BMJ Open

Instruments for measuring patient health education competence among nursing staff: Protocol for a COSMIN-based systematic review

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2023-072905
Article Type:	Protocol
Date Submitted by the Author:	16-Feb-2023
Complete List of Authors:	Wang, Shuyi; Central South University, Xiangya School of Nursing Liu, Ke; Central South University, Xiangya School of Nursing SHI, Zeya; Hunan Provincial People's Hospital, Chen, Qirong; Central South University, Xiangya School of Nursing; Xiangya Center for Evidence-Based Practice & Healthcare Innovation: A Joanna Briggs Institute Affiliated Group Tang, Siyuan; Central South University, Xiangya School of Nursing; XiangyaCenterforEvidence-BasedNursingPractice&HealthcareInnovation:AJBIAffiliatedGroup
Keywords:	Health Education, PUBLIC HEALTH, Primary Health Care

Instruments for measuring patient health education competence among nursing staff: Protocol for a COSMIN-based systematic review

Shuyi Wang[a], Ke Liu[a], Zeya Shi[b], Qirong Chen[a,c,*], Siyuan Tang[a,c,*]

a. Xiangya School of Nursing, Central South University,

172 Tongzipo Road, Changsha, Hunan, 410013, China.

b. Nursing Department, Hunan Provincial People's Hospital,

61 Jiefang West Road, Changsha 410005, China

c. Xiangya Center for Evidence-Based Practice & Healthcare Innovation: A Joanna Briggs Institute Affiliated Group,

172 Tongzipo Road, Changsha, Hunan, 410013, China.

***Corresponding author:**

Qirong Chen

Address: 172 Tongzipo Road, Changsha, Hunan, 410013, China.

Phone number: (86)15084726637

Email address: qirong.chen@csu.edu.cn

Siyuan Tang

Address: 172 Tongzipo Road, Changsha, Hunan, 410013, China.

Email address: sytang263@csu.edu.cn

Phone number: (86)15874107355

Keywords: Health education; Competency assessment; Nursing staff; Evaluation; Psychometrics

Word count: 2303

ABSTRACT

Introduction Health education, as a crucial strategic measure of disease prevention and

control in the 21st century, has become an important part of healthcare. As the main deliverers of patient health education, nursing staff's patient health education competence (PHEC) has received much attention. Instruments for assessing the PHEC of nursing staff have been developed internationally, but there is a lack of systematic reviews and evaluations of the psychometric properties of these instruments. To effectively select appropriate PHEC assessment instruments in specific contexts, a systematic and comprehensive review and evaluation of these measurement instruments are needed. The goal of this systematic review is to systematically evaluate the psychometric properties of existing PHEC instruments.

Methods and analysis In this study, eight databases will be searched between March 1, 2023, and March 31, 2023, to retrieve studies that include instrument(s) measuring the PHEC of nursing staff. Two researchers will independently perform literature screening, data extraction, and literature evaluation. In case of disagreement, a third researcher will be involved in the resolution. The measurement properties of PHEC assessment instruments will be systematically reviewed based on the consensus-based standards for the selection of health measurement instruments (COMSIN) methodology and guideline.

Ethics and dissemination Ethical approval is not applicable for this study. We will share the findings from the study at national and/or international conferences and in a peer-reviewed journal in the fields of health education and/or patient education.

PROSPERO registration number CRD42023393293

Strengths and limitations of this study

- To our knowledge, this will be the first COSMIN-based systematic review of patient health education competence (PHEC) assessment instruments for nursing staff.

- The Preferred Reporting Items for Systematic Reviews and Meta-analyses protocols (PRISMA-P) 2020 checklist will be used to guide the implementation and report of the protocol and systematic review.
- The consensus-based standards for the selection of health measurement instruments (COSMIN) methodology will be used to evaluate the methodological quality of included studies on measurement properties of the instruments and the quality of included instruments.
- The systematic review may fail to include relevant literature published outside of the searched databases.

1. INTRODUCTION

Health education has been identified by the World Health Organization (WHO) as one of the three crucial strategic measures of disease prevention and control in the 21st century, and it is the most economical and effective measure for improving public health.[1] Health education for patients is a crucial part of the healthcare services provided by health professionals. In clinical practice, health professionals are often required to develop patient education programs to impart to patients knowledge and skills for health restoration and promotion.[2,3] Health education for patients can improve their understanding of their own health status and disease management measures, which can relieve patients' anxiety and improve their compliance and satisfaction with medical staff, thus improving their health status and quality of life.[4] These better patient outcomes could reduce the burden of disease on patients and society at the economic level.[5,6] As the world's largest group of health professionals and the health professionals who have the closest contact with patients, nursing staff plays an important role in patient health education.[5,7]

Patient health education competency (PHEC) refers to the specific qualities that

health educators should have to conduct effective health education activities with patients.[8,9] PHEC is an essential professional competency for nursing staff and determines the quality of patient education.[10-13] However, in existing studies, the PHEC of clinical nurses is often the lowest-rated area of nursing competency.[14,15] Therefore, the development and strengthening of PHEC for nurses are extremely important to improve the quality of patient education, patient care, patient safety, and the development of nursing careers. In addition, we should pay attention to nursing students' PHEC because they are the mainstay of the clinical nurse workforce.

