cancer care in
the United States: 2010-2020. J Natl Cancer Inst. 2011;103:117-128.

463  6. ACTION Study Group. Financial catastrophe, treatment discontinuation and death
associated with surgically operable cancer in South-East Asia: Results from the
ACTION Study. Surgery. 2015;157:971-982.

466  7. Mackintosh M, Channon A, Karan A, et al. What is the private sector?
Understanding private provision in the health systems of low-income and
middle-income countries. Lancet. 2016;388:596-605.

469  8. Yousuf Zafar S. Financial toxicity of cancer care: It’s time to intervene. J Natl
Cancer Inst. 2015;108:djv370.

471  9. Carrera PM, Kantarjian HM, Blinder VS. The financial burden and distress of
patients with cancer: understanding and stepping-up action on the financial
toxicity of cancer treatment. CA Cancer J Clin. 2018,68:153-165

474  10. Witte J, Mehlis K, Surmann B, et al. Methods for measuring financial toxicity
after cancer diagnosis and treatment: a systematic review and its implications.
Ann Oncol. 2019;30:1061-1070.

477  11. Gordon LG, Merollini KM, Lowe A, Chan RJ. A systematic review of financial
toxicity among cancer survivors: we can’t pay the co-pay. Patient. 2017;10:295-309.

12. De Souza JA, Yap BJ, Hlubocky FJ et al. The development of a financial toxicity patient-reported outcome in cancer: the COST measure. Cancer. 2014;120:3245-3253.

13. Given BA, Given CW, Stommel M. Family and out-of-pocket costs for women with breast cancer. Cancer Pract. 1994;2:187-193.

14. Head BA, Faul AC. Development and validation of a scale to measure socioeconomic well-being in persons with cancer. J Support Oncol. 2008;6(4):183-192.

15. Prawitz AD, Garman ET, Sorhaindo B et al. InCharge financial distress/financial well-being scale: development, administration, and score interpretation. Financial Counsel Plann. 2006;17:34–50.

16. Wright P, Smith AB, Keding A, Velikova G. The social difficulties inventory (SDI): development of subscales and scoring guidance for staff. Psychooncology. 2011;20:36-43.

17. Malin JL, Ko C, Ayanian JZ et al. Understanding cancer patients’ experience and outcomes: development and pilot study of the Cancer Care Outcomes Research and Surveillance patient survey. Support Care Cancer. 2006;14:837-848.

18. Ripamonti C, Leporati R, De Feo G, et al. Italian version of the Edmonton Symptom Assessment System (ESAS)-Total Care (TC): development and psychometric validation in patients undergoing cancer treatment or follow-up.
19. Salsman JM, Danhauer SC, Moore JB, et al. Systematic review of financial burden assessment in cancer: Evaluation of measures and utility among adolescents and young adults and caregivers. Cancer. 2021;127(11):1739-1748.

20. Prinsen CAC, Mokkink LB, Bouter LM, Alonso J, Patrick DL, de Vet HCW, Terwee CB. COSMIN guideline for systematic reviews of patient-reported outcome measures. Qual Life Res. 2018;27(5):1147-1157.

21. Page MJ, McKenzie JE, Bossuyt PM , et al. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. BMJ. 2021;372:n71.

22. Zhu Z, Xing W, Lizarondo L, Peng J, Hu Y, So WK. Psychometric properties of self-reported financial toxicity measures in cancer survivors: a systematic review protocol using COSMIN methodology. BMJ open. 2020;10(5):e036365.

23. Avis NE, Smith KW, McGraw S, Smith RG, Petronis VM, Carver CS. Assessing quality of life in adult cancer survivors (QLACS). Qual Life Res. 2005;14(4):1007-1023.

24. Chan DN, Choi KC, Ng MS, et al. Translation and validation of the Traditional Chinese version of the COmprehensive Score for financial Toxicity-Functional Assessment of Chronic Illness Therapy (Version 2). Health Qual Life Outcomes. 2021;19(1):17.

25. Dar MA, Chauhan R, Sharma KK, Trivedi V, Dhirila S, Murti K. Assessing the reliability and validity of comprehensive score for financial toxicity (COST) among radiation oncology patients in India: a cross-sectional pilot study.
26. Dar MA, Chauhan R, Murti K, Trivedi V, Dhingra S. Development and Validation of Subjective Financial Distress Questionnaire (SFDQ): A Patient Reported Outcome Measure for Assessment of Financial Toxicity Among Radiation Oncology Patients. Front Oncol. 2022;11:819313.

27. De Souza JA, Yap BJ, Wroblewski K, et al. Measuring financial toxicity as a clinically relevant patient-reported outcome: the validation of the COmprehensive Score for financial Toxicity (COST). Cancer. 2017;123(3):476-484.

28. Durber K, Halkett GK, McMullen M, Nowak AK. Measuring financial toxicity in Australian cancer patients–Validation of the COmprehensive Score for financial Toxicity (FACT COST) measuring financial toxicity in Australian cancer patients. Asia Pac J Clin Oncol. 2021;17(4):377-387.

29. Harley C, Pini S, Kenyon L, Daffu-O'Reilly A, Velikova G. Evaluating the experiences and support needs of people living with chronic cancer: development and initial validation of the Chronic Cancer Experiences Questionnaire (CCEQ). BMJ Support Palliat Care. 2019;9(1):e15.

30. Head BA. Development and validation of a scale to measure health-related socioeconomic well-being in persons with a cancer diagnosis. Electronic Theses and Dissertations. Paper 594. https://doi.org/10.18297/etd/594

31. Hueniken K, Douglas CM, Jethwa AR, et al. Measuring financial toxicity incurred after treatment of head and neck cancer: Development and validation of the...
Financial Index of Toxicity questionnaire. Cancer. 2020;126(17):4042-4050.

32. de Alcantara Nogueira L, Koller FJ, Marcondes L, et al. Validation of the comprehensive score for financial toxicity for Brazilian culture. Ecancermedicalscience. 2020;14:1158.

33. Ripamonti CI, Chiesi F, Di Pede P, et al. The validation of the Italian version of the COmprehensive Score for financial Toxicity (COST). Supportive Care Cancer. 2020;28(9):4477-4485.

34. Sharif SP, Sim OF, Moshtagh M, Naghavi N, Nia HS. Psychometric evaluation of the COmprehensive Score for financial Toxicity scale among Iranian cancer patients. Research Square. 2020.

35. Shilling V, Starkings R, Jenkins V, Cella D, Fallowfield L. Development and validation of the patient roles and responsibilities scale in cancer patients. Qual Life Res. 2018;27(11):2923-2934.

36. Tremblay D, Poder TG, Vasiliadis HM, Touati N, Fortin B, Lévesque L, Longo C. Translation and Cultural Adaptation of the Patient Self-Administered Financial Effects (P-SAFE) Questionnaire to Assess the Financial Burden of Cancer in French-Speaking Patients. Healthcare (Basel). 2020;8(4):366.

37. Veenstra CM, Regenbogen SE, Hawley ST, et al. A composite measure of personal financial burden among patients with stage III colorectal cancer. Med Care. 2014;52(11):957-962.

38. Wright EP, Kiely M, Johnston C, Smith AB, Cull A, Selby PJ. Development and evaluation of an instrument to assess social difficulties in routine oncology.
practice. Qual Life Res. 2005;14(2):373-386.

39. Yu HH, Yu ZF, Li H, Zhao H, Sun JM, Liu YY. The COmprehensive Score for financial Toxicity in China: Validation and Responsiveness. J Pain Symptom Manage. 2021;61(6):1297-1304.e1.

40. Zebrack BJ, Donohue JE, Gurney JG, Chesler MA, Bhatia S, Landier W. Psychometric evaluation of the impact of cancer (IOC-CS) scale for young adult survivors of childhood cancer. Qual Life Res. 2010;19(2):207-218.

41. Zhao L, Portier K, Stein K, Baker F, Smith T. Exploratory factor analysis of the Cancer Problems in Living Scale: a report from the American Cancer Society's Studies of Cancer Survivors. J Pain Symptom Manage. 2009;37(4):676-686.

42. Riva S, Arenare L, Di Maio M, et al. Cross-sectional study to develop and describe psychometric characteristics of a patient-reported instrument (PROFFIT) for measuring financial toxicity of cancer within a public healthcare system. BMJ Open. 2021;11(10):e049128.

43. De Lorenzo F, Apostolidis K, Florindi F, Makaroff LE. Improving European policy to support cancer survivors. J Cancer Policy. 2018;15:72-75.

44. Bhoo-Pathy N, Ng CW, Lim GC, et al. Financial toxicity after cancer in a setting with universal health coverage: A call for urgent action. J Oncol Pract. 2019;15(6):e537-e546.

45. Philip CC, Mathew A, John MJ. Cancer care: Challenges in the developing world. Cancer Res Stat Treat. 2018;1(1):58.

46. Mokkink LB, Boers M, van der Vleuten CPM, et al. COSMIN Risk of Bias tool to
assess the quality of studies on reliability or measurement error of outcome measurement instruments: a Delphi study. BMC Med Res Methodol. 2020;20:293.

47. Tucker-Seeley RD, Thorpe Jr RJ. Material–psychosocial–behavioral aspects of financial hardship: a conceptual model for cancer prevention. Gerontologist. 2019;59(Suppl 1):S88-S93.

48. Coleman JS. Foundations of social theory. Harvard university press; 1994.

49. A 'Wreckers Theory' of Financial Distress. [Internet]. [cited 2021 July 31]. 2005. Available from: https://www.econstor.eu/bitstream/10419/19625/1/200540dkp.pdf

50. Bergeron K, Abdi S, DeCorby K, Mensah G, Rempel B, Manson H. Theories, models and frameworks used in capacity building interventions relevant to public health: a systematic review. BMC Public Health. 2017;17(1):914.

51. Acker J, Baca-Zinn M, Collins PH, et al. Social class and stratification: Classic statements and theoretical debates. Rowman & Littlefield Publishers; 2006.

