Title

The validity of instruments to measure knowledge in population-based cancer screening targeting individuals at average risk – a systematic review

Rikke Nicoline Stokholm[1,2]*, MScPH (ORCID ID: 0000-0001-8618-4944)

Louise Stenholt[3], Medical librarian

Henrik Hein Lauridsen[4], Associate Professor (ORCID ID: 0000-0001-5198-5215)

Adrian Edwards[1,5], Professor (ORCID ID: 0000-0002-6228-4446)

Berit Andersen[1,2], Professor (ORCID ID: 0000-0003-4074-6504)

Mette Bach Larsen[1,2], Associate Professor (ORCID ID: 0000-0002-0727-5571)

1) Department of Public Health Programmes and University Research Clinic for Cancer Screening, Randers Regional Hospital, Randers, Denmark

2) Department of Clinical Medicine, Aarhus University, Denmark

3) Medical Library, Regional Hospital Central Jutland, Denmark

4) Department of Sports Science and Clinical Biomechanics, University of Southern Denmark, Denmark

5) Division of Population Medicine, School of Medicine, Cardiff University, UK

Word count: Abstract: 249, Manuscript: 4.088

***Corresponding author**

Rikke Nicoline Stokholm, Department of Public Health Programmes and University Research Clinic for Cancer Screening, Randers Regional Hospital, Skovlyvej 15, DK-8930 Randers NO, Denmark.

Phone: +45 7842 0172, E-mail: ristok@rm.dk

Abbreviations

COnsensus-based Standards for the selection of health Measurement Instruments (COSMIN)

Patient reported outcome measure (PROM)

Informed Choice in Mammography Screening Questionnaire (IMQ)

Colorectal cancer screening (CRC)

Colorectal Cancer Screening Decision Quality Instrument (CRC-DQI)

HPV Testing Knowledge Scale (HTKS)

Intraclass Correlation Coefficient (ICC)

Item response theory (IRT)

Abstract

Objectives: Relevant knowledge is essential for informed choices about (non)participation in population-based cancer screening. Many instruments have been proposed to assess residents' knowledge about cancer screening programmes but their measurement properties are unknown. This systematic review aims to identify and critically evaluate the measurement properties of instruments to measure knowledge about cancer screening in individuals eligible for population-based screening.

Methods: A literature search was undertaken in PubMed, PsycINFO, Embase, CINAHL, Scopus and Web of Science in August 2023. The review included any study reporting one or more measurement properties of the questionnaire or sub-scale used measuring knowledge of cancer screening including breast, colorectal and/or cervical cancer screening. Studies including males aged 45 or older and females aged 20 or older were included. Two independent reviewers screened the articles and assessed the included articles using the COnsensus-based Standards for the selection of health Measurement Instruments (COSMIN).

Results: We included 24 instruments, which varied in number and characteristics of items. All instruments were assessed as having an inadequate instrument development. The results of the measurement properties, structural validity, internal consistency and reliability were assessed as indeterminate, while construct validity and responsiveness were sufficient.

Conclusion: This systematic review identified no instruments to measure knowledge about cancer screening where the measurement properties were sufficiently evaluated. There is a lack of focus on content validity, and further validation of the instruments is needed. The results indicate a lack of shared understanding or agreement of what constitutes relevant knowledge about cancer screening.

Keywords: Systematic review, population-based cancer screening, knowledge, measurement properties, instrument

Introduction

Many instruments have been proposed to assess residents' knowledge about participation in cancer screening programmes, but their measurement properties are poorly elucidated. This impedes consistent use of instruments, making it challenging to conduct evidence synthesis and meta-analysis about effective interventions and programmes.

In recent years there has been an increased focus on the importance of ensuring that residents make informed choices about participation in population-based cancer screening (Ghanouni et al., 2016). An informed choice is defined as *'one that is based on relevant knowledge, consistent with the decision-maker's values and behaviourally implemented'* (Marteau et al., 2001). Thus, obtaining relevant knowledge is central for an informed choice.

It is usually expected that residents make the decision about whether or not to participate in cancer screening without consulting health care professionals. Therefore, it is essential to have comprehensible and useful self-administered information materials, which can contribute to knowledge about cancer screening and support individuals to make informed decisions about participation or not (European Colorectal Cancer Screening Guidelines Working Group: von Karsa et al., 2013). Several studies have developed and evaluated interventions, which aim to improve informed choice and residents' knowledge about cancer screening (Gabel et al., 2020a; Gabel et al., 2020b; Hersch et al., 2015; Larsen et al., 2021; Mathieu et al., 2007; Mathieu et al., 2010; Reder and Kolip, 2017; Roberto et al., 2020; Smith et al., 2010; Trevena et al., 2008). However, these studies lack standardized instruments for assessing knowledge related to cancer (Ghanouni et al., 2016). This is problematic in terms of the validity of reported findings. Furthermore, it hampers comparison of results across patient populations making it more difficult to make general recommendation on requirements for informed choice.

It can be difficult and time consuming to choose the most appropriate measurement instrument, and biased inferences may be drawn if the chosen instrument does not measure the construct of interest, or if the instrument used has not been validated in the target population in which the investigation is conducted (Mokkink et al., 2006; Terwee et al., 2007). The best method to assess the validity of instruments is through a systematic review that appraises the instruments´ measurement properties according to key criteria such as the COSMIN guideline for patient reported outcome measures (Scholtes et al., 2011; Terwee et al., 2007).

This systematic review aims to identify and critically appraise the measurement properties of instruments measuring knowledge about cancer screening in individuals eligible for population-based screening.