Accurate measurement of PHEC is important because it can be used to assess the PHEC status of nursing staff and to develop targeted strategies based on the nursing staff's PHEC. Moreover, it can be used in research to assess the effectiveness of relevant PHEC interventions. Currently, relevant measurement instruments have been developed internationally: for example, a scale for measuring the PHEC of registered nurses developed by Lin et al. in 2017,[16] a PHEC competency assessment scale developed by Hwang et al. based on a literature review and the Delphi method,[17] and a Spanish version of the nurse PHEC scale developed by Pueyo-Garrigues et al.[18] Although related instruments are available for assessing PHEC in nursing staff, these evaluation instruments have been developed in different settings and their validation varies considerably, with none considered the gold standard.

In this study, we defined PHEC as the specific qualities that must be possessed by health educators to provide health education to patients, including knowledge, skills, beliefs or attitudes, self-concept, personality qualities, and motivation. Although there has been a review of PHEC measurement instruments for nursing staff, this review has some limitations on its rigor.[19] First, this review included not only measurement instruments for PHEC but also systems for evaluating PHEC, which are different from

measurement instruments. Second, this review did not systematically evaluate the measurement properties of instruments for measuring PHEC based on related guidelines. However, a systematic and comprehensive review of PHEC measurement instruments is crucial for guiding the selection of instruments and/or guiding the development and refinement of high-quality instruments in the future. The consensus-based standards for the selection of health measurement instruments (COSMIN) methodology provides resources to systematically review measurement instruments and evaluate them in terms of both methodological quality and quality of measurement properties to select instruments that are of high quality for study purposes and provide an evidence-based foundation for future high-level instrument development.[20] Thus, this study will conduct a comprehensive and rigorous systematic review of PHEC assessment instruments based on the COSMIN methodology, which aims to evaluate the measurement properties of these instruments, provide a reference for nursing staff and researchers to accurately and effectively assess PHEC, and provide recommendations for researchers to develop and improve PHEC assessment instruments.

This systematic review will answer the following questions: (1) What instruments are available for assessing the PHEC of nursing staff? (2) What are the characteristics of these instruments? (3) What is the methodological quality of studies on the measurement properties of these instruments? (4) How about these instruments' measurement properties, interpretability, and feasibility? (5) What are the similarities and differences between these instruments? (6) What are the knowledge and research gaps in this area?

2. METHODS

The review proposed by this protocol will follow the COSMIN methodology for

conducting systematic reviews of psychometric properties and will be reported following the Preferred Reporting Items for Systematic Reviews and Meta-analyses protocols (PRISMA-P) 2020 checklist.[21,22] We registered the protocol in the International Prospective Register of Systematic Reviews (PROSPERO, CRD42023393293).

Inclusion and exclusion criteria of studies

Inclusion criteria

Studies will be included if they (1) address instrument(s) for measuring the PHEC of nurses or nursing students, (2) describe the processes of development and evaluation of one or more measurement properties for eligible instrument(s), (3) discuss instruments designed to measure the PHEC of health professionals (the literature explicitly mentions that it applies to nursing staff as well), and (4) have full-text availability. If full-text versions of the studies are not available online, the authors of these articles will be contacted, and articles for which valid information was not available after contacting the authors will be excluded.

Exclusion criteria

Studies will be excluded if they are (1) not primary studies (e.g. biographies, addresses, and editorials) or are case studies, (2) reports that used the instruments only for outcome measurements, (3) secondary studies (e.g. reviews and/or systematic reviews), or (4) duplicate published studies.

Search strategy

A systematic search will be performed between March 1, 2023, and March 31, 2023, in six English databases (i.e. CINAHL, EMBASE, Ovid Medline, PubMed, PsycINFO, and Web of Science) and two Chinese databases (i.e. CNKI and WANFANG DATA). We include Chinese databases since the researchers speak Chinese as their native

language. We will also search for references in all eligible literature to prevent omissions. The search time limit is from the library's creation date to the search date. A literature search will be conducted using a combination of subject terms and free words. The major search concepts will be nursing, health education, competence, instrument, and measurement properties. Related comprehensive and sensitive search strategies developed by other researchers will also be used in this literature search, including (1) the search filter developed by the University of Oxford for finding PROMs,[23] (2) the sensitive PubMed search filter for measuring attributes developed by Terwee et al., and (3) corresponding search filters applicable to other databases.[24] Our study will examine results reported by nurses or nursing students, so the first filter will be adjusted appropriately (e.g. we will remove those sections that are relevant to the quality of life and patient-reported outcomes). The search strategy constructed for PubMed is described in Table S1 in the supplementary file. The search strategy for the Chinese databases is shown in Table S2 in the supplementary file.