52. Miller A. Returning to work after head and neck cancer. Curr Opin Otolaryngol Head Neck Surg. 2020;28(3):155-160.

53. Koch R, Wittekindt C, Altendorf–Hofmann A, Singer S, Guntinas–Lichius O. Employment pathways and work-related issues in head and neck cancer survivors. Head Neck. 2015;37(4):585-593.

54. Giuliani M, Papadakos J, Broadhurst M, et al. The prevalence and determinants of return to work in head and neck cancer survivors. Support Care Cancer.
55. Dewa CS, Trojanowski L, Tamminga SJ, Ringash J, McQuestion M, Hoch JS. Work-related experiences of head and neck cancer survivors: an exploratory and descriptive qualitative study. Disabil Rehabil. 2018;40(11):1252-1258.

56. Zhu Z, Xing W, Zhang X, Hu Y, So WK. Cancer survivors' experiences with financial toxicity: A systematic review and meta-synthesis of qualitative studies. Psychooncology. 2020;29(6):945-959.

57. Chebli P, Lemus J, Avila C, et al. Multilevel determinants of financial toxicity in breast cancer care: perspectives of healthcare professionals and Latina survivors. Support Care Cancer. 2020;28(7):3179-3188.

58. Kong YC, Wong LP, Ng CW, et al. Understanding the financial needs following diagnosis of breast cancer in a setting with universal health coverage. Oncologist. 2020;25(6):497-504.
Table 1 Overview of the included studies

| Author (year), country | PROM | Country | Language(s) of PROM | Study design | Target population | Sample size | Measurement domain | Number of items | Total score range |
|------------------------|------|---------|---------------------|--------------|-------------------|-------------|--------------------|-----------------|------------------|
| Avis et al., 2005      | QLACS financial problems domain | US | English | Development and validation study | Age: 71.4±11.5 yrs Male: 42% Cancer type: Breast, bladder, head and neck, gynecologic, prostate, colorectal cancer | 242 | 1 domain regarding financial problems | 4 | 4-28 |
| Chan et al., 2021      | COST-v2, Traditional Chinese version | Hong Kong | Traditional Chinese | Validation study | Age: 59.9±11.1 yrs Male: 35.3% Cancer type: Breast, gynecological, head and neck, gastric and colorectal, genitourinary, lung, hematological, skin, bone and soft tissue, brain and central nervous system cancer and others | 640 | No subdomain | 12 | 0-44 |
| Dar et al., 2021       | COST-v1, India version | India | Hindi or English | Validation study | Age: 49.5±16.8 yrs Male: 82.8% Cancer type: Tongue, gingival buccal sulcus, buccal mucosa, supraglottic larynx, hypopharynx, parotid and others | 29 | No subdomain | 11 | 0-44 |
| Dar et al., 2021       | SFDQ | India | Indian or English | Development and validation study | Age: 18-59 68.3%; ≥60 31.7% Male: 85.9% Cancer type: head and neck cancer | 142 | Five domains: financial resources; financial spending; psychosocial affect; coping care & coping lifestyles; support | 14 | 0-28 |
| Study | Version | Country | Language | Type | Description | Mean Age | Male (%) | Cancer Type | Subdomains |
|-------|---------|---------|----------|------|-------------|----------|----------|-------------|------------|
| de Alcantara Nogueira et al., 2020 | COST-v1, Brazilian version | Brazil | Brazilian Portuguese | Validation study | Mean age: 56 yrs  Male: 40.5%  Cancer type: Not specified | 126 | No subdomain | 11 | 0-44 |
| De Souza et al., 2014 | COST-v1 | US | English | Development study and validation study | Age: 58.4±11.5 yrs  Male: 41.6%  Cancer type: Not specified (diagnosis of AJCC stage IV cancer) | 233 | No subdomain | 11 | 0-44 |
| De Souza et al., 2017 | COST-v1 | Australia | English | Validation study | Age: ≤50 yrs 23%; 51-64 yrs 30%; ≥65 yrs 48%  Male: 46%  Cancer type: Thoracic, breast, sarcoma, skin, central nervous system, gynecological, head and neck, colorectal, upper gastrointestinal, urological, and miscellaneous cancer | 257 | No subdomain | 11 | 0-44 |
| Harley et al., 2019 | CCEQ financial advice domain | UK | English | Development and validation study | Pilot study:  Age: 65 (41-90) yrs  Male: 48.5%  Cancer type: Breast, colorectal/gastrointestinal, gynecological, prostate, and renal cancer  Final study:  Age: 67 (41-88) yrs  Male: 50.0%  Cancer type: Breast, colorectal/gastrointestinal, gynecological, prostate, and renal cancer | 103 for pilot study 313 for final study | 1 domain regarding financial advice | 5 | 5-25 |
| Study                                      | Short Name | Country     | Language     | Type                        | Age Description                  | Sample Size | Subdomains | Range |
|--------------------------------------------|------------|-------------|--------------|-----------------------------|----------------------------------|-------------|------------|-------|
| Head, 2008                                 | SWBS       | US          | English      | Development and validation study | Age: 59.6±12.7 yrs Male: 35.7%  | 266         | Two domains: material and social capital. | 0-68   |
| Head & Faul, 2008                          |            |             |              |                             | Cancer type: Breast, melanoma, head and neck, prostate, rectum/anus, colon, endometrium, lung/tracheal/bronchus, and non-Hodgkin’s lymphoma |            |            |       |
| Hueniken et al., 2020                      | FIT        | Canada      | English      | Development and validation study | Age: 61.6 (25.5-88.5) yrs Male: 77.2%  | 430         | No subdomain | 0-100 |
|                                          |            |             |              |                             | Cancer type: Oropharynx, oral cavity, larynx, nasopharynx, laryngopharynx cancers, and others |            |            |       |
| Ripamonti et al., 2020                     | COST-v2, Italian version | Italy | Italian | Validation study | Age: 61.5±12.7 yrs Male: 52.5%  | 118         | No subdomain | 0-44  |
|                                          |            |             |              |                             | Cancer type: Breast, lung, colon, gastric, hepatocellular, endometrial, prostate, sarcoma, bladder, head and neck, Hodgkin lymphoma, non-Hodgkin lymphoma, leukemia, myeloma, and others |            |            |       |
| Riva et al., 2021                          | PROFFIT    | Italy       | Italian or English | Development and validation study | Age: 29-82 yrs Male: 41.3%  | 184         | No subdomain | 0-100 |
|                                          |            |             |              |                             | Cancer type: breast, lower gastrointestinal tract, genitourinary, thoracic, upper gastrointestinal tract, and others |            |            |       |
| Sharif et al., 2020                        | COST-v2, Persian version | Iran  | Persian    | Validation study | Age: 50.0±14.3 yrs Male: 51.0%  | 398         | No subdomain | 0-44  |
|                                          |            |             |              |                             | Cancer type: Not specific |            |            |       |
| Shilling et al., 2018                      | PRRS financial wellbeing domain | UK    | English    | Development and validation study | Age: ≤50 yrs 25%; 51-65 yrs 41%; ≥66 yrs 34% Male: 23%  | 135         | 1 domain regarding financial wellbeing | 0-24  |
|                                          |            |             |              |                             | Cancer type: breast, gynecological, lung, and melanoma cancers |            |            |       |
| Study Authors and Year | Instrument and Language Version | Country | Language | Study Type | Age Details | Male (%) | Cancer Types | N | Subdomains | Score | HR or Other |
|------------------------|---------------------------------|---------|----------|------------|-------------|----------|--------------|---|------------|-------|-------------|
| Tremblay et al., 2020  | P-SAFE, French Version          | Canada  | French   | Cross-adaption study | Age: 50-59 yrs 57%; 60-69 yrs 29%; ≥70 yrs 34%  
Male: 14%  
Cancer type: colorectal, lung, breast, and prostate cancer. | | | | | | | 7 | NR | NR |
| Veenstra et al., 2014  | PFB                             | US      | English  | Validation study | Age: <50 yrs 17%; 50-64 yrs 37%; 65-74 yrs 23%; >75 yrs 24%  
Male: 53%  
Cancer type: Stage III colorectal cancer | | | | | | | 956 | No subdomain | 7 | 0-7 |
| Wright et al., 2005    | SDI-21 providing for the family domain | UK      | English  | Development and validation study | Age: 53.8±14.1 yrs  
Male: Not specific  
Cancer type: brain, lung cancers and others | | | | | | | 271 | 1 domain regarding providing for the family | 5 | 0-20 |
| Wright et al., 2011    | SDI-16 money matters domain     | UK      | English  | Development and validation study | Age: 56 (18-88) yrs for men; 56 (21-88) yrs for women  
Male: 48%  
Cancer type: breast, gastrointestinal, hematological, gynecological, germ cell, head and neck, lung, genitourinary, and others | | | | | | | 652 | 1 domain regarding money matters | 5 | 0-20 |
| Yu et al., 2021        | COST-v1, Simplified Chinese version | Mainland China | Chinese | Validation study | Age: 57.0±9.2 yrs  
Male: 45.7%  
Cancer type: lung, stomach, colorectal, and breast cancer | | | | | | | 440 | No subdomain | 11 | 0-44 |
| Zebrack et al., 2010   | IOC-CS financial problems domain | US      | English  | Validation study | Age: 26.7±5.3 yrs  
Male: 48.0%  
Cancer type: hematological, brain, and solid tumors/soft tissue tumors | | | | | | | 519 | 1 domain regarding financial problems | 3 | 1-15 |
| Zhao et al., 2009      | CPILS employment/financial domain | US      | English  | Validation study | Age: 55 yrs 48.8%; >55 51.2%  
Male: 41.6%  
Cancer type: breast, prostate, colorectal, bladder, uterine, kidney, lung, and ovarian cancer | | | | | | | 5901 | 1 domain regarding employment/finances | 6 | 0-12 |
melanoma of skin; non-Hodgkin lymphoma.