Methods

This systematic review adheres to the COSMIN methodology for systematic reviews of Patient-Reported Outcome Measures (Mokkink, 2018; Terwee, 2018). To evaluate the methodological quality of the included studies, the COSMIN Risk of Bias Checklist was used, and measurement properties of the instruments were evaluated using the COSMIN criteria (Prinsen et al., 2018; Terwee et al., 2007). The protocol was registered in PROSPERO (registration number: CRD42021289395).

Search strategy

The search was developed and conducted in collaboration with an experienced medical librarian (LS). A comprehensive literature search was performed using six different bibliographic databases: PubMed, Embase, PsychInfo, CINAHL, Scopus and Web of Science (complete search strategies are shown in Appendix A). The search period spanned from the inception of each database until November 19th 2021 and was updated on August 2nd 2023 to ensure that subsequent relevant studies were included. The search was customized for each database using both controlled thesaurus terms and natural language terms for synonyms (Appendix A). The validated COSMIN search filter on measurement properties developed for PubMed (Terwee et al., 2009) was used with minor adjustments such as new MeSH-terms, field codes etc. and adapted accordingly to the other databases. A follow-up search of references cited in included studies was conducted.

Eligibility criteria

The review included any study reporting one or more measurement properties of instruments or sub-scales used for measuring cancer screening knowledge of individuals eligible for population-based cancer screening for breast, colorectal and/or cervical cancer. However, studies were excluded if the development process could not be assessed. According to the European Union, the relevant populations include males aged 45 or older and females aged 20 or older (Council of the European Union, 2022; European Commission et al., 2008; European Commission et al., 2013; European Commission et al., 2010). Studies covering this age range were included even if they also

included younger and/or older individuals. Experimental and quasi-experimental study designs and observational studies were included if the study reported measurement properties of the instrument or sub-scale used to measure knowledge about cancer screening. After submission of the protocol, a criterion was added to exclude studies outside Europe, North America and Australasia these studies could be less comparable to studies in Europe, North America and Australasia. Studies were also excluded if there were no separate scores for the level of knowledge about cancer screening (Table 1).

Study selection

The studies identified from the systematic search were uploaded to Covidence Systematic Review Software and duplicates were removed. Initially, the titles and abstracts were screened using the eligibility criteria. Potentially eligible studies were then assessed in full against the eligibility criteria. All assessments were carried out by two independent reviewers (RNS and MBL), and any disagreements between the reviewers were resolved through discussion.

Data extraction

First, extraction of characteristics of the included studies was conducted (country, instrument, study design, number of participants, target population, mode of administration, (sub)scale(s) and number of items, response options, language of instrument and measurement properties).
Second, data extraction on measurement properties was conducted (content validity including PROM development, structural validity, internal consistency, reliability, criterion validity, construct validity and responsiveness).

Risk of Bias: evaluation of measurement properties and quality of evidence

According to COSMIN, the evaluation of each measurement property includes three phases. Initially, the COSMIN Risk of Bias Checklist was applied to assess *methodological quality* of studies. COSMIN Risk of Bias Checklist consists of 10 elements which are assessed in the following order: 1) instrument development, 2) content validity, 3) structural validity, 4) internal consistency, 5) cross-cultural validity, 6) reliability, 7) measurement error, 8) criterion validity, 9) hypothesis testing for construct validity, 10) responsiveness. Each element is rated as very good, adequate, doubtful or inadequate quality. Each element is assessed by evaluating corresponding standards. A standard is rated 'very good' if there is evidence that the quality of the study to which

the standard is referring is adequate, 'adequate' if relevant information is not reported in the article but it can be assumed that the quality aspect is adequate, 'doubtful' if it is doubtful whether the quality aspect is adequate and 'inadequate' if the evidence provided for the quality aspect is not adequate. To determine the overall rating for the element, the lowest rating of any standard is chosen (Mokkink, 2018; Terwee, 2018). If there is high quality evidence that content validity including the patient reported outcome measure (PROM) development were inadequate, COSMIN recommends the instruments not to be further assessed (Mokkink, 2018; Prinsen et al., 2018).

Evidence of good measurement properties

In the second sub-step, the result of each study on any measurement property should be evaluated in accordance with the revised criteria for good measurement properties. For each of the measurement properties, criteria are listed that the property should meet to be rated as sufficient (+), insufficient (-) or indeterminate (?) (Mokkink, 2018). To assess construct validity and responsiveness of the included instruments, hypotheses based on the intervention and results were formulated by the review team:

1. Participants in the intervention group would exhibit greater knowledge of cancer screening than those in the control group.
2. Knowledge level increased significantly between pre-test and post-test.
3. A correlation between two cognitive measures is expected so that one increases if another does and vice versa.

Confidence in the data

Finally, the third sub-step includes summarising the evidence and grading the quality of evidence using the modified GRADE approach (Mokkink, 2018; Terwee, 2018). This step could not be performed since there was not enough evidence to summarise and grade any of the instruments. The quality assessment was performed by two reviewers independently (RNS and MBL) and the final results were determined through consensus process.

Results

Study selection

The search strategy identified 7915 studies from the databases and 10 studies from searching selected reference lists. After removing duplicates, 5327 studies were screened by title and abstract. Of the remaining 370 studies, 38 were excluded as they could not be retrieved. Thus, 332 studies were assessed for eligibility by reading full text, and 303 were excluded mainly due to being outside Europe, North America or Australasia, no measurement properties were reported or the study did not report a measure of residents' knowledge of cancer screening (a list of these studies are found in Appendix B). Forty-three studies were excluded because the study only reported construct validity (without further description of the development of the instrument), and 16 studies were not evaluated by the COSMIN Risk of Bias Checklist as the studies referred to another study, which was assessed to have inadequate PROM development or content validity and only reported internal consistency or construct validity. Therefore, 25 studies (24 instruments) were included in the systematic review (Figure 1).