Study screening

Covidence will be used to manage the references. First, duplicates from the eight databases will be removed with Covidence. After the initial screening, both researchers will independently review and screen titles, abstracts, and full-text articles with the support of Covidence. In case of disagreement, a third researcher will be consulted to screen the literature. The screening processes of this study are shown in Figure 1.

Data extraction

The two researchers will independently extract data from the included papers and resolve their differences through discussion. We will extract the data on the characteristics of the instruments (including instrument name, developer(s)/year developed, construct(s), targeted population, mode of administration, recall period,

(sub)scale(s)/(number of items), response options, range of scores/scoring, original language, and available translations; see Table S3 in the supplementary file), the characteristics of the included populations (including sample size, age, gender, setting, country, and language; see Table S4 in the supplementary file,), the results on the psychometric properties (Table S5 in the supplementary file), and information about the interpretability (Table S6 in the supplementary file) and feasibility (Table S7 in the supplementary file) of the included instruments.

The term 'outcome measure instrument development' will be used instead of the original 'patient-reported outcome measure development' to more accurately reflect the inclusion of studies that examined outcomes reported by nurses or nursing students rather than patients.

Quality appraisal and Data synthesis

Two researchers will independently assess the quality of eligible studies using the COSMIN Risk of Bias checklist, which is divided into three sections: content validity (instrument development and content validity), internal structure (structural validity, internal consistency, and cross-cultural validity/measurement invariance), and other measurement properties (reliability, measurement error, criterion validity, hypothesis testing for construct validity, and responsiveness).[20,21,25] Each measurement property will be evaluated with 3 to 35 items, and the items will be rated on a five-level score of 'very good', 'adequate', 'doubtful', 'inadequate', or 'not applicable.' Based on the 'the worst score counts' principle, each measurement property's overall methodological quality score is expressed by taking the lowest rating of any standard in the box. Subsequently, the two researchers will apply the updated criteria for good measurement properties alone to evaluate the reliability and validity of the instruments themselves, and the quality of the evidence will be graded using the Grading of

Recommendations Assessment, Development, and Evaluation (GRADE) approach. In case of disagreement, a third researcher will be consulted.

We will work according to the following three steps. In the first step, two investigators will apply the COSMIN Risk of Bias checklist to evaluate the methodological quality of each eligible study individually.[25] The final consensus on the results of the methodological quality will be presented in Tables S8 and S8-1 in the supplementary file. In the second step, the updated criteria for good measurement properties will be applied to evaluate the quality of evidence for each measured property, and the evaluation results will be shown in Tables S5 and S5-1 in the supplementary file.[21,26] This section mainly evaluates the strengths and weaknesses of the measurement properties. Among these, the quality of content validity will be evaluated according to the COSMIN methodology for content validity in three aspects: the relevance, comprehensiveness, and comprehensibility of items, which can be 'sufficient (+)', 'insufficient (-)', 'indeterminate (?)', or 'inconsistent (±)'.[26,27] The quality of the remaining measurement properties (structural validity, internal consistency, cross-cultural validity, measurement invariance, reliability, measurement error, criterion validity, construct validity, and responsiveness) will be evaluated by applying the COSMIN quality criteria, which can be 'sufficient (+)', 'insufficient (-)', and 'indeterminate (?)'.[21] The corresponding results will be reported in the rating columns of Table S5 in the supplementary file, and the results of rating content validity will be presented separately in Table S5-1 in the supplementary file. In the third step, a modified GRADE approach will be used to rate the quality of the above evidence, reflecting the level of confidence in the quality of the evidence. To evaluate the content's validity, three of these factors are applicable: risk of bias, inconsistency, and indirectness.[27] Assuming that the level of evidence quality for each of the remaining

measurement properties is high, the quality of the evidence will be downgraded by considering the following factors: risk of bias, inconsistency, imprecision, and indirectness.[21] The quality of evidence will be divided into four levels: 'high', 'moderate', 'low', or 'very low'.[20,21] The corresponding results will be displayed in Table S9 in the supplementary file. Two investigators will independently grade and cross-check the results. In case of disputes, final decisions will be made in consultation with the third investigator.

Patient and public involvement

Neither patients nor the public will be involved in this study.

Ethics and dissemination

Ethical approval is not applicable for this study. We will share the findings from the study at national and/or international conferences and in a peer-reviewed journal in the fields of health education and/or patient education.