CCEQ: Chronic Cancer Experiences Questionnaire; COST: Comprehensive Score for Financial Toxicity; CPILS: Cancer Problems in Living Scale; FIT: Financial Index of Toxicity; IOC-CS: Impact of Cancer-Childhood Survivors; PFB: Personal Financial Burden; PROFFIT: Patient-Reported Outcome for Fighting Financial Toxicity; PROM: Patient-Reported Outcome Measures; PRRS: Patient Roles and Responsibilities Scale; P-SAFE: Patient Self-Administered Financial Effects questionnaire; QLACS: Quality of Life in Adult Cancer Survivors; SDI: Social Difficulties Inventory Cancer Care Outcomes; SFDQ: Subjective Financial Distress Questionnaire; SWBS: Socioeconomic Wellbeing Scale
## Table 2 Methodological quality assessment of the measures

| Author (year) | PROM development | Content validity | Structural validity | Internal consistency | Cross-cultural validity | Reliability | Measurement error | Criterion validity | Hypothesis testing | Responsiveness |
|---------------|------------------|------------------|--------------------|---------------------|------------------------|-------------|-------------------|-------------------|-------------------|----------------|
| Avis et al., 2005 | QLACS financial problems domain | Inadequate | R: NR | C1: NR | C2: NR | Adequate | Very good | NR | NR | Adequate | Doubtful | NR |
| Chan et al., 2021 | COST-v2, Traditional Chinese version | Doubtful | R: Doubtful | C1: NR | C2: Doubtful | Very good | Very good | NR | Doubtful | NR | NR | Very good | NR |
| Dar et al., 2021 | COST-v1, India version | Inadequate | R: NR | C1: NR | C2: NR | Inadequate | Very good | NR | NR | NR | NR | NR | NR |
| Dar et al., 2021 | SFDQ | Inadequate | R: Adequate | C1: NR | C2: NR | Very good | Very good | NR | NR | NR | NR | NR | NR |
| de Alcantara, 2020 | COST-v1, Brazilian version | Inadequate | C1: NR | C2: NR | R: NR | Very good | Very good | Inadequate | NR | NR | NR | NR | NR | NR |
| De Souza et al., 2014 | COST-v1 | Doubtful | R: Adequate | C1: Adequate | C2: Adequate | NR | Very good | NR | Adequate | NR | NR | Very good | NR |
| Durber et al., 2021 | COST-v1, Australia version | Inadequate | R: NR | C1: NR | C2: NR | NR | Very good | NR | Adequate | NR | NR | Very good | NR |
| Reference           | Scale/Content                     | R: Adequate | C1: Adequate | C2: Adequate | Adequate | Very good | NR | NR | NR | NR | Very good | NR |
|---------------------|-----------------------------------|-------------|--------------|--------------|----------|-----------|----|----|----|----|-----------|----|
| Harley et al., 2019 | CCEQ financial advice domain      | Doubtful    | Adequate     | Adequate     | Adequate | Very good | NR | NR | NR | NR | Very good | NR |
| Head, 2008          | SWBS                              | Inadequate  | R: NR        | C1:Doubt     | C2: NR   | Very good | Very good | NR | NR | NR | Very good | Very good | NR |
| Head & Faul, 2008   |                                   | Inadequate  | R: NR        | C1: NR       | C2: NR   | Adequate  | Very good | Doubtful | NR | NR | NR | Very good | Very good | NR |
| Hueniken et al., 2020 | FIT                              | Inadequate  | R: NR        | C1: NR       | C2: NR   | Very good | NR | NR | NR | NR | Very good | NR |
| Ripamonti et al., 2020 | COST-v2, Italian version           | Inadequate  | R: NR        | C1: NR       | C2: NR   | Very good | NR | NR | NR | NR | Very good | NR |
| Riva et al., 2021   | PROFFIT                           | Adequate    | R: Adequate  | C1: Adequate | Adequate | Adequate | NR | Adequate | NR | NR | NR | NR | NR |
| Sharif et al., 2020 | COST-v2, Persian version          | Inadequate  | R: NR        | C1: NR       | C2: NR   | Very good | Very good | NR | NR | NR | NR | NR | NR |
| Shilling et al., 2018 | PRRS financial wellbeing domain | Inadequate  | R: NR        | C1: NR       | C2: NR   | Adequate  | Very good | NR | Inadequate | NR | Inadequate | NR | NR | NR |
| Tremblay et al., 2020 | P-SAFE, French Version               | Inadequate  | R: NR        | C1: NR       | C2: NR   | NR | NR | Inadequate | NR | NR | NR | NR | NR | NR |
| Veenstra et al., 2014 | PFB                              | Inadequate  | R: NR        | C1: NR       | C2: NR   | Very good | Inadequate | NR | NR | NR | NR | NR | NR |
| Wright et al., 2005 | SDI-21 providing for the family domain | Inadequate | R: Doubt     | C1: NR       | C2: NR   | Very good | Very good | NR | NR | NR | NR | NR | NR |
| Wright et al., 2011 | SDI-16 money matters              | Inadequate  | R: Doubt     | C1: NR       | C2: NR   | Very good | Very good | Inadequate | NR | NR | NR | NR |

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml
| Domain                        | C2: NR | R: NR | C1: NR | C2: NR | Very good | Very good | NR      | Very good | NR      | NR      | Doubt    | Doubt    |
|-------------------------------|--------|-------|--------|--------|-----------|-----------|---------|-----------|---------|---------|----------|----------|
| Yu et al., 2021               | COST-v1, *Simplified Chinese version* | Inadequate | Very good | Very good | NR      | Very good | NR      | NR      | Doubt    | Doubt    |
| Zebrack et al., 2010          | IOC-CS financial problems domain | Inadequate | Adequate | Very good | NR      | Adequate  | NR      | Inadequate | Very good | NR      |
| Zhao et al., 2009             | CPILS employment/financial domain | Inadequate | Very good | Very good | NR      | NR      | NR      | NR      | Very good | NR      |

CCEQ: Chronic Cancer Experiences Questionnaire; COST: Comprehensive Score for Financial Toxicity; CPILS: Cancer Problems in Living Scale; FIT: Financial Index of Toxicity; IOC-CS: Impact of Cancer-Childhood Survivors; NA: Not Applicable; NR: Not Report; PFB: Personal Financial Burden; PROFFIT: Patient-Reported Outcome for Fighting Financial Toxicity; PROM: Patient-Reported Outcome Measures; PRRS: Patient Roles and Responsibilities Scale; P-SAFE: Patient Self-Administered Financial Effects questionnaire; QLACS: Quality of Life in Adult Cancer Survivors; SDI: Social Difficulties Inventory; SFDQ: Subjective Financial Distress Questionnaire; SWBS: Socioeconomic Wellbeing Scale.
## Table 3 Rating of measurement properties

| PROM                        | Author (year)                                    | Structural validity | Internal consistency | Reliability | Measurement error | Hypothesis testing | Cross-cultural validity | Criterion validity | Responsiveness |
|-----------------------------|--------------------------------------------------|---------------------|----------------------|-------------|-------------------|--------------------|------------------------|------------------|-----------------|
| CCEQ financial advice domain | Harley et al., 2019                              | - (no data)         | -                    | NR          | NR                | +                  | NR                     | NR               | NR              |
|                             | (Cronbach’s α=0.71–0.88)                         |                     |                      |             |                   |                    |                        |                  |                 |
| COST-v1                     | De Souza et al., 2014; De Souza et al., 2017     | NR                  | + (Cronbach’s α=0.92)| NR          | NR                | +                  | NR                     | NR               | NR              |
|                             | (ICC=0.80)                                       |                     |                      |             |                   |                    |                        |                  |                 |
| COST-v1, Australia version  | Durber et al., 2021                              | NR                  | + (Cronbach’s α=0.89)| NR          | NR                | +                  | NR                     | NR               | NR              |
|                             | (ICC=0.80)                                       |                     |                      |             |                   |                    |                        |                  |                 |
| COST-v1, Brazilian version  | de Alcantara Nogueira et al., 2020               | - (RMSEA=1.20)      | + (Cronbach’s α=0.83)| NR          | NR                | NR                 | NR                     | NR               | NR              |
|                             | (no clear hypothesis)                            |                     |                      |             |                   |                    |                        |                  |                 |
| COST-v1, India version      | Dar et al., 2021                                 | - (EFA: χ²=60.82)   | ? (Cronbach’s α=0.92)| NR          | NR                | NR                 | NR                     | NR               | NR              |
| COST-v1 Simplified Chinese version | Yu et al., 2021                         | - (CFI=0.86; SRMR=0.08) | + (Cronbach’s α=0.85–0.88) | NR          | +                  | NR                 | (r=-0.57–0.88)         | NR               | ?               |
|                             | (ICC=0.85)                                       |                     |                      |             |                   |                    |                        |                  |                 |
| COST-v2, Italian version    | Ripamonti et al., 2020                           | ? (CFI=1.00; RMSEA=0.04) | ? (Cronbach’s α=0.83) | NR          | +                  | NR                 | (r=-0.66–0.79)         | NR               | + (β=0.55)     |
|                             | (ICC=0.70)                                       |                     |                      |             |                   |                    |                        |                  |                 |
| COST-v2, Persian version    | Sharif et al., 2020                              | + (CFI=0.97; RMSEA=0.07) | + (Cronbach’s α=0.89) | NR          | NR                |NR                  | (no clear hypothesis)  | NR               | NR              |
|                             | (ICC=0.71)                                       |                     |                      |             |                   |                    |                        |                  |                 |
| COST-v2, Traditional Chinese version | Chan et al., 2021                         | - (CFI=0.91; RMSEA=0.15) | + (Cronbach’s α=0.86) | NR          | +                  | NR                 | (r=-0.46–0.47)         | NR               | NR              |
|                             | (ICC=0.71)                                       |                     |                      |             |                   |                    |                        |                  |                 |
| CPILS employment/financial  | Zhao et al., 2009                                | - (EFA: no model)   | + (Cronbach’s α=0.83) | NR          | NR                | NR                 | NR                     | NR               | NR              |
|                             | (no clear hypothesis)                            |                     |                      |             |                   |                    |                        |                  |                 |
| Domain                        | Study Authors           | Methodology                          | Cronbach’s α | ICC/Weighted Kappa | Spearman’s r | Significance |
|-------------------------------|-------------------------|--------------------------------------|---------------|-------------------|------------|--------------|
| FIT                           | Hueniken et al., 2020  | EFA: no model data                   | + (0.78-0.97) | + (0.70)          | + (0.26-0.61) | NR           |
| IOC-CS financial problems domain | Zebrack et al., 2010  | EFA: no model data                   | + (0.70-0.86) | + (0.75)          | + (>75% were not significant) | NR           | ? (no data) |
| PFB                           | Veenstra et al., 2014  | CFA: no model data                   | + (0.79)      | + (0.80)          | + (0.80)    | NR           |
| PROFFIT                       | Riva et al., 2021      | EFA: no model data                   | + (0.87)      | + (0.80)          | + (0.70-0.65) | NR           | + (0.65)     |
| PRRS financial wellbeing domain | Shilling et al., 2018  | EFA: no model data                   | + (0.90)      | + (0.86)          | + (0.80)    | NR           | + (0.71-0.65) |
| P-SAFE, French Version        | Tremblay et al., 2020  | CFA: no model data                   | -             | + (0.90)          | -           | ? (no data) |
| QLACS financial problems domain | Avis et al., 2005      | CFA: no model data                   | + (0.72)      | + (0.50-0.57)     | - (r=0.72)  | NR           |
| SDI-16 money matters domain   | Wright et al., 2011    | EFA: no model data                   | + (0.71-0.82) | -                 | + (33% were not significant) | NR           | ? (r=0.72)  |
| SDI-21 providing for the family domain | Wright et al., 2005  | EFA: no model data                   | + (0.50-0.86) | + (0.54-0.80)     | + (0.54-0.80) | NR           | + (0.54-0.80) |
| SFDQ                          | Dar et al., 2021       | CFA: no model data                   | + (0.85-0.88) | + (0.85-0.88)     | + (0.85-0.88) | NR           | NR           |
| SWBS | Head, 2008; Head & Faul, 2008 | (CFA: no model data) | (Cronbach’s α=0.92) | NR | NR | (r=-0.57-0.60) | NR | (r=-0.12-0.03) | NR |