Study characteristics

In total 17 of the included studies derived from North America. Six studies were conducted in, and one study was conducted in Australia. One study was conducted in United Kingdom, Australia and USA (Table 2).

Ten studies used an instrument to measure knowledge about breast cancer screening. Eleven studies were about colorectal cancer screening, and four studies were about cervical cancer screening. Most studies had a randomized controlled trial design, a cross sectional design or a pre-test-post-test design, while only eight studies were designed as actual development and validation studies. The study populations ranged from 40 to 7142 participants, and the ages of the target populations ranged from 18 to 89 years old and thus within the age group for residents eligible for population-based cancer screening (Table 2).

Instrument characteristics

Of the 10 studies which examined an instrument to measure knowledge about breast cancer screening, 7 instruments were '*ad hoc'* instruments (Chamot and Perneger, 2002; Goel et al., 2011; Hickey et al., 2013; Kagawa-Singer et al., 2009; Price, 1994; Schonberg et al., 2014; van Agt et al., 2012), which means that the instrument was developed and evaluated for the particular study. One study examined the Breast Cancer Knowledge Test developed by McCance et al. (McCance et al., 1990), one study examined a revised version of the Breast Cancer Knowledge Test (Lee et al., 2017) and one the Informed Choice in Mammography Screening Questionnaire (IMQ) (Knowledge index) (Reder et al., 2019). Of the 11 studies examining an instrument about colorectal cancer

screening (CRC) knowledge, one study examined the CRC Knowledge, Perception and Screening Survey (Green and Kelly, 2004), another study examined the Colorectal Cancer Knowledge Questionnaire (Weinrich et al., 1992), one examined the knowledge subscale of Colorectal Cancer Screening Decision Quality Instrument (CRC-DQI) (Sepucha et al., 2014), while a fifth study examined the Knowledge Assessment Survey (Sanchez et al., 2013). The remaining seven studies examined *ad hoc* instruments (Table 2). For cervical cancer screening one study examined an *ad hoc* instrument (Breitkopf et al., 2005), while one examined the HPV testing knowledge instrument (Waller et al., 2013) and two studies examined the HPV Testing Knowledge Scale (Griffin-Mathieu et al., 2022; Haward et al., 2022).

The number of items included in the instruments ranged from 3 to 28 items. Most instruments had questions about both the cancer disease in general and the cancer screening test. The commonest reporting response options were 'true, false, don´t know', 'correct, incorrect answers, don't know' or 'yes and no'. However, some studies did not report response options. Three studies reported the mode of administration to be self-administered (Breitkopf et al., 2005; Chamot and Perneger, 2002; Reder et al., 2019), three were interview-based (Kagawa-Singer et al., 2009; Peterson et al., 2007; Ramírez-Amill et al., 2017), one was reported as telephone survey (Price, 1994) and three by web-based questionnaires (Gabel et al., 2019; Haward et al., 2022; Waller et al., 2013). The remaining studies did not report mode of administration of the instrument. Four studies reported the language of the instrument to be both English and French (Chamot and Perneger, 2002; Griffin-Mathieu et al., 2022; Haward et al., 2022), one study Hmong (China and Southeast Asia) (Kagawa-Singer et al., 2009), one study German and Turkish (Reder et al., 2019), and two studies both English and Spanish (Breitkopf et al., 2005; Sanchez et al., 2013). For the remaining studies the languages were not reported. PROM development, structural validity, internal consistency and construct validity were the most frequently assessed measurement properties (Table 2).

COSMIN Risk of Bias assessment

Using the Risk of Bias checklist all the included studies were assessed as having an inadequate PROM development (Table 3) and there was not enough information reported to assess the content validity. In order to provide a more nuanced appraisal of the existing instruments for the measurement of knowledge about cancer screening, we have decided to further assessed all the

included studies knowing that COSMIN guidelines do not recommend to further assess instruments with inadequate content validity including PROM development.

The *ad hoc* instruments for the measurement of knowledge about breast, colorectal and/or cervical cancer included in this systematic review were assessed as having an inadequate PROM development. Several of the *ad hoc* instruments were assessed as having a clear description of the construct to be measured. However, unclear description of the target population for which the instruments were developed, that the instruments were not developed in a sample representing the target population and lack of cognitive interviews or other pilot test led to an inadequate total quality of PROM development. Cognitive interviews or other pilot tests were performed for nine of the 15 *ad hoc* instruments. However, the quality of the comprehensibility or comprehensiveness examinations were doubtful or inadequate. For some of the *ad hoc* instruments internal consistency, reliability, construct validity and/or responsiveness were reported. The Risk of Bias assessment of these measurement properties ranges from very good to inadequate (Table 4). For the instruments where only internal consistency was reported, the measurement property was assessed as doubtful. For most of the *ad hoc* instruments internal consistency data were reported without information about structural validity or unidimensionality. In such cases, COSMIN recommends rating the standard *'was an internal consistency statistic calculated for each unidimensional scale or subscale separately?'* as doubtful (D). For most of the instruments where responsiveness was reported, the measurement property was assessed as doubtful or inadequate due to an inappropriate measure of responsiveness (Table 4). To assess the reliability a two-week time interval was considered appropriate. In three of the *ad hoc* instruments reliability was assessed to be inadequate due to inappropriate time intervals (Breitkopf et al., 2005; Price, 1994; Weinrich et al., 1992).

The Breast Cancer Knowledge Test (McCance et al., 1990) was assessed as having clear description of the construct to be measured, and cognitive interviews were performed and assumable in a sample representing the target population (McCance et al., 1990). However, the instrument was assessed as having inadequate total quality of PROM development due to unclear descriptions of the target population for which the PROM was developed and the PROM was not developed in a sample representing that same target population (Table 3). Structural validity and internal consistency were reported. However due to unclear description of the methods used, the structural validity was assessed as inadequate, while the internal consistency was assessed as very good.