3. DISCUSSION

To our knowledge, this will be the first COSMIN-based systematic review of PHEC assessment instruments for nursing staff, which will be reported following the Preferred Reporting Items for Systematic Reviews and Meta-analyses protocols (PRISMA-P) 2020 checklist. This systematic review will provide a comprehensive rating of the level of evidence for each measurement property of the PHEC assessment instruments, which will be based on an evaluation of the measurement properties of all included instruments and the methodological quality of the studies. Through this study, we will be able to develop recommendations on the use of existing qualified instruments in clinical practice and research that could assist nursing staff and researchers in the accurate and valid assessment of PHEC. This review may provide an evidence-based foundation for the development, design, validation, and use of future instruments by

identifying problems in instrument development and validation and therefore help researchers to develop and improve these instruments.

Author contributions All authors have read and agreed to the published version of the manuscript.

Conceptualization: QC, ST, SW;

Methodology: QC, ST, ZS;

Data curation: QC, SW, KL;

Writing—original draft preparation: SW, KL, QC;

Writing—review and editing: QC, ST, ZS;

Supervision: QC and ST.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient consent Not required.

Data sharing statement No additional data available.

List of Figures:

List of Supplementary File Tables:

Table S5-1: Rating of the content validity of instruments

Table S6: Information on interpretability of instruments

Table S7: Information on feasibility of instruments

Table S8: Quality of studies on measurement properties

Table S8-1: Quality of the instrument development

Table S9: Quality of the evidence for measurement properties of the instruments

References:

1. Organization WH. Health education: theoretical concepts, effective strategies and core competencies: a foundation document to guide capacity development of health educators. 2012. Available from: https://apps.who.int/iris/bitstream/handle/10665/119953/EMRPUB_2012_EN_1362.pdf?sequen
2. Russell KA. Nurse practice acts guide and govern nursing practice. *J Nurs Regul* 2012;3(3):36-42. doi: https://doi.org/10.1016/S2155-8256(15)30197-6
3. Wittink H, Oosterhaven J. Patient education and health literacy. *Musculoskelet Sci Pract* 2018;38:120-7. doi: 10.1016/j.msksp.2018.06.004.
4. Turner A, Anderson JK, Wallace LM, et al. An evaluation of a self-management program for patients with long-term conditions. *Patient Educ Couns* 2015;98(2):213-9. doi: 10.1016/j.pec.2014.08.022.
5. Lin P, Gao C, Su S, et al. Teaching principles: The applications in nursing practice. Taipei: Farseeing. 2005.
6. Siegel KR, Ali MK, Zhou X, et al. Cost-effectiveness of interventions to manage diabetes: has the evidence changed since 2008? *Diabetes Care* 2020;43(7):1557-92. doi: 10.2337/dci20-0017.

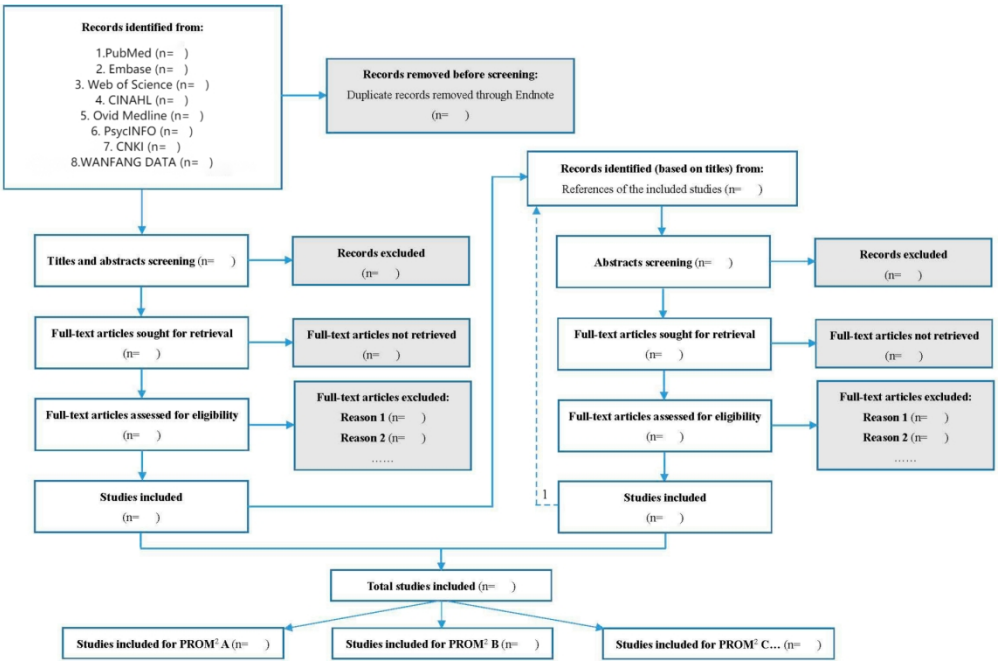
7. Koutsopoulou S, Papathanassoglou ED, Katapodi MC, et al. Nurses as information providers to cancer patients: Critical review of the evidence. *J Clin Nurs* 2010;19(5-6):749-65. doi: 10.1111/j.1365-2702.2009.02954.x.
8. Hwang H-L, Kuo T-Y. Competency in delivering health education: A concept analysis. *J Interprof Educ Pract* 2018;11:20-5. doi: https://doi.org/10.1016/j.xjep.2018.02.005
9. Karimi Moonaghi H, Emami Zeydi A, Mirhaghi A. Patient education among nurses: bringing evidence into clinical applicability in Iran. *Invest Educ Enferm* 2016;34(1):137-51. doi: 10.17533/udea.iee.v34n1a16.
10. Flinkman M, Leino-Kilpi H, Numminen O, et al. Nurse Competence Scale: a systematic and psychometric review. *J Adv Nurs* 2017;73(5):1035-50. doi: 10.1111/jan.13183.
11. Ko Y, Yu S. Core nursing competency assessment tool for graduates of outcome-based nursing education in South Korea: A validation study. *Jpn J Nurs Sci* 2019;16(2):155-71. doi: 10.1111/jjns.12223.
12. Nilsson J, Engström M, Florin J, et al. A short version of the nurse professional competence scale for measuring nurses' self-reported competence. *Nurse Educ Today* 2018;71:233-9. doi: 10.1016/j.nedt.2018.09.028.
13. Wilkinson CA. Competency assessment tools for registered nurses: An integrative review. *J Contin Educ Nurs* 2013;44(1):31-7. doi: 10.3928/00220124-20121101-53.
14. Cowan DT, Wilson-Barnett DJ, Norman IJ, et al. Measuring nursing competence: Development of a self-assessment tool for general nurses across Europe. *Int J Nurs Stud* 2008;45(6):902-13. doi: 10.1016/j.ijnurstu.2007.03.004.
15. Pueyo-Garrigues M, Whitehead D, Pardavila-Belio MI, et al. Health education: a