CCEQ: Chronic Cancer Experiences Questionnaire; CFA: Confirmatory Factor Analysis; CFI: Comparative Fit Index; COST: Comprehensive Score for Financial Toxicity; CPILS: Cancer Problems in Living Scale; EFA: Exploratory Factor Analysis; FIT: Financial Index of Toxicity; ICC: Intraclass Correlation Coefficient; IOC-CS: Impact of Cancer-Childhood Survivors; NA: Not Applicable; NR: Not Report; PFB: Personal Financial Burden; PROM: Patient-Reported Outcome Measures; PROFFIT: Patient-Reported Outcome for Fighting Financial Toxicity; PRRS: Patient Roles and Responsibilities Scale; P-SAFE: Patient Self-Administered Financial Effects questionnaire; QLACS: Quality of Life in Adult Cancer Survivors; RMSEA: Root Mean Square Error of Approximation; SDI: Social Difficulties Inventory Cancer Care Outcomes; SFDQ: Subjective Financial Distress Questionnaire; SWBS: Socioeconomic Wellbeing Scale
### Table 4 Certainty of evidence of measurement properties

| PROM                   | Author (year)                                      | Structural validity | Internal consistency | Reliability | Measurement error | Hypothesis testing | Cross-cultural validity | Criterion validity | Responsiveness |
|------------------------|---------------------------------------------------|---------------------|----------------------|-------------|-------------------|---------------------|------------------------|-------------------|-----------------|
| CCEQ financial advice domain | Harley et al., 2019                              | Moderate            | High                 | -           | -                 | High                | -                      | -                 | -               |
| COST-v1                | De Souza et al., 2014; De Souza et al., 2017      | -                   | High                 | Moderate    | -                 | High                | -                      | -                 | -               |
| COST-v1, Australia version | Durbet et al., 2021                             | -                   | High                 | Moderate    | -                 | High                | -                      | -                 | -               |
| COST-v1, Brazilian version | de Alcantara Nogueira et al., 2020                | High                | High                 | -           | -                 | -                   | Very low               | -                 | -               |
| COST-v1, India version | Dar et al., 2021                                  | Very Low            | Low                  | -           | -                 | -                   | -                      | -                 | -               |
| COST-v1, Simplified Chinese version | Yu et al., 2021                                 | High                | High                 | High        | Low               | -                   | Low                    | -                 | Low             |
| COST-v2, Italian version | Ripamonti et al., 2020                          | Very low            | High                 | Very low    | -                 | High                | -                      | High               | -               |
| COST-v2, Persian version | Sharif et al., 2020                              | High                | High                 | -           | -                 | High                | -                      | Very low           | -               |
| COST-v2, Traditional Chinese version | Chan et al., 2021                              | High                | High                 | Low         | -                 | High                | -                      | -                 | -               |
| CPILS employment/financial domain | Zhao et al., 2009                               | High                | High                 | -           | -                 | High                | -                      | -                 | -               |
| FIT                    | Hueniken et al., 2020                            | Low                 | Moderate             | Very Low    | -                 | Moderate            | -                      | -                 | Moderate        |
| IOC-CS financial problems domain | Zebrack et al., 2010                          | Moderate            | High                 | Moderate    | -                 | High                | -                      | Very low           | -               |
| PFB                    | Veenstra et al., 2014                            | Moderate            | Very low             | -           | -                 | -                   | -                      | -                 | -               |
| PROFFIT                | Riva et al., 2021                                 | Moderate            | High                 | Moderate    | -                 | -                   | -                      | -                 | -               |
| Measure                                      | Domain                                      | Year       | Level of Financial Wellbeing | Level of Financial Problems | Level of Providing for the Family | Psychométrical Properties |
|----------------------------------------------|---------------------------------------------|------------|-------------------------------|------------------------------|-------------------------------|--------------------------|
| PRRS financial wellbeing domain              |                                             | 2018       | Moderate                      | High                         | Very low                      | Very low                 |
| P-SAFE, French Version                       |                                             | 2020       | -                             | -                            | -                             | Very low                 |
| QLACS financial problems domain              |                                             | 2005       | Moderate                      | High                         | Low                           | Moderate                 |
| SDI-16 money matters domain                  |                                             | 2011       | High                          | High                         | Very low                      | Very low                 |
| SDI-21 providing for the family domain       |                                             | 2005       | High                          | High                         | -                             | -                        |
| SFDQ                                         |                                             | 2022       | High                          | High                         | -                             | -                        |
| SWBS                                         |                                             | 2008       | High                          | High                         | High                          | -                        |
| SWBS                                         |                                             | 2008       | High                          | High                         | -                             | -                        |

641 CCEQ: Chronic Cancer Experiences Questionnaire; COST: Comprehensive Score for Financial Toxicity; CPILS: Cancer Problems in Living Scale; FIT: Financial Index of Toxicity; IOC-CS: Impact of Cancer-Childhood Survivors; PFB: Personal Financial Burden; PROM: Patient-Reported Outcome Measures; PROFFIT: Patient-Reported Outcome for Fighting Financial Toxicity; PRRS: Patient Roles and Responsibilities Scale; P-SAFE: Patient Self-Administered Financial Effects questionnaire; QLACS: Quality of Life in Adult Cancer Survivors; RMSEA: Root Mean Square Error of Approximation; SDI: Social Difficulties Inventory Cancer Care Outcomes; SFDQ: Subjective Financial Distress Questionnaire; SWBS: Socioeconomic Wellbeing Scale
Figure 1 PRISMA Flow Chat of Selection Process

Appendix I Search Strategy and Results

Appendix II COSMIN Risk of Bias Checklist
PubMed (N=4003); MEDLINE (OVID) (N=4383); Embase (OVID) (N=1870); Web of Science (N=1002); CINAHL (EBSCO) (N=879); PsycINFO (N=45); ProQuest (N=90); Cochrane Library (N=4453)

Records excluded (N=11669)

Total number after duplication (N=11731)

Full text excluded (N=39) - Phenomena of interest (n=7) - Population group (n=15) - Study design (n=9) - Abstract (n=5) - No psychometric property (n=3)

Studies included in the review (N=23)

Figure 1 PRISMA flow chat of selection process
Appendix I Search strategy and results

Search strategy for PubMed

Search time: 2022-3-13 15:07 (UTC+8)