The Colorectal Cancer Knowledge questionnaire (Weinrich et al., 1992), CRC Knowledge, Perception and Screening Survey (Green and Kelly, 2004) and the CRC Knowledge Assessment Survey (Sanchez et al., 2013) were all assessed as having clear descriptions of the construct to be measured (Table 3). However, the instruments were assessed as having inadequate total quality of PROM development due to unclear descriptions of the target population for which the PROMs were developed and/or the PROMs were not developed in a sample representing those same target populations (Table 3). For these instruments internal consistency was reported and assessed as doubtful due to lack of information about structural validity or unidimensionality.

The knowledge index of the IMQ (Reder et al., 2019) was assessed as having very good description of the construct to be measured and the population for which the PROM was developed. Despite this, the total quality of PROM design was doubtful due to a doubtful concept elicitation and also description of the origin of construct (Table 3). Pilot testing was performed. However, it was doubtful whether the test was performed in a sample representing the target population, and the comprehensibility and comprehensiveness were not investigated. Structural validity and internal consistency were assessed as very good and construct validity as adequate (Table 4).

The CRC-DQI Knowledge subscale (Sepucha et al., 2014) was assessed as having clear descriptions of the construct to be measured and the context of use. The study was performed in a sample representing the target population. Cognitive interviews were performed and also assumable in a sample representing the target population. The examination of comprehensibility was assessed as doubtful due to an unclear description of the method and analysis. However, reporting of comprehensiveness and content validity assessment were unclear (Table 3). Reliability and construct validity were assessed as inadequate. (Table 4).

For the HTKS (Griffin-Mathieu et al., 2022; Haward et al., 2022) the description of the construct to be measured was assessed as very good. Further the origin of the construct and the description of the target population for which the PROM was developed were assessed as very good. However, the PROM development study was not performed in a sample representing the target population for which the PROM was developed. Pilot tests were performed in a sample representing the target population and comprehensibility was assessed. However, the investigation of comprehensibility

was assessed as doubtful due to an unclear description of the method and analysis (Table 3). Structural validity and internal validity were assessed as very good, while criterion validity was assessed as inadequate (Table 4).

The HPV Testing Knowledge scale (Waller et al., 2013) was assessed as having clear descriptions of the construct to be measured (Table 3). However, the instrument was assessed as having inadequate total quality of PROM development due to unclear description of the target population for which the PROM was developed and the PROM was not developed in a sample representing that same target population. Further, there was no information about whether cognitive interviews or other pilot tests were performed (Table 3). Structural validity and internal consistency were assessed as very good, while reliability and construct validity were assessed as doubtful (Table 4).

Evidence of good measurement properties

For most of the *ad hoc* instrument the internal consistency was assessed as indeterminate due to lack of information on the structural validity. Reliability was also assessed as indeterminate because of no information about Intraclass Correlation Coefficient (ICC) (continuous scores) or weighted kappa (dichotomous scores). Although responsiveness in most of the studies was calculated by using an inappropriate method, the results were in accordance with the hypothesis and therefore the responsiveness was assessed as sufficient. For the instruments of Gabel et al. (Gabel et al., 2019) and Wolf et al. (Wolf et al., 2005), structural validity was assessed as indeterminate due to lack of data about local independence and monotonicity (Table 5).

For the Breast Cancer Knowledge Test (McCance et al., 1990) structural validity and internal consistency were assessed as indeterminate due to lack of information about structural validity (Table 5).

The results of the IMQ knowledge index showed indeterminate structural validity and internal consistency, due to a lack of information to assess the criteria of local independence. Further, the results showed sufficient construct validity since the results were in accordance with the hypothesis (Reder et al., 2019) (Table 5).

The results of the CRC-DQI Knowledge subscale showed insufficient reliability due to an ICC lower than 0.70. Further, the results showed sufficient construct validity since the hypotheses were met (Sepucha et al., 2014) (Table 5).

The HPV Testing Knowledge Scale showed indeterminate structural validity and internal consistency due to a lack of report of important aspects of the item response theory (IRT) analysis. Further, indeterminate criterion validity was found (Griffin-Mathieu et al., 2022; Haward et al., 2022) (Table 5).

For the HPV knowledge measure structural validity was assessed as indeterminate due to lack of information about local independence. Therefore, internal consistency was also assessed as indeterminate. Reliability was also assessed as indeterminate because ICC was not reported. Construct validity was also assessed as sufficient because the results were in accordance with the hypothesis (Table 5).

Discussion

This is the first systematic review of the measurement properties of instruments measuring knowledge about breast, colorectal or cervical cancer screening using the COSMIN guidelines. We included 25 studies and 24 different instruments. The included instruments varied in number and characteristics of items, and most of the instruments were *ad hoc* instruments developed for each specific study. All included instruments were assessed as having inadequate PROM development, and there was not enough information to assess content validity of the instruments. Nevertheless, all included instruments were further assessed for risk of bias in development and against the updated criteria for good measurement properties. The results of these assessments showed indeterminate structural validity, internal consistency, criterion validity and reliability, while construct validity and responsiveness were assessed as sufficient. The assessments indicated a lack of focus on content validity including PROM development and structural validity.