Rogerian concept analysis. *Int J Nurs Stud* 2019;94:131-8. doi: 10.1016/j.ijnurstu.2019.03.005.

16. Lin LY, Wang RH. Patient Education Competence Scale for Registered Nurses in Taiwan: Scale development and psychometric validation. *Jpn J Nurs Sci* 2017;14(2):117-25. doi: 10.1111/jjns.12141.
17. Hwang HL, Kuo ML, Tu CT. Health education and competency scale: Development and testing. *J Clin Nurs* 2018;27(3-4):e658-e67. doi: 10.1111/jocn.14116.
18. Pueyo-Garrigues M, Pardavila-Belio MI, Whitehead D, et al. Nurses' knowledge, skills and personal attributes for competent health education practice: An instrument development and psychometric validation study. *J Adv Nurs* 2021;77(2):715-28. doi: 10.1111/jan.14632.
19. Mi Y, Wu D, Wei ZZ, et al. Research progress on evaluation tool for health education competency of nursing staff. *Chinese Nursing Research* 2020;34(11):1983-7.
20. Prinsen CA, Mokkink LB, Bouter LM, et al. COSMIN guideline for systematic reviews of patient-reported outcome measures. *Qual Life Res* 2018;27:1147-57. doi: 10.1007/s11136-018-1798-3.
21. Mokkink LB, Prinsen C, Patrick DL, et al. COSMIN methodology for systematic reviews of patient-reported outcome measures (PROMs). User manual. 2018;78(1):6-63. Available from: https://www.cosmin.nl/wp-content/uploads/COSMIN-syst-review-for-PROMs-manual_version-1_feb-2018-1.pdf
22. Page MJ, Moher D, Bossuyt PM, et al. PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. *BMJ* 2021;372.

doi: 10.1136/bmj.n160.

23. Mackintosh A, Comabella C, Hadi M, et al. PROM group construct & instrument type filters. Department of Public Health, University of Oxford, Oxford. 2010. Available from: https://cosmin.nl/wp-content/uploads/prom-search-filter-oxford-2010.pdf
24. Terwee CB, Jansma EP, Riphagen II, et al. Development of a methodological PubMed search filter for finding studies on measurement properties of measurement instruments. *Qual Life Res* 2009;18:1115-23. doi: 10.1007/s11136-009-9528-5.
25. Mokkink LB, De Vet HC, Prinsen CA, et al. COSMIN risk of bias checklist for systematic reviews of patient-reported outcome measures. *Qual Life Res* 2018;27:1171-9. doi: 10.1007/s11136-017-1765-4.
26. Terwee CB, Prinsen CA, Chiarotto A, et al. COSMIN methodology for evaluating the content validity of patient-reported outcome measures: a Delphi study. Qual Life Res 2018;27:1159-70. doi: 10.1007/s11136-018-1829-0.
27. Terwee CB, Prinsen C, Chiarotto A, et al. COSMIN methodology for assessing the content validity of PROMs–user manual. Amsterdam: VU University Medical Center. 2018. Available from: https://cosmin.nl/wp-content/uploads/COSMIN-methodology-for-content-validity-user-manual-v1.pdf