| Search | Query | Items found |
|--------|-------|-------------|
| #1     | Cancer[Title/Abstract] OR neoplasms[MeSH] | 4176348 |
| #2     | Patient[Title/Abstract] OR survivor[Title/Abstract] OR patients[MeSH] OR “cancer survivors”[MeSH] OR survivors[MeSH] | 2721864 |
| #3     | Cost[Title/Abstract] OR bill[Title/Abstract] OR expense[Title/Abstract] OR “productivity loss”[Title/Abstract] OR “out-of-pocket”[Title/Abstract] OR “economic burden”[Title/Abstract] OR “financial toxicity”[Title/Abstract] OR “financial hardship”[Title/Abstract] OR “financial effect”[Title/Abstract] OR “financial stress”[Title/Abstract] OR “economic burden”[Title/Abstract] OR “economic hardship”[Title/Abstract] OR “co-payment”[Title/Abstract] OR “cost of illness”[MeSH] | 573062 |
| #4     | Scale[Title/Abstract] OR “patient reported outcome measure”[Title/Abstract] OR PROM? [Title/Abstract] OR measure* [Title/Abstract] OR “Patient Reported Outcome Measures”[MeSH] OR “Surveys and Questionnaires”[MeSH] | 5065964 |
| #5     | (instrumentation[sh] OR methods[sh] OR “Validation Studies”[pt] OR “Comparative Study”[pt] OR “psychometrics”[MeSH] OR psychometr*[tiab] OR clinimetr*[tw] OR clinometry*[tw] OR “outcome assessment (health care)”[MeSH] OR “outcome assessment”[tiab] OR “outcome measure”*[tw] OR “observer variation”[MeSH] OR “observer variation”[tiab] OR “Health Status Indicators”[MeSH] OR “reproducibility of results”[MeSH] OR reproducib*[tiab] OR “discriminant analysis”[MeSH] OR reliab*[tiab] OR unrelia*[tiab] OR valid*[tiab] OR “coefficient of variation”[tiab] OR coefficient[tiab] OR homogeneity[tiab] OR homogeneous[tiab] OR “internal consistency”[tiab] OR (cronbach*[tiab] AND (alpha[tiab] OR alpha[tiab])) OR (item[tiab] AND (correlation*[tiab] OR selection*[tiab] OR reduction*[tiab]))) OR agreement*[tw] OR precision*[tw] OR imprecision*[tw] OR “precise values”*[tw] OR test-retest[tiab] OR (test[tiab] AND retest[tiab]) OR (reliab*[tiab] AND ((test[tiab] OR retest[tiab]) OR stability[tiab] OR interrater[tiab] OR inter-rater[tiab] OR intrarater[tiab] OR intra-rater[tiab] OR intertest[tiab] OR inter-tester[tiab] OR intratest[tiab] OR intra-tester[tiab] OR interobserver[tiab] OR inter-observer[tiab] OR intraobserver[tiab] OR intra-observer[tiab] OR inter-technician[tiab] OR inter-technician[tiab] OR intratechnician[tiab] OR inter-technician[tiab] OR inter-examiner[tiab] OR inter-examiner[tiab] OR intra-examiner[tiab] OR intra-examiner[tiab] OR inter-essay[tiab] OR inter-assay[tiab] OR intra-essay[tiab] OR intra-assay[tiab] OR inter-individual[tiab] OR intra-individual[tiab] OR inter-individual[tiab] OR inter-participant[tiab] OR intraparticipant[tiab] OR intra-participant[tiab] OR kappa[tiab] OR kappa*[tiab] OR kappas[tiab] OR kappas[tiab] OR repeatab*[tw] OR ((repliab*[tw] OR repeated*[tw]) AND (measure*[tw] OR measures*[tw] OR result*[tw] OR results*[tw] OR test*[tw] OR tests*[tw]) OR generaliza*[tiab] OR generalisa*[tiab] OR concordance[tiab] OR (intraclass[tiab] AND correlation*[tiab]) OR discriminative[tiab]) OR “known group”[tiab] OR “factor analysis”[tiab] OR “factor analyses”[tiab] OR “factor structure”[tiab] OR “factor structures”[tiab] OR dimension*[tiab] OR subscale*[tiab] OR (multitrait[tiab] AND scaling[tiab] AND (analysis[tiab] OR analyses[tiab])) OR “item discriminant”[tiab] OR “interscale correlation”*[tiab] OR error[tiab] OR errors[tiab] OR “individual variability”[tiab] | 10114928 |
|   | exp / (ti,ab.)                                                                 |   |
|---|-------------------------------------------------------------------------------|---|
| 7 | exp Psychometrics/                                                             | 83581 |
| 8 | psychometr*.ti,ab.                                                             | 46299 |
| 9 | (clinimetr* or clinometr*).tw.                                                | 1135  |
| 10| exp Outcome Assessment Health Care/                                            | 1270817 |
| 11| outcome assessment.ti,ab.                                                      | 3870  |
| 12| outcome measure*.tw.                                                           | 230270 |
| 13| exp Observer Variation/                                                        | 44491 |
| 14| observer variation.ti,ab.                                                      | 1055  |
| 15| exp Health Status Indicators/                                                  | 335480 |
| 16| exp Reproducibility of Results/                                               | 442637 |
| 17| reproducing*.ti,ab.                                                            | 152648 |
| 18| exp Discriminant Analysis/                                                     | 11624 |
| 19| (reliab* or unreliab* or valid* or coefficient or homogeneity or internal consistency).ti,ab. | 1334500 |
| 20| (cronbach* and (alpha or alphas)).ti,ab.                                      | 22858 |
| 21| (item and (correlation* or selection* or reduction*)).ti,ab.                   | 22439 |
| 22| (agreement or precision or imprecision or precise values or test-retest).ti,ab. | 369056 |
| 23| (test and retest).ti,ab.                                                       | 26981 |
| 24| (reliab* and (test or retest)).ti,ab.                                         | 87641 |
| 25| (stability or interrater or inter-rater or intrarater or intra-rater or intertaster or intratester or intra-tester or interobserver or inter-observer or intraobserver or intraobserver or intertechnician or inter-technician or intratechnician or intra-technician or interexaminer or inter-examiner or intraexaminer or intra-examiner or interassay or interassay or intra-assay or interindividual or inter-individual or intra-individual or intra-individual or interparticipant or inter-participant or intraparticipant or intra-participant or kappa or kappa’s or kappas or repeatable*.ti,ab. | 551861 |
| 26| ((replicab* or repeated) and (measure or measures or findings or result or results or test or tests)).ti,ab. | 192657 |
| 27| (generaliza* or generalisa* or concordance).ti,ab.                           | 87243 |
| 28| (intracllass and correlation*).ti,ab.                                         | 26027 |
| 29| (discriminative or known group or factor analysis or factor analyses or dimension* or subscale*).ti,ab. | 547916 |
| 30| (multitrait and scaling and (analysis or analyses)).ti,ab.                    | 134 |
| 31| (item discriminant or interscale correlation* or error or errors or individual variability).ti,ab. | 273771 |
| 32| (variability and (analysis or values)).ti,ab.                                  | 97699 |
| 33| (uncertainty and (measurement or measuring)).ti,ab.                           | 5737 |
| 34| (standard error of measurement or sensitiv* or responsive*).ti,ab.            | 1522843 |
| 35| ((minimal or minimally or clinical or clinically) and (important or significant or detectable) and (change or difference)).ti,ab. | 236237 |
| 36| (small* and (real or detectable) and (change or difference)).ti,ab.           | 7151  |
| 37| (meaningful change or ceiling effect or floor effect or Item response model or IRT or Rasch or Differential item functioning or DIF or computer adaptive testing or item bank or cross-cultural equivalence).ti,ab. | 13325 |

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml
### Search strategy for EMBASE (Ovid)

Search time: 2022-3-13 16:17 (UTC+8)