COSMIN was developed in 2005 due to a lack of clarity about the terminology and definitions of measurement properties, inconsistency in methods used to determine measurement properties and a lack of explicit and comprehensive criteria for what constitute good measurement properties (COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN)

webpage). Without clear standards the evidence-based selection of measurement instruments is strongly hampered (Mokkink et al., 2006). A systematic review conducted in 2006 found heterogeneity in the measures used in studies of informed decision making for prostate, breast and colorectal cancer screening (Mullen et al., 2006), which is consistent with our findings. Mullen et al. 2006 reported that in many cases, the measurement properties of the instruments were either not or minimally reported and the validity of the measures was reported infrequently (Mullen et al., 2006). The instruments included in the systematic review by Mullen et al. (Mullen et al., 2006) were developed before 2005, and the same applies to several of the instruments included in our systematic review. Before the development of COSMIN, only general guidelines were available (Mokkink et al., 2006), which may explain the inconsistent reporting of measurement properties. The COSMIN methodology to evaluate content validity of PROMs was published in 2018 (Terwee et al., 2018), potentially explaining the inadequate development of several of the included instruments before that date. In general, more recent instruments such as IMQ knowledge index, CRC-DQI Knowledge subscale, HPV Testing Knowledge Scale and HPV Knowledge measure accomplish more of the COSMIN criteria.

Developing a measurement instrument is a challenging and time consuming process (Mokkink et al., 2006). The results of the systematic review indicate that there has been little rigorous development of validated instruments to measure knowledge about cancer screening, despite the intense interest and resources devoted to cancer screening programmes. According to the COSMIN guidelines, this means that study findings are not based on high quality instruments, which affects inferences drawn from trials about improving residents' levels of knowledge about cancer screening. Knowledge about population-based cancer screening is a critical component of informed decision making (Mullen et al., 2006).

The results of the knowledge subscales included in the instruments of IMQ, CRC-DQI, the HPV Testing Knowledge Scale and the HPV Knowledge measure indicate a foundation for well-developed instruments for the measurement of knowledge related to breast, colorectal and cervical cancer screening, respectively. Nevertheless, to ascertain the effectiveness of information materials, which can support informed choices about participation or not in screening, further assessments of measurement properties of these three instruments are essential in order to validate the instruments and develop them further if required. In addition, the COSMIN guidelines suggest both that

literature should be reviewed and that scientific experts and the target population should be involved in developing and testing an instrument, to define and elaborate the instrument's construct, choose the measurement method, select and formulate items and scoring issues. Few studies have investigated what defines relevant knowledge about cancer screening in order to make an informed choice, and the results of these studies demonstrate a lack of consensus on concept (Jepson et al., 2007; Kirkegaard et al., 2016). To conceptualize relevant knowledge about cancer screening it is crucial through qualitative interviews to engage scientific experts within the field and representatives from the target population. Further, the target population should be involved in pilot- and field-testing of the measurement instrument (de Vet et al., 2011). To fulfil these guidelines a content validity study and further evaluation of measurement properties need to be conducted for all three instruments.

Strength and limitations

Strengths of this review include its systematic literature search, which was conducted in collaboration with an experienced medical librarian and the application of the validated COSMIN search filter on measurement properties (Terwee et al., 2009). Second, we used the COSMIN criteria to assess methodological quality of the included studies, supporting transparent conclusions. Strict application of the COSMIN guideline would have led to no further assessment of any of the instruments because of inadequate PROM development and inability to assess content validity. We nevertheless extended the assessments to highlight the extent of existing work, identifying which instruments provide a basis for well-developed instruments for the measurement of knowledge related to breast, colorectal and cervical cancer screening.

This systematic review is limited to studies where the target population is from Europe, North America and Australasia. This eligibility criterion has been chosen to enhance comparability across studies. However, variations may exists depending on the healthcare system and population composition. Further, it only includes colorectal, breast and cervical cancer screening knowledge instruments. Consequently, in both cases we may have excluded potentially high-quality instruments for other screening domains, where instruments to measure knowledge may also be relevant.

Conclusion

In total, 24 instruments to measure knowledge about cancer screening were included and assessed using the COSMIN guidelines in this systematic review. In general, the results indicate inadequate focus on content validity including PROM development, which is considered the most important measurement property by COSMIN. Further assessments of content validity and measurement properties need to be conducted.

Co-author contributions

Louise Stenholt: Literature search, reviewing and editing. **Henrik Hein Lauridsen:** Methodology, reviewing and editing. **Adrian Edwards:** Conceptualization, methodology, reviewing and editing. **Berit Andersen:** Conceptualization, methodology, reviewing and editing. **Mette Bach Larsen:** Conceptualization, formal analysis, review and editing.

Funding

This research did not receive any grants from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of interest

The authors declare no conflict of interest.

Appendix A – Literature search

Information associated with the literature search can be found in appendix A.

Appendix B – excluded studies

The excluded studies can be found in Appendix B.

References

Breitkopf, C.R., Pearson, H.C., Breitkopf, D.M., 2005. Poor knowledge regarding the Pap test among low-income women undergoing routine screening. Perspect Sex Reprod Health 37:78-84. https://doi.org/10.1363/psrh.37.078.05.

Chamot, E., Perneger, T.V., 2002. Men's and women's knowledge and perceptions of breast cancer and mammography screening. Prev Med 34:380-5. https://doi.org/10.1006/pmed.2001.0999.

COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) webpage, About the COSMIN initiative. Retrieved 6 Dec, 2023, from https://www.cosmin.nl/about/.

Council of the European Union, 2022. Council Recommendation on strengthening prevention through early detection: A new EU approach on cancer screening replacing Council Recommendation 2003/878/EC. General Secretariat of the Council Brussels p. 22.

de Vet, H.C.W., Terwee, C.B., Mokkink, L.B., Knol, D.L., 2011. Measurement in Medicine: A Practical Guide. Cambridge University Press.