705x463mm (72 x 72 DPI)

Table S1. Search strategy for PubMed

1	#1	nurs*
2	#2	((((((((((Health Education[Mesh] OR Patient Education as Topic[Mesh]) OR (health education*[tiab])) OR (education, health[tiab])) OR (patient education*[tiab])) OR (education program*[tiab])) OR (educational program*[tiab])) OR (education, patient*[tiab])) OR (education of patient*[tiab])) OR (nursing education*[tiab])) OR (hospital education*[tiab])) OR (educational activit*[tiab])) OR ((educat*[tiab]) AND (individual patient*[tiab]))
3	#3	((((((((professional competence[MeSH]) OR (professional competence[tiab])) OR (competence, professional[tiab])) OR (generalization of expertise[tiab])) OR (expertise generalization[tiab])) OR (technical expertise[tiab])) OR (expertise, technical[tiab])) OR (competenc*[tiab] OR capabilit*[tiab] OR capacit*[tiab] OR abilit*[tiab])) OR ((skill*[tiab] OR belief[tiab] OR attitude[tiab]) AND (knowledge[tiab]))
4	#4	(report[tiab] OR reported[tiab] OR reporting[tiab] OR rated[tiab] OR rating[tiab] OR ratings[tiab] OR based[tiab] OR assessed[tiab] OR assessment[tiab] OR assessments[tiab] OR disability[tiab] OR function[tiab] OR functional[tiab] OR functions[tiab] OR subjective[tiab] OR utility[tiab] OR utilities[tiab] OR wellbeing[tiab] OR well being[tiab]) AND (index[tiab] OR indices[tiab] OR instrument[tiab] OR instruments[tiab] OR measure[tiab] OR measures[tiab] OR questionnaire[tiab] OR questionnaires[tiab] OR profile[tiab] OR profiles[tiab] OR scale[tiab] OR scales[tiab] OR score[tiab] OR scores[tiab] OR status[tiab] OR survey[tiab] OR surveys[tiab])
5	#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 reproducib*[tiab] OR "discriminant analysis"[MeSH] OR reliab*[tiab] OR unreliab*[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 alphas[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 intertester[tiab] OR inter-tester[tiab] OR intratester[tiab] OR intra-tester[tiab] OR interobserver[tiab] OR inter-observer[tiab] OR intraobserver[tiab] OR intra-observer[tiab] OR intertechnician[tiab] OR inter-technician[tiab] OR intratechnician[tiab] OR intra-technician[tiab] OR interexaminer[tiab] OR inter-examiner[tiab] OR

		intraexaminer[tiab] OR intra-examiner[tiab] OR interassay[tiab] OR inter-assay[tiab] OR intraassay[tiab] OR intra-assay[tiab] OR interindividual[tiab] OR inter-individual[tiab] OR intraindividual[tiab] OR intra-individual[tiab] OR interparticipant[tiab] OR inter-participant[tiab] OR intraparticipant[tiab] OR intra-participant[tiab] OR kappa[tiab] OR kappa's[tiab] OR kappas[tiab] OR repeatab*[tw] OR ((replicab*[tw] OR repeated[tw]) AND (measure[tw] OR measures[tw] OR findings[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] OR "interval variability"[tiab] OR "rate variability"[tiab] OR (variability[tiab] AND (analysis[tiab] OR values[tiab])) OR (uncertainty[tiab] AND (measurement[tiab] OR measuring[tiab])) OR "standard error of measurement"[tiab] OR sensitiv*[tiab] OR responsive*[tiab] OR (limit[tiab] AND detection[tiab]) OR "minimal detectable concentration"[tiab] OR interpretab*[tiab] OR ((minimal[tiab] OR minimally[tiab] OR clinical[tiab] OR clinically[tiab]) AND (important[tiab] OR significant[tiab] OR detectable[tiab]) AND (change[tiab] OR difference[tiab])) OR (small*[tiab] AND (real[tiab] OR detectable[tiab]) AND (change[tiab] OR difference[tiab])) OR "meaningful change"[tiab] OR "ceiling effect"[tiab] OR "floor effect"[tiab] OR "Item response model"[tiab] OR IRT[tiab] OR Rasch[tiab] OR "Differential item functioning"[tiab] OR DIF[tiab] OR "computer adaptive testing"[tiab] OR "item bank"[tiab] OR "cross-cultural equivalence"[tiab])
6	#6	("addresses"[Publication Type] OR "biography"[Publication Type] OR "case reports"[Publication Type] OR "comment"[Publication Type] OR "directory"[Publication Type] OR "editorial"[Publication Type] OR "festschrift"[Publication Type] OR "interview"[Publication Type] OR "lectures"[Publication Type] OR "legal cases"[Publication Type] OR "legislation"[Publication Type] OR "letter"[Publication Type] OR "news"[Publication Type] OR "newspaper article"[Publication Type] OR "patient education handout"[Publication Type] OR "popular works"[Publication Type] OR "congresses"[Publication Type] OR "consensus development conference"[Publication Type] OR "consensus development conference, nih"[Publication Type] OR "practice guideline"[Publication Type]) NOT ("animals"[MeSH Terms] NOT "humans"[MeSH Terms])
7	#7	#1 AND #2 AND #3 AND #4 AND #5 NOT #6

Note: "*" to include all derivatives of that word or concept.