| Search | Query | Items found |
|--------|-------|-------------|
| #1     | cancer.ab. or cancer.ti. or neoplasms.hw. | 2728162 |
| #2     | patient?.ab. or patient?.ti. or survivor?.ab. or survivor?.ti. or patients.hw. or cancer survivors.hw. or survivors.hw. | 10910535 |
| #3     | Cost.ab. or Cost.ti. or bill?.ab. or bill?.ti. or expense.ab. or expense.ti. or productivity loss.ab. or productivity loss.ti. or out-of-pocket.ab. or out-of-pocket.ti. or economic burden.ab. or economic burden.ti. or financial toxicity.ab. or financial toxicity.ti. or financial hardship.ab. or financial hardship.ti. or financial burden.ab. or financial burden.ti. or financial effect.ab. or financial effect.ti. or financial stress.ab. or financial stress.ti. or economic burden.ab. or economic burden.ti. or economic hardship.ab. or economic hardship.ti. or co-payment.ab. or co-payment.ti. or (cost of illness).hw. | 743789 |
| #4     | Scale?.ab. or Scale?.ti. or patient reported outcome measur*.ab. or patient reported outcome measur*.ti. or PROM?.ab. or PROM?.ti. or measure*.ab. or measure*.ti. or Patient Reported Outcome Measures*.ab. or Patient Reported Outcome Measures*.ti. or (Surveys and Questionnaires).hw. | 5577750 |
| #5     | exp Intermethod comparison/ OR exp data collection method/ OR exp validation study/ OR exp feasibility study/ OR exp pilot study/ OR exp psychometry/ OR exp reproducibility/ OR reproducib*:ab,ti OR audit:ab,ti OR psychometr*:ab,ti OR clinimetr*:ab,ti OR clinometr*:ab,ti OR exp observer variation/ OR observer variation:ab,ti OR exp discriminant analysis/ OR exp validity/ OR relia*:ab,ti OR valid*:ab,ti OR coefficient:ab,ti OR internal consistency:ab,ti OR (cronbach*:ab,ti AND (alpha:ab,ti OR alphas:ab,ti)) OR item correlation:ab,ti OR item item correlations:ab,ti OR item selection:ab,ti OR item selection:ab,ti OR item reduction:ab,ti OR item reductions:ab,ti OR agreement:ab,ti OR precision:ab,ti OR imprecision:ab,ti OR precise values:ab,ti OR test-retest:ab,ti OR (test:ab,ti AND retest:ab,ti) OR (reliab*:ab,ti AND (test:ab,ti OR retest:ab,ti)) OR stability:ab,ti OR interrater:ab,ti OR inter-rater:ab,ti OR intra-rater:ab,ti OR intra-rater:ab,ti OR intertester:ab,ti OR inter-tester:ab,ti OR intratester:ab,ti OR intratester:ab,ti OR interobserver:ab,ti OR inter-observer:ab,ti OR intraobserver:ab,ti OR intraobserver:ab,ti OR intratechnician:ab,ti OR inter-technician:ab,ti OR intratechnician:ab,ti OR intratechnician:ab,ti OR interexaminer:ab,ti OR inter-examiner:ab,ti OR interexaminer:ab,ti OR interexaminer:ab,ti OR interassay:ab,ti OR inter-assay:ab,ti OR intraassay:ab,ti OR intra-assay:ab,ti OR interindividual:ab,ti OR inter-individual:ab,ti OR intra-individual:ab,ti OR intra-individual:ab,ti OR interparticipant:ab,ti OR inter-participant:ab,ti OR intraparticipant:ab,ti OR intraparticipant:ab,ti OR kappa:ab,ti OR kappa:ab,ti OR kappa:ab,ti OR coefficient of variation:ab,ti OR coefficient of variation:ab,ti | 2229098 |
| Search | Query                                                                 | Items found |
|--------|----------------------------------------------------------------------|-------------|
| #1     | TI cancer OR AB cancer OR MH neoplasms                               | 478227      |
| #2     | AB patient? OR TI patient? OR AB survivor? OR TI survivor? OR MH patients OR MH “cancer survivors” OR MH survivors | 2004224     |
| #3     | AB Cost OR TI Cost OR AB bill? OR TI bill? OR AB expense OR TI expense OR AB “productivity loss” OR TI “productivity loss” OR AB out-of-pocket OR TI out-of-pocket OR AB “economic burden” OR TI “economic burden” OR AB “financial toxicity” OR TI “financial toxicity” OR AB “financial hardship” OR TI “financial hardship” OR AB “financial burden” OR TI “financial burden” OR AB “financial effect” OR TI “financial effect” OR AB “financial stress” OR TI “financial stress” OR AB “economic burden” OR TI “economic burden” OR AB “economic hardship” OR TI “economic hardship” OR AB co-payment OR TI co-payment OR MH “cost of illness” | 211023      |
| #4     | AB Scale? OR TI Scale? OR AB “patient reported outcome measure” OR TI “patient reported outcome measure” OR AB PROM? OR TI PROM? OR AB measure* OR TI measure* OR AB “Patient Reported Outcome Measures” OR MH “Patient Reported Outcome Measures” OR MH “Surveys and Questionnaires” | 942889      |
| #5     | (MH “Psychometrics”) or ( TI psychometr* or AB psychometr* ) or ( TI clinimetr* or AB | 702880      |
clinimetr* or (TI clinimetr* OR AB clinimetr*) or (MH “Outcome Assessment”) or (TI outcome assessment or AB outcome assessment) or (TI outcome measure* or AB outcome measure*) or (MH “Health Status Indicators”) or (MH “Reproducibility of Results”) or (MH “Discriminant Analysis”) or (TI reproducib* or AB reproducib*) or (TI reliab* or AB reliab*) or (TI unreliab* or AB unreliab*) or (TI valid* or AB valid*) or (TI coefficient or AB coefficient) or (TI homogeneity or AB homogeneity) or (TI “coefficient of variation” or AB “coefficient of variation”) or (TI “internal consistency” or AB “internal consistency”) or (MH “Internal Consistency”) or (MH “Reliability”) or (MH “Measurement Error”) or (MH “Content Validity”) or “hypothesis testing” or “structural validity” or “cross-cultural validity” or (MH “Criterion-Related Validity”) or “responsiveness” or “interpretability” or (TI reliab* or AB reliab*) and (TI test or AB test) OR (TI retest or AB retest) or (TI stability or AB stability) or (TI interrater or AB interrater) or (TI inter-rater or AB inter-rater) or (TI intrarater or AB intrarater) or (TI intra-rater or AB intra-rater) or (TI intertester or AB intertester) or (TI inter-tester or AB inter-tester) or (TI intratester or AB intratester) or (TI intra-tester or AB intra-tester) or (TI interobserver or AB interobserver) or (TI inter-observer or AB inter-observer) or (TI intraobserver or AB intraobserver) or (TI intra-observer or AB intra-observer) or (TI intertechnician or AB intertechnician) or (TI inter-technician or AB inter-technician) or (TI intratechnician or AB intratechnician) or (TI intra-technician or AB intra-technician) or (TI interexaminer or AB interexaminer) or (TI inter-examiner or AB inter-examiner) or (TI intraexaminer or AB intraexaminer) or (TI intra-examiner or AB intra-examiner) or (TI inter-assay or AB inter-assay) or (TI inter-assay or AB inter-assay) or (TI intr-assay or AB intr-assay) or (TI intra-assay or AB intra-assay) or (TI interindividual or AB interindividual) or (TI inter-individual or AB inter-individual) or (TI intrindividual or AB intrindividual) or (TI intra-individual or AB intra-individual) or (TI interparticipant or AB interparticipant) or (TI inter-participant or AB inter-participant) or (TI intraparticipant or AB intraparticipant) or (TI intra-participant or AB intra-participant) or (TI kappa or AB kappa) or (TI kappa’s or AB kappa’s) or (TI kappas or AB kappas) or (TI repeatab* or AB repeatab*) or (TI responsive* or AB responsive*) or (TI interpretab* or AB interpretab*)

#6 1 and 2 and 3 and 4 and 5 879

Search strategy for PsycINFO (EBSCO)
Search time: 2022-3-13 17:32 (UTC+8)

| Search | Query | Items found |
|--------|-------|-------------|
| #1     | TI cancer OR AB cancer OR MH neoplasms | 1240 |
| #2     | AB patient? OR TI patient? OR AB survivor? OR TI survivor? OR MH patients OR MH “cancer survivors” OR MH survivors | 17860 |
| #3     | AB Cost OR TI Cost OR AB bill? OR TI bill? OR AB expense OR TI expense OR AB “productivity loss” OR TI “productivity loss” OR AB out-of-pocket OR TI out-of-pocket OR AB “economic burden” OR TI “economic burden” OR AB “financial toxicity” OR TI “financial toxicity” OR AB “financial hardship” OR TI “financial hardship” OR AB “financial burden” OR TI “financial burden” OR AB “financial effect” OR TI “financial effect” OR AB “financial | 4217 |
stress” OR TI “financial stress” OR AB “economic burden” OR TI “economic burden” OR AB “economic hardship” OR TI “economic hardship” OR AB co-payment OR TI co-payment OR MH “cost of illness”

|   | #4 | AB Scale? OR TI Scale? OR AB “patient reported outcome measure” OR TI “patient reported outcome measure” OR AB PROM? OR TI PROM? OR AB measure* OR TI measure* OR AB “Patient Reported Outcome Measures” OR MH “Patient Reported Outcome Measures” OR MH “Surveys and Questionnaires” |
|---|---|---|
|   |   | 50879 |

### Search strategy for Web of Science

Search time: 2022-3-13 18:17 (UTC+8)

| Search | Query | Items found |
|---|---|---|
| #1 | TI=cancer OR TS=(cancer OR neoplasms) | 6889664 |
| #2 | TI=(patient? OR survivor?) OR TS=(patients OR “cancer survivors” OR survivors) | 12513517 |
| #3 | TI=(Cost OR bill? OR expense OR “productivity loss” OR out-of-pocket OR “economic burden” OR “financial toxicity” OR “financial hardship” OR “financial burden” OR “financial effect” OR “financial stress” OR “economic burden” OR “economic hardship”) OR TS=“cost of illness” | 463395 |
| #4 | TI=(Scale? OR “patient reported outcome measure” OR PROM? OR measure*) OR TS=(“Patient Reported Outcome Measures” OR “Surveys and Questionnaires”) | 2955804 |
| #6 | #4 AND #3 AND #2 AND #1 | 1002 |

### Search strategy for ProQuest Dissertations and Theses

Search time: 2022-3-13 18:41 (UTC+8)

| Search | Query | Items found |
|---|---|---|
| #1 | ti(cancer OR neoplasms) OR su(cancer OR neoplasms) | 48268 |
| #2 | ti(patient? OR survivor? OR “cancer survivors”) OR su(patient? OR survivor? OR cancer survivors) | 51934 |
| #3 | ti(Cost OR bill? OR expense OR “productivity loss” OR out-of-pocket OR “economic burden” OR “financial toxicity” OR “financial hardship” OR “financial burden” OR “financial effect” OR “financial stress”) OR su(Cost OR bill? OR expense OR “productivity loss” OR out-of-pocket OR “economic burden” OR “financial toxicity” OR “financial hardship” OR “financial burden” OR “financial effect” OR “financial stress”) OR ti(“economic burden” OR “economic hardship” OR “cost of illness”) OR su(“economic burden” OR “economic hardship” OR “cost of illness”) | 45357 |
| #4 | All : Scale? OR “patient reported outcome measure” OR PROM? OR measure* | 2854129 |
| #6 | #4 AND #3 AND #2 AND #1 | 90 |

### Search strategy for Cochrane Library (Wiley)

Search time: 2022-3-13 19:05 (UTC+8)

| Search | Query | Items found |
|---|---|---|
|   |   |   |
| Step | Query                                                                 | Result |
|------|----------------------------------------------------------------------|--------|
| 1    | (Cancer):ti,ab,kw                                                   | 177850 |
| 2    | MeSH descriptor: [Neoplasms] explode all trees                       | 86823  |
| 3    | (patient? OR survivor? OR "cancer survivors"):ti,ab,kw              | 1070252|
| 4    | MeSH descriptor: [Patients] explode all trees                        | 2946   |
| 5    | MeSH descriptor: [Cancer Survivors] explode all trees               | 476    |
| 6    | MeSH descriptor: [Survivors] explode all trees                       | 1760   |
| 7    | (Cost OR bill? OR expense OR "productivity loss" OR out-of-pocket OR "economic burden" OR "financial toxicity" OR "financial hardship" OR "financial burden" OR "financial effect" OR "financial stress" OR "economic burden" OR "economic hardship" OR "cost of illness"):ti,ab,kw | 67224  |
| 8    | MeSH descriptor: [Cost of Illness] explode all trees                | 864    |
| 9    | (Scale? OR "patient reported outcome measur*" OR PROM? OR measure*):ti,ab,kw | 602173 |
| 10   | MeSH descriptor: [Patient Reported Outcome Measures] explode all trees | 929    |
| 11   | MeSH descriptor: [Surveys and Questionnaires] explode all trees      | 58363  |
| 12   | ("Validation Studies" OR "Comparative Study"):pt                    | 169332 |
| 13   | MeSH descriptor: [Psychometrics] explode all trees                   | 2951   |
| 14   | (psychometr*):ti,ab,kw                                             | 6755   |
| 15   | #1 OR #2                                                            | 208929 |
| 16   | #3 OR #4 OR #5 OR #6                                                | 1070725|
| 17   | #7 OR #8 OR #9 OR #10 OR #11                                        | 666816 |
| 18   | #12 OR #13 OR #14                                                  | 175091 |
| 19   | #15 AND #16 AND #17 AND #18                                         | 4453   |
Appendix II COSMIN Checklists