Denters, M.J., Deutekom, M., Essink-Bot, M.L., Bossuyt, P.M., Fockens, P., Dekker, E., 2015. Assessing knowledge and attitudes towards screening among users of Faecal Immunochemical Test (FIT). Health Expect 18:839-49. https://doi.org/10.1111/hex.12056.

European Colorectal Cancer Screening Guidelines Working Group: von Karsa, L., Patnick, J., Segnan, N., Atkin, W., Halloran, S., Lansdorp-Vogelaar, I., Malila, N., Minozzi, S., Moss, S., et al., 2013. European guidelines for quality assurance in colorectal cancer screening and diagnosis: overview and introduction to the full supplement publication. Endoscopy 45:51-59. https://doi.org/10.1055/s-0032-1325997.

European Commission, Directorate-General for Health Consumers, Anttila, A., Arbyn, M., Jordan, J., 2008. European guidelines for quality assurance in cervical cancer screening. Publications Office. https://doi.org/doi/10.2772/44215.

European Commission, Directorate-General for Health Consumers, Broeders, M., Wolf, C., Perry, N., 2013. European guidelines for quality assurance in breast cancer screening and diagnosis. Publications Office. https://doi.org/10.2772/13196.

European Commission, Directorate-General for Health Consumers, Executive Agency for Health Consumers, World Health Organization, Karsa, L., Patnick, J., Segnan, N., 2010. European guidelines for quality assurance in colorectal cancer screening and diagnosis. Publications Office. https://doi.org/10.2772/1458.

Fowler, B.A., Rodney, M., Roberts, S., Broadus, L., 2005. Collaborative breast health intervention for African American women of lower socioeconomic status. Oncol Nurs Forum 32:1207-16. https://doi.org/10.1188/05.ONF.1207-1216.

Gabel, P., Edwards, A., Kirkegaard, P., Larsen, M.B., Andersen, B., 2020a. The LEAD trial-The effectiveness of a decision aid on decision making among citizens with lower educational attainment who have not participated in FIT-based colorectal cancer screening in Denmark: A randomised controlled trial. Patient Educ Couns 103:359-68. https://doi.org/10.1016/j.pec.2019.08.029.

Gabel, P., Larsen, M.B., Edwards, A., Kirkegaard, P., Andersen, B., 2019. Knowledge, attitudes, and worries among different health literacy groups before receiving first invitation to colorectal cancer screening: Cross-sectional study. Prev Med Rep 14:100876. https://doi.org/10.1016/j.pmedr.2019.100876.

Gabel, P., Larsen, M.B., Edwards, A., Kirkegaard, P., Andersen, B., 2020b. Effectiveness of a decision aid for colorectal cancer screening on components of informed choice according to educational attainment: A randomised controlled trial. PLoS One 15:e0241703. https://doi.org/10.1371/journal.pone.0241703.

Ghanouni, A., Renzi, C., Meisel, S.F., Waller, J., 2016. Common methods of measuring 'informed choice' in screening participation: Challenges and future directions. Prev Med Rep 4:601-07. https://doi.org/10.1016/j.pmedr.2016.10.017.

538 Goel, M.S., Gracia, G., Baker, D.W., 2011. Development and pilot testing of a culturally
539 sensitive multimedia program to improve breast cancer screening in Latina women.
540 Patient Educ Couns 84:128-31. https://doi.org/10.1016/j.pec.2010.05.014.
541 Green, P.M., Kelly, B.A., 2004. Colorectal Cancer Knowledge, Perceptions, and Behaviors in
542 African Americans. Cancer nursing 27:206-15. https://doi.org/10.1097/00002820-
543 200405000-00004.
544 Griffin-Mathieu, G., Haward, B., Tatar, O., Zhu, P., Perez, S., Shapiro, G.K., McBride, E.,
545 Thompson, E.L., Smith, L.W., et al., 2022. Ensuring a Successful Transition From
546 Cytology to Human Papillomavirus-Based Primary Cervical Cancer Screening in Canada
547 by Investigating the Psychosocial Correlates of Women's Intentions: Protocol for an
548 Observational Study. JMIR Res Protoc 11:e38917. https://doi.org/10.2196/38917.
549 Haward, B., Tatar, O., Zhu, P., Griffin-Mathieu, G., McBride, E., Waller, J., Brotherton, J.,
550 Lofters, A., Mayrand, M.H., et al., 2023. Are Canadian Women Prepared for the
551 Transition to Primary HPV Testing in Cervical Screening? A National Survey of
552 Knowledge, Attitudes, and Beliefs. Curr Oncol 30:7055-72.
553 https://doi.org/10.3390/curroncol30070512.
554 Haward, B., Tatar, O., Zhu, P., Griffin-Mathieu, G., Perez, S., Shapiro, G.K., McBride, E.,
555 Zimet, G.D., Rosberger, Z., 2022. Development and validation of the cervical cancer
556 knowledge scale and HPV testing knowledge scale in a sample of Canadian women. Prev
557 Med Rep 30:102017. https://doi.org/10.1016/j.pmedr.2022.102017.
558 Hersch, J., Barratt, A., Jansen, J., Irwig, L., McGeechan, K., Jacklyn, G., Thornton, H., Dhillon,
559 H., Houssami, N., et al., 2015. Use of a decision aid including information on
560 overdetection to support informed choice about breast cancer screening: a randomised
561 controlled trial. Lancet 385:1642-52. https://doi.org/10.1016/s0140-6736(15)60123-4.
562 Hickey, S., Merz, E.L., Malcarne, V.L., Gunsauls, D.C., Huang, J., Sadler, G.R., 2013. Breast
563 cancer education for the Deaf community in American Sign Language. Oncol Nurs
564 Forum 40:E86-91. https://doi.org/10.1188/13.ONF.E86-E91.
565 Jepson, R.G., Hewison, J., Thompson, A., Weller, D., 2007. Patient perspectives on information
566 and choice in cancer screening: a qualitative study in the UK. Soc Sci Med 65:890-9.
567 https://doi.org/10.1016/j.socscimed.2007.04.009.
568 Kagawa-Singer, M., Tanjasiri, S.P., Valdez, A., Yu, H., Foo, M.A., 2009. Outcomes of a breast
569 health project for Hmong women and men in California. Am J Public Health 99 Suppl
570 2:S467-73. https://doi.org/10.2105/AJPH.2008.143974.
571 Kirkegaard, P., Mortensen, G.L., Mortensen, S.L., Larsen, M.B., Gabel, P., Andersen, B., 2016.
572 Making decisions about colorectal cancer screening. A qualitative study among citizens
573 with lower educational attainment. Eur J Public Health 26:176-81.
574 https://doi.org/10.1093/eurpub/ckv207.
575 Larsen, M.B., Stokholm, R., Kirkegaard, P., Laursen, H.S., Gabel, P., Andersen, B., 2021.
576 Making decisions on your own: Self-administered decision aids about colorectal cancer
577 screening – A systematic review and meta-analyses. Patient Education and Counseling.
578 https://doi.org/https://doi.org/10.1016/j.pec.2021.07.035.
579 Lee, H., Ghebre, R., Le, C., Jang, Y.J., Sharratt, M., Yee, D., 2017. Mobile Phone Multilevel and
580 Multimedia Messaging Intervention for Breast Cancer Screening: Pilot Randomized
581 Controlled Trial. JMIR MHEALTH AND UHEALTH 5.
582 https://doi.org/10.2196/mhealth.7091.
583 Marteau, T.M., Dormandy, E., Michie, S., 2001. A measure of informed choice. Health Expect
584 4:99-108. https://doi.org/10.1046/j.1369-6513.2001.00140.x.
585 Mathieu, E., Barratt, A., Davey, H.M., McGeechan, K., Howard, K., Houssami, N., 2007.
586 Informed choice in mammography screening: a randomized trial of a decision aid for
587 70-year-old women. Arch Intern Med 167:2039-46.
588 https://doi.org/10.1001/archinte.167.19.2039.
589 Mathieu, E., Barratt, A.L., McGeechan, K., Davey, H.M., Howard, K., Houssami, N., 2010.
590 Helping women make choices about mammography screening: an online randomized
591 trial of a decision aid for 40-year-old women. Patient Educ Couns 81:63-72.
592 https://doi.org/10.1016/j.pec.2010.01.001.