Table S2. Search strategy for Chinese databases

1	#1	TKA = 护理 OR TKA = 护士 OR TKA = 护生
2	#2	TKA = 健康教育能力 OR TKA = 患者教育能力 OR TKA = 健康教育胜任力 OR TKA = 患者教育胜任力
3	#3	SU = 评估 OR SU = 测量 OR SU = 评价 OR SU = 收集 OR SU = 调查 OR SU = 工具 OR SU = 问卷 OR SU = 量表 OR SU = 仪器 OR SU = 研究
4	#4	TKA = 信度 OR TKA = 效度 OR TKA = 反应度 OR TKA = 内部一致性 OR TKA = 稳定性 OR TKA = 相关系数 OR TKA = 克朗巴赫系数 OR TKA = 探索性因子分析 OR TKA = 验证性因子分析 OR TKA = 探索性因素分析 OR TKA = 验证性因素分析 OR TKA = 检验 OR TKA = 结果
5	#5	#1 AND #2 AND #3 AND #4

Note: TKA = title/abstract; SU = title/abstract/keywords.

Table S3. Characteristics of the included instruments

Instrument name	Developer(s)/ year developed	Construct (s)	Target population	Mode of administration	Recall period	(Sub)scale (s) (number of items)	Response options	Range of scores/scoring	Original language	Available translations

Table S4. Characteristics of the included study populations

		Population			Instrument administration			
Instrument	**Ref**	**N**	**Age Mean (SD, range) yr**	**Gender % female**	**Setting**	**Country**	**Language**	**Response rate**
A	1							
	2							
	3							
B	1							

Table S5. Rating the measurement properties of the instruments

Instrument	Study 1			Study 2			Study 3			OVERALL					
	RATING	RATING	RATING	RATING	RATING	RATING	RATING	RATING	RATING	OVERALL RATING	OVERALL RATING	OVERALL RATING	QUALITY OF EVIDENCE	QUALITY OF EVIDENCE	QUALITY OF EVIDENCE
	+ / - / ?	+ / - / ?	+ / - / ?	+ / - / ?	+ / - / ?	+ / - / ?	+ / - / ?	+ / - / ?	+ / - / ?	+ / - / ?	+ / - / ?	+ / - / ?	High, moderate, low, very low	High, moderate, low, very low	High, moderate, low, very low
	rater 1	rater 2	consensus	rater 1	rater 2	consensus	rater 1	rater 2	consensus	rater 1	rater 2	consensus	rater 1	rater 2	consensus
Structural validity															
internal consistency															
Cross-cultural validity															
Measurement invariance															
Reliability															
Measurement error															
Criterion validity															
Construct validity															
Responsiveness															

Note: “+” = sufficient; “-” = insufficient; “?” = indeterminate.

Table S5-1. Rating of the content validity of instruments

Instrument (Reference – study type/Rating of reviewers)	Content Validity													
	Relevance[1]						Comprehensiveness[1]		Comprehensibility[1]					CONTENT VALIDITY RATING[2]
	1. Are the included items relevant for the construct of interest?	2. Are the included items relevant for the target population of interest?	3. Are the included items relevant for the context of use of interest?	4. Are the response options appropriate?	5. Is the recall period appropriate?	RELEVANCE RATING[2]	6. Are all key concepts included?	COMPREHENSIVENESS RATING[2]	7. Are the PROM instructions understood by the population of interest as intended?	8. Are the PROM items and response options understood by the population of interest as intended?	9. Are the PROM items appropriately worded?	10. Do the response options match the question?	COMPREHENSIBILITY RATING[2]	
A (Ref 1- instrument development study)														
A (Ref 2 - Content validity study)														
A (Ref 3 - Content validity study)														
Rating of reviewers														
B (Ref 1- instrument development study)														
B (Ref 2 - Content validity study)														
Rating of reviewers														

……														

Note:1. Rating for the 10 criteria for relevance, comprehensiveness, comprehensibility can be + / - /±/ ?: ' + '= sufficient, ' - '= insufficient, '±' = inconsistent,'?' =indeterminate.

2. The RELEVANCE, COMPREHENSIVENESS, COMPREHESIBILITY, AND CONTENT VALIDITY rating can be + / - /±/ ? : ' + '= sufficient, ' - '= insufficient, '±' = inconsistent, '?' =indeterminate.