eTable 1 COSMIN risk of bias checklist

| PROM Development                                                                 | Results |
|----------------------------------------------------------------------------------|---------|
| 1. Is a clear description provided of the construct to be measured?             |         |
| 2. Is the origin of the construct clear: was a theory, conceptual framework or disease model used or clear rationale provided to define the construct to be measured? |         |
| 3. Is a clear description provided of the target population for which the PROM was developed? |         |
| 4. Is a clear description provided of the context of use                         |         |
| 5. Was the PROM development study performed in a sample representing the target population for which the PROM was developed? |         |
| 6. Was an appropriate qualitative data collection method used to identify relevant items for a new PROM? |         |
| 7. Were skilled group moderators/interviewers used?                              |         |
| 8. Were the group meetings or interviews based on an appropriate topic or interview guide? |         |
| 9. Were the group meetings or interviews recorded and transcribed verbatim?      |         |
| 10. Was an appropriate approach used to analyse the data?                        |         |
| 11. Was at least part of the data coded independently?                          |         |
| 12. Was data collection continued until saturation was reached?                  |         |
| 13. For quantitative studies (surveys): was the sample size appropriate?         |         |
| 14. Was a cognitive interview study or other pilot test conducted?               |         |
| 15. Was the cognitive interview study or other pilot test performed in a sample representing the target population? |         |
| 16. Were patients asked about the comprehensibility of the PROM?                 |         |
| 17. Were all items tested in their final form?                                   |         |
| 18. Was an appropriate qualitative method used to assess the comprehensibility of the PROM instructions, items, response options, and recall period? |         |
| 19. Was each item tested in an appropriate number of patients?                   |         |
| 20. Were skilled interviewers used?                                              |         |
| 21. Were the interviews based on an appropriate interview guide?                 |         |
| 22. Were the interviews recorded and transcribed verbatim?                       |         |
| 23. Was an appropriate approach used to analyse the data?                        |         |
| 24. Were at least two researchers involved in the analysis?                      |         |
| 25. Were problems regarding the comprehensibility of the PROM instructions, items, response options, and recall period appropriately addressed by adapting the PROM? |         |
| 26. Were patients asked about the comprehensiveness of the PROM?                 |         |
| 27. Was the final set of items tested?                                           |         |
| 28. Was an appropriate method used for assessing the comprehensiveness of the PROM? |         |
| 29. Was each item tested in an appropriate number of patients?                   |         |
| 30. Were skilled interviewers used?                                              |         |
| 31. Were the interviews based on an appropriate interview guide?                 |         |
| 32. Were the interviews recorded and transcribed verbatim?                       |         |
| Question                                                                 | Answer |
|-------------------------------------------------------------------------|--------|
| 33. Was an appropriate approach used to analyse the data?               |        |
| 34. Were at least two researchers involved in the analysis?             |        |
| 35. Were problems regarding the comprehensiveness of the PROM appropriately addressed by adapting the PROM? |        |
| Content validity                                                        |        |
| 1. Was an appropriate method used to ask patients whether each item is relevant for their experience with the condition? |        |
| 2. Was each item tested in an appropriate number of patients?           |        |
| 3. Were skilled group moderators/interviewers used?                     |        |
| 4. Were the group meetings or interviews based on an appropriate topic or interview guide? |        |
| 5. Were the group meetings or interviews recorded and transcribed verbatim? |        |
| 6. Was an appropriate approach used to analyse the data?                |        |
| 7. Were at least two researchers involved in the analysis?              |        |
| 8. Was an appropriate method used for assessing the comprehensiveness of the PROM? |        |
| 9. Was each item tested in an appropriate number of patients?           |        |
| 10. Were skilled group moderators/interviewers used?                    |        |
| 11. Were the group meetings or interviews based on an appropriate topic or interview guide? |        |
| 12. Were the group meetings or interviews recorded and transcribed verbatim? |        |
| 13. Was an appropriate approach used to analyse the data?                |        |
| 14. Were at least two researchers involved in the analysis?              |        |
| 15. Was an appropriate qualitative method used for assessing the comprehensibility of the PROM instructions, items, response options, and recall period? |        |
| 16. Was each item tested in an appropriate number of patients?           |        |
| 17. Were skilled group moderators/interviewers used?                    |        |
| 18. Were the group meetings or interviews based on an appropriate topic or interview guide? |        |
| 19. Were the group meetings or interviews recorded and transcribed verbatim? |        |
| 20. Was an appropriate approach used to analyse the data?                |        |
| 21. Were at least two researchers involved in the analysis?              |        |
| 22. Was an appropriate method used to ask professionals whether each item is relevant for the construct of interest? |        |
| 23. Were professionals from all relevant disciplines included?           |        |
| 24. Was each item tested in an appropriate number of professionals?      |        |
| 25. Was an appropriate approach used to analyse the data?                |        |
| 26. Were at least two researchers involved in the analysis?              |        |
| 27. Was an appropriate method used for assessing the comprehensiveness of the PROM? |        |
| 28. Were professionals from all relevant disciplines included?           |        |
| 29. Was each item tested in an appropriate number of professionals?      |        |
| 30. Was an appropriate approach used to analyse the data?                |        |
| 31. Were at least two researchers involved in the analysis?              |        |

Structural validity
1. For CTT: Was exploratory or confirmatory factor analysis performed?

2. For IRT/Rasch: does the chosen model fit to the research question?

3. Was the sample size included in the analysis adequate?

4. Were there any other important flaws in the design or statistical methods of the study?

**Internal consistency**

1. Was an internal consistency statistic calculated for each unidimensional scale or subscale separately?

2. For continuous scores: Was Cronbach’s alpha or omega calculated?

3. For dichotomous scores: Was Cronbach’s alpha or KR-20 calculated?

4. For IRT-based scores: Was standard error of the theta (SE (θ)) or reliability coefficient of estimated latent trait value (index of (subject or item) separation) calculated?

5. Were there any other important flaws in the design or statistical methods of the study?

**Cross-cultural validity**

1. Were the samples similar for relevant characteristics except for the group variable?

2. Was an appropriate approach used to analyse the data?

3. Was the sample size included in the analysis adequate?

4. Were there any other important flaws in the design or statistical methods of the study?

**Reliability**

1. Were patients stable in the interim period on the construct to be measured?

2. Was the time interval appropriate?

3. Were the test conditions similar for the measurements? e.g. type of administration, environment, instructions

4. For continuous scores: Was an intraclass correlation coefficient (ICC) calculated?

5. For dichotomous/nominal/ordinal scores: Was kappa calculated?

6. For ordinal scores: Was a weighted kappa calculated?

7. For ordinal scores: Was the weighting scheme described? e.g. linear, quadratic

8. Were there any other important flaws in the design or statistical methods of the study?

**Measurement error**

1. Were patients stable in the interim period on the construct to be measured?

2. Was the time interval appropriate?

3. Were the test conditions similar for the measurements? (e.g. type of administration, environment, instructions)

4. For continuous scores: Was the Standard Error of Measurement (SEM), Smallest Detectable Change (SDC) or Limits of Agreement (LoA) calculated?

5. For dichotomous/nominal/ordinal scores: Was the percentage (positive and negative) agreement calculated?

6. Were there any other important flaws in the design or statistical methods of the study?

**Criterion validity**

1. For continuous scores: Were correlations, or the area under the receiver operating curve calculated?

2. For dichotomous scores: Were sensitivity and specificity determined?

3. Were there any other important flaws in the design or statistical methods of the study?

Hypotheses testing for construct validity
1. Is it clear what the comparator instrument(s) measure(s)?
2. Were the measurement properties of the comparator instrument(s) sufficient?
3. Was the statistical method appropriate for the hypotheses to be tested?
4. Were there any other important flaws in the design or statistical methods of the study?
5. Was an adequate description provided of important characteristics of the subgroups?
6. Was the statistical method appropriate for the hypotheses to be tested?
7. Were there any other important flaws in the design or statistical methods of the study?

**Responsiveness**

1. For continuous scores: Were correlations between change scores, or the area under the Receiver Operator Curve (ROC) curve calculated?
2. For dichotomous scales: Were sensitivity and specificity (changed versus not changed) determined?
3. Were there any other important flaws in the design or statistical methods of the study?
4. Is it clear what the comparator instrument(s) measure(s)?
5. Were the measurement properties of the comparator instrument(s) sufficient?
6. Was the statistical method appropriate for the hypotheses to be tested?
7. Were there any other important flaws in the design or statistical methods of the study?
8. Was an adequate description provided of important characteristics of the subgroups?
9. Was the statistical method appropriate for the hypotheses to be tested?
10. Were there any other important flaws in the design or statistical methods of the study?
11. Was an adequate description provided of the intervention given?
12. Was the statistical method appropriate for the hypotheses to be tested?
13. Were there any other important flaws in the design or statistical methods of the study?
### Table 2 - Criteria for good measurement properties

| Measurement property | Rating | Criteria |
|----------------------|--------|----------|
| **Structural validity** | + | CTT:
  CFA: CFI or TLI or comparable measure >0.95 OR RMSEA <0.06 OR SRMR <0.082
  IRT/Rasch:
  No violation of unidimensionality3: CFI or TLI or comparable measure >0.95 OR RMSEA <0.06 OR SRMR <0.08
  AND
  no violation of local independence: residual correlations among the items after controlling for the dominant factor <0.20 OR Q3’s < 0.37
  AND
  no violation of monotonicity: adequate looking graphs OR item scalability >0.30
  AND
  adequate model fit:
  IRT: χ² >0.01
  Rasch: infit and outfit mean squares ≥ 0.5 and ≤ 1.5 OR Z standardized values >-2 and <2 |
|                      | - | Criteria for ‘+’ not met |
| **Internal consistency** | + | At least low evidence for sufficient structural validity AND Cronbach’s alpha(s) ≥ 0.70 for each unidimensional scale or subscale |
|                      | - | Criteria for “At least low evidence for sufficient structural validity” not met |
| **Reliability** | + | ICC or weighted Kappa ≥ 0.70 |
|                      | ? | ICC or weighted Kappa not reported |
|                      | - | ICC or weighted Kappa < 0.70 |
| **Measurement error** | + | SDC or LoA < MIC |
|                      | ? | MIC not defined |
|                      | - | SDC or LoA > MIC |
| **Hypotheses testing for construct validity** | + | The result is in accordance with the hypothesis |
|                      | ? | No hypothesis defined (by the review team) |
|                      | - | The result is not in accordance with the hypothesis |
| **Cross-cultural** | + | No important differences found between group factors (such as cultural differences) |