McCance, K.L., Mooney, K.H., Smith, K.R., Field, R., 1990. Validity and reliability of a breast cancer knowledge test. Am J Prev Med 6:93-8.
Mokkink, L.B., Prinsen, C. A., Patrick, D. L., Alonso, J., Bouter, L. M., de Vet, H. C., & Terwee, C. B, 2018. COSMIN methodology for systematic reviews of Patient-Reported Outcome Measures (PROMs) – user manual. from www.cosmin.nl.
Mokkink, L.B., Terwee, C.B., Knol, D.L., Stratford, P.W., Alonso, J., Patrick, D.L., Bouter, L.M., de Vet, H.C.W., 2006. Protocol of the COSMIN study: COnsensus-based Standards for the selection of health Measurement INstruments. BMC medical research methodology 6:2-2. https://doi.org/10.1186/1471-2288-6-2.
Mullen, P.D., Allen, J.D., Glanz, K., Fernandez, M.E., Bowen, D.J., Pruitt, S.L., Glenn, B.A., Pignone, M., 2006. Measures used in studies of informed decision making about cancer screening: a systematic review. Ann Behav Med 32:188-201. https://doi.org/10.1207/s15324796abm3203_4.
Peterson, N.B., Dwyer, K.A., Mulvaney, S.A., Dietrich, M.S., Rothman, R.L., 2007. The influence of health literacy on colorectal cancer screening knowledge, beliefs and behavior. J Natl Med Assoc 99:1105-12.
Price, J.H., 1994. Economically disadvantaged females' perceptions of breast cancer and breast cancer screening. Journal of the National Medical Association 86:899-906.
Prinsen, C.A.C., Mokkink, L.B., Bouter, L.M., Alonso, J., Patrick, D.L., de Vet, H.C.W., Terwee, C.B., 2018. COSMIN guideline for systematic reviews of patient-reported outcome measures. Quality of life research 27:1147-57. https://doi.org/10.1007/s11136-018-1798-3.
Ramírez-Amill, R., Soto-Salgado, M., Vázquez-Santos, C., Corzo-Pedrosa, M., Cruz-Correa, M., 2017. Assessing Colorectal Cancer Knowledge Among Puerto Rican Hispanics: Implications for Cancer Prevention and Control. Journal of Community Health 42:1141-47. https://doi.org/10.1007/s10900-017-0363-2.
Reder, M., Berens, E.-M., Spallek, J., Kolip, P., 2019. Development of the Informed Choice in Mammography Screening Questionnaire (IMQ): factor structure, reliability, and validity. BMC PSYCHOLOGY 7. https://doi.org/10.1186/s40359-019-0291-2.
Reder, M., Kolip, P., 2017. Does a decision aid improve informed choice in mammography screening? Results from a randomised controlled trial. PLoS One 12:e0189148. https://doi.org/10.1371/journal.pone.0189148.
Roberto, A., Colombo, C., Candiani, G., Satolli, R., Giordano, L., Jaramillo, L., Castagno, R., Mantellini, P., Falini, P., et al., 2020. A dynamic web-based decision aid to improve informed choice in organised breast cancer screening. A pragmatic randomised trial in Italy. Br J Cancer 123:714-21. https://doi.org/10.1038/s41416-020-0935-2.
Sanchez, J.I., Palacios, R., Thompson, B., Martinez, V., O'Connell, M.A., 2013. Assessing Colorectal Cancer Screening Behaviors and Knowledge among At-Risk Hispanics in Southern New Mexico. J Cancer Ther 4:15-25. https://doi.org/10.4236/jct.2013.46A2003.
Scholtes, V.A., Terwee, C.B., Poolman, R.W., 2011. What makes a measurement instrument valid and reliable? Injury 42:236-40. https://doi.org/10.1016/j.injury.2010.11.042.
Schonberg, M.A., Hamel, M.B., Davis, R.B., Griggs, M.C., Wee, C.C., Fagerlin, A., Marcantonio, E.R., 2014. Development and Evaluation of a Decision Aid on Mammography Screening for Women 75 Years and Older. JAMA INTERNAL MEDICINE 174:417-24. https://doi.org/10.1001/jamainternmed.2013.13639.
Sepucha, K.R., Feibelmann, S., Cosenza, C., Levin, C.A., Pignone, M., 2014. Development and evaluation of a new survey instrument to measure the quality of colorectal cancer screening decisions. BMC Med Inform Decis Mak 14:72. https://doi.org/10.1186/1472-6947-14-72.
Shokar, N.K., Carlson, C.A., Weller, S.C., 2008. Factors associated with racial/ethnic differences in colorectal cancer screening. J Am Board Fam Med 21:414-26. https://doi.org/10.3122/jabfm.2008.05.070266.
Smith, S.G., Raine, R., Obichere, A., Wolf, M.S., Wardle, J., von Wagner, C., 2015. The effect of a supplementary ('Gist-based') information leaflet on colorectal cancer knowledge