---

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml
| Validity/Measurement Invariance | Description                                                                 |
|--------------------------------|-----------------------------------------------------------------------------|
| ?                              | No multiple group factor analysis OR DIF analysis performed                   |
| -                              | Important differences between group factors OR DIF was found                  |
| Criterion Validity             | +                             | Correlation with gold standard ≥ 0.70 OR AUC ≥ 0.70                        |
|                                | ?                             | Not all information for ‘+’ reported                                       |
|                                | -                             | Correlation with gold standard < 0.70 OR AUC < 0.70                       |
| Responsiveness                 | +                             | The result is in accordance with the hypothesis7 OR AUC ≥ 0.70            |
|                                | ?                             | No hypothesis defined (by the review team)                                 |
|                                | -                             | The result is not in accordance with the hypothesis7 OR AUC < 0.70        |

AUC: area under the curve; CFA: confirmatory factor analysis; CFI: comparative fit index; CTT: classical test theory; DIF: differential item functioning; ICC: intraclass correlation coefficient; IRT: item response theory; LoA: limits of agreement; MIC: minimal important change; RMSEA: Root Mean Square Error of Approximation; SEM: Standard Error of Measurement; SDC: smallest detectable change; SRMR: Standardized Root Mean Residuals; TLI: Tucker-Lewis index; “+”: sufficient; “-”: insufficient; “?”: indeterminate.
**eTable 3** Modified GRADE approach for assessing certainty of evidence

| Domain              | Grade     | Reason                                                                                                                                 |
|---------------------|-----------|--------------------------------------------------------------------------------------------------------------------------------------|
| Risk of bias        | -0 level: No | There are multiple studies of at least adequate quality, or there is one study of very good quality available                        |
|                     | -1 level: Serious | There are multiple studies of doubtful quality available, or there is only one study of adequate quality                           |
|                     | -2 level: Very serious | There are multiple studies of inadequate quality, or there is only one study of doubtful quality available                     |
|                     | -3 level: Extremely serious | There is only one study of inadequate quality available                                                                              |
| Inconsistency       | -0 level: No | There is no inconsistency among pooled studies or there is only one study in subgroups                                               |
|                     | -1 level: Serious | There are severe inconsistencies among pooled studies                                                                               |
|                     | -2 level: Very serious | There are very severe inconsistencies among pooled studies.                                                                         |
| Imprecision         | -0 level: No | Total sample size > 50-100                                                                                                               |
|                     | -1 level: Serious | Total sample size = 50-100                                                                                                           |
|                     | -2 level: Very serious | Total sample size < 50                                                                                                               |
| Indirectness        | -0 level: No | There is no indirectness between results and conclusion                                                                                |
|                     | -1 level: Serious | There is severe indirectness between results and conclusion                                                                           |
|                     | -2 level: Very serious | There is very severe indirectness between results and conclusion                                                                     |

*The starting point of quality level is high evidence. The quality of evidence is subsequently downgraded to moderate, low, or very low evidence.*
| Section and Topic   | Item # | Checklist item                                                                                                                                                                                                 | Location where item is reported |
|--------------------|--------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------|
| **TITLE**          | 1      | Identify the report as a systematic review.                                                                                                                                                                    | Title                           |
| **ABSTRACT**       | 2      | See the PRISMA 2020 for Abstracts checklist.                                                                                                                                                                  | P3-4                            |
| **INTRODUCTION**   | 3      | Describe the rationale for the review in the context of existing knowledge.                                                                                                                                  | P7-9                            |
| **METHODS**        | 4      | Provide an explicit statement of the objective(s) or question(s) the review addresses.                                                                                                                       | P9                              |
| Eligibility criteria | 5     | Specify the inclusion and exclusion criteria for the review and how studies were grouped for the syntheses.                                                                                                | P10-11                          |
| Information sources | 6     | Specify all databases, registers, websites, organisations, reference lists and other sources searched or consulted to identify studies. Specify the date when each source was last searched or consulted. | P9-10                           |
| Search strategy    | 7      | Present the full search strategies for all databases, registers and websites, including any filters and limits used.                                                                                   | Appendix I                      |
| Selection process  | 8      | Specify the methods used to decide whether a study met the inclusion criteria of the review, including how many reviewers screened each record and each report retrieved, whether they worked independently, and if applicable, details of automation tools used in the process. | P11                             |
| Data collection process | 9   | Specify the methods used to collect data from reports, including how many reviewers collected data from each report, whether they worked independently, any processes for obtaining or confirming data from study investigators, and if applicable, details of automation tools used in the process. | P11-12                          |
| Data items         | 10a    | List and define all outcomes for which data were sought. Specify whether all results that were compatible with each outcome domain in each study were sought (e.g. for all measures, time points, analyses), and if not, the methods used to decide which results to collect. | P12                             |
| Study risk of bias assessment | 11 | Specify the methods used to assess risk of bias in the included studies, including details of the tool(s) used, how many reviewers assessed each study and whether they worked independently, and if applicable, details of automation tools used in the process. | P11-12                          |
| Effect measures    | 12     | Specify for each outcome the effect measure(s) (e.g. risk ratio, mean difference) used in the synthesis or presentation of results.                                                                          | NA                              |
| Synthesis methods  | 13a    | Describe the processes used to decide which studies were eligible for each synthesis (e.g. tabulating the study intervention characteristics and comparing against the planned groups for each synthesis (item #5)). | P12-13                          |
|                    | 13b    | Describe any methods required to prepare the data for presentation or synthesis, such as handling of missing summary statistics, or data conversions.                                                                 | P12-13                          |
|                    | 13c    | Describe any methods used to tabulate or visually display results of individual studies and syntheses.                                                                                                      | P12-13                          |
|                    | 13d    | Describe any methods used to synthesize results and provide a rationale for the choice(s). If meta-analysis was performed, describe the model(s), method(s) to identify the presence and extent of statistical heterogeneity, and software package(s) used. | P12-13                          |
|                    | 13e    | Describe any methods used to explore possible causes of heterogeneity among study results (e.g. subgroup analysis, meta-regression).                                                                      | NA                              |
|                    | 13f    | Describe any sensitivity analyses conducted to assess robustness of the synthesized results.                                                                                                                  | NA                              |
| Reporting bias assessment | 14       | Describe any methods used to assess risk of bias due to missing results in a synthesis (arising from reporting biases).                                                                  | NA                              |
| Certainty assessment | 15        | Describe any methods used to assess certainty (or confidence) in the body of evidence for an outcome.                                                                                                       | P13                             |
| Section and Topic | Item # | Checklist item | Location where item is reported |
|------------------|--------|----------------|---------------------------------|
| **RESULTS**      |        |                |                                 |
| Study selection  | 16a    | Describe the results of the search and selection process, from the number of records identified in the search to the number of studies included in the review, ideally using a flow diagram. | Figure 1 |
|                  | 16b    | Cite studies that might appear to meet the inclusion criteria, but which were excluded, and explain why they were excluded. | P14 |
| Study characteristics | 17 | Cite each included study and present its characteristics. | Table 1 |
| Risk of bias in studies | 18 | Present assessments of risk of bias for each included study. | Table 3 |
| Results of individual studies | 19 | For all outcomes, present, for each study: (a) summary statistics for each group (where appropriate) and (b) an effect estimate and its precision (e.g. confidence/credible interval), ideally using structured tables or plots. | Table 1 & 2 |
| Results of syntheses | 20a | For each synthesis, briefly summarise the characteristics and risk of bias among contributing studies. | Table 2 |
|                  | 20b    | Present results of all statistical syntheses conducted. If meta-analysis was done, present for each the summary estimate and its precision (e.g. confidence/credible interval) and measures of statistical heterogeneity. If comparing groups, describe the direction of the effect. | NA |
|                  | 20c    | Present results of all investigations of possible causes of heterogeneity among study results. | NA |
|                  | 20d    | Present results of all sensitivity analyses conducted to assess the robustness of the synthesized results. | Table 2 |
| Reporting biases | 21    | Present assessments of risk of bias due to missing results (arising from reporting biases) for each synthesis assessed. | NA |
| Certainty of evidence | 22 | Present assessments of certainty (or confidence) in the body of evidence for each outcome assessed. | Table 2 |
| **DISCUSSION**   |        |                |                                 |
| Discussion       | 23a    | Provide a general interpretation of the results in the context of other evidence. | P16-17 |
|                  | 23b    | Discuss any limitations of the evidence included in the review. | P21 |
|                  | 23c    | Discuss any limitations of the review processes used. | P21 |
|                  | 23d    | Discuss implications of the results for practice, policy, and future research. | P21-22 |
| **OTHER INFORMATION** | | | |
| Registration and protocol | 24a | Provide registration information for the review, including register name and registration number, or state that the review was not registered. | P9 |
|                  | 24b    | Indicate where the review protocol can be accessed, or state that a protocol was not prepared. | P9 |
|                  | 24c    | Describe and explain any amendments to information provided at registration or in the protocol. | No amendments |
| Support          | 25    | Describe sources of financial or non-financial support for the review, and the role of the funders or sponsors in the review. | P22 |
| Competing interests | 26 | Declare any competing interests of review authors. | P22 |
| Availability of data, code and other materials | 27 | Report which of the following are publicly available and where they can be found: template data collection forms; data extracted from included studies; data used for all analyses; analytic code; any other materials used in the review. | NA |