and screening intention: a randomized controlled trial. J Behav Med 38:261-72. https://doi.org/10.1007/s10865-014-9596-z.
Smith, S.K., Barratt, A., Trevena, L., Simpson, J.M., Jansen, J., McCaffery, K.J., 2012. A theoretical framework for measuring knowledge in screening decision aid trials. Patient Educ Couns 89:330-6. https://doi.org/10.1016/j.pec.2012.07.009.
Smith, S.K., Trevena, L., Simpson, J.M., Barratt, A., Nutbeam, D., McCaffery, K.J., 2010. A decision aid to support informed choices about bowel cancer screening among adults with low education: randomised controlled trial. Bmj 341:c5370. https://doi.org/10.1136/bmj.c5370.
Steckelberg, A., Hülfenhaus, C., Haastert, B., Mühlhauser, I., 2011. Effect of evidence based risk information on "informed choice" in colorectal cancer screening: randomised controlled trial. BMJ 342:d3193. https://doi.org/10.1136/bmj.d3193.
Terwee, C.B., Bot, S.D.M., de Boer, M.R., van der Windt, D.A.W.M., Knol, D.L., Dekker, J., Bouter, L.M., de Vet, H.C.W., 2007. Quality criteria were proposed for measurement properties of health status questionnaires. Journal of clinical epidemiology 60:34-42. https://doi.org/10.1016/j.jclinepi.2006.03.012.
Terwee, C.B., Jansma, E.P., Riphagen, I.I., de Vet, H.C.W., 2009. Development of a Methodological PubMed Search Filter for Finding Studies on Measurement Properties of Measurement Instruments. Quality of Life Research 18:1115-23. https://doi.org/10.1007/s11136-009-9528-5.
Terwee, C.B., Prinsen, C.A.C., Chiarotto, A., Westerman, M.J., Patrick, D.L., Alonso, J., Bouter, L.M., de Vet, H.C.W., Mokkink, L.B., 2018. COSMIN methodology for evaluating the content validity of patient-reported outcome measures: a Delphi study. Qual Life Res 27:1159-70. https://doi.org/10.1007/s11136-018-1829-0.
Terwee, C.B., Prinsen, C. A., Chiarotto, A., de Vet, H. C., Bouter, L. M., Alonso, J., Westerman, M. J., Patrick, D. L., & Mokkink, L. B., 2018. COSMIN methodology for assessing the content validity of PROMs – user manual. from www.cosmin.nl.
Trevena, L.J., Irwig, L., Barratt, A., 2008. Randomized trial of a self-administered decision aid for colorectal cancer screening. J Med Screen 15:76-82. https://doi.org/10.1258/jms.2008.007110.
van Agt, H., Fracheboud, J., van der Steen, A., de Koning, H., 2012. Do women make an informed choice about participating in breast cancer screening? A survey among women invited for a first mammography screening examination. Patient Educ Couns 89:353-9. https://doi.org/10.1016/j.pec.2012.08.003.
Waller, J., Ostini, R., Marlow, L.A.V., McCaffery, K., Zimet, G., 2013. Validation of a measure of knowledge about human papillomavirus (HPV) using item response theory and classical test theory. Preventive medicine 56:35-40. https://doi.org/10.1016/j.ypmed.2012.10.028.
Weinrich, S.P., Weinrich, M.C., Boyd, M.D., Johnson, E., Frank-Stromborg, M., 1992. Knowledge of colorectal cancer among older persons. Cancer nursing 15:322-30. https://doi.org/10.1097/00002820-199210000-00002.
Wolf, M.S., Rademaker, A., Bennett, C.L., Ferreira, M.R., Dolan, N.C., Davis, T.C., Medio, F., Liu, D., Lee, J., et al., 2005. Development of a brief survey on colon cancer screening knowledge and attitudes among veterans. Prev Chronic Dis 2:A11.

694

695

696