Table S6. Information on interpretability of instruments

Instrument (ref)		**Percentage of missing items and percentage of missing total scores**	**Floor and ceiling effects**	**Scores and change scores available for relevant (sub)groups**	**Minimal important change (MIC) or minimal important difference (MID)**	**Information on response shift**
Instrument A (ref 1)						
Instrument A (ref 2)						
Instrument A (ref 3)						
Instrument B (ref 1)						
……						

Table S7. Information on feasibility of instruments

Feasibility aspects	**Instrument A**	**Instrument B**	**Instrument C**	**Instrument D**
Patient's comprehensibility				
Clinician's comprehensibility				
Type and ease of administration				
Length of the instrument				
Completion time				
Patient's required mental and physical ability level				
Ease of standardization				
Ease of score calculation				
Copyright				
Cost of an instrument				
Required equipment				
Availability in different settings				
Regulatory agency's requirement for approval				

Table S8. Quality[1] of studies on measurement properties

Instrument	Content validity[2]					Structural validity	Internal consistency	Cross-cultural validity	Reliability	Measurement error	Criterion validity	Co
	Asking patients			Asking experts								
	Relevance	Comprehensiveness	Comprehensibility	Relevance	Comprehensiveness							Converg validit
A												
B												
……												

Note: 1. Quality: V = very good, A = adequate, D = doubtful, I = inadequate.
2. Given that the criteria and rating systems for evaluating the content validity of instruments are different from those for other measurement properties, the quality results of content validity are not included in this table but separately shown in following Table S8-1.

Table S8-1. Quality[1] of the instrument development

Instrument	PROM design							Cognitive interview (CI) study[3]				TOTAL PROM DEVELOPMENT
	General design requirements					Concept elicitation[2]	Total PROM design	General design requirements	Comprehensibility	Comprehensiveness	Total CI study	
	Clear construct	Clear origin of construct	Clear target population for which the PROM was developed	Clear context of use	PROM developed in sample representing the target population			CI study performed in sample representing the target population				
A												
B												
……												

Note: 1. Quality: V = very good, A = adequate, D = doubtful, I = inadequate.
2.The concept elicitation will not be further rated if the instrument(s) was not developed in the sample representing the target population.
3. Empty cells indicate that a CI study (or part of it) was not performed.

Table S9. Quality of the evidence for measurement properties of the instruments (Summary of findings)

Instrument	Content validity		Structural validity		Internal consistency		Cross-cultural validity		Reliability		Measurement error		Criterion validity		Hypotheses testing		Responsiveness	
	Overall Rating[1]	Quality of Evidence[3]	Overall Rating[2]	Quality of evidence[3]	Overall rating[2]	Quality of evidence[3]	Overall rating[2]	Quality of evidence[3]	Overall rating[2]	Quality of evidence[3]	Overall rating[2]	Quality of evidence[3]	Overall rating[2]	Quality of evidence[3]	Overall rating[2]	Quality of evidence[3]	Overall rating[2]	Quality of evidence[3]
Instrument A																		
Instrument B																		
Instrument C																		
……																		

Note:1. Overall ratings for the content validity (relevance, comprehensiveness, comprehensibility) can only be **+** / **-** /±: ‘ **+** ’= sufficient, ‘ **-** ’= insufficient, ‘±’ = inconsistent.

2. Overall ratings for other measurement properties can be **+** / **-** /±/ ?: ‘ **+** ’= sufficient, ‘ **-** ’= insufficient, ‘±’ = inconsistent, ‘?’ =indeterminate.

3. Ratings for quality of evidence: High, Moderate, Low, Very low.

Reporting checklist for systematic review (with or without a meta-analysis).

Based on the PRISMA guidelines.

Instructions to authors

Complete this checklist by entering the page numbers from your manuscript where readers will find each of the items listed below.

Your article may not currently address all the items on the checklist. Please modify your text to include the missing information. If you are certain that an item does not apply, please write "n/a" and provide a short explanation.

Upload your completed checklist as an extra file when you submit to a journal.

In your methods section, say that you used the PRISMAreporting guidelines, and cite them as:

Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, Shamseer L, Tetzlaff JM, Akl EA, Brennan SE, Chou R, Glanville J, Grimshaw JM, Hróbjartsson A, Lalu MM, Li T, Loder EW, Mayo-Wilson E, McDonald S, McGuinness LA, Stewart LA, Thomas J, Tricco AC, Welch VA, Whiting P, Moher D. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews

		Reporting Item	Page Number
Title			
Title	#1	Identify the report as a systematic review	1
Abstract			
Abstract	#2	Report an abstract addressing each item in the PRISMA 2020 for Abstracts checklist	1
Introduction			
Background/rationale	#3	Describe the rationale for the review in the context of existing knowledge	3
Objectives	#4	Provide an explicit statement of the objective(s) or question(s) the review addresses	5
Methods			

Eligibility criteria	#5	Specify the inclusion and exclusion criteria for the review and how studies were grouped for the syntheses	6
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	6
Search strategy	#7	Present the full search strategies for all databases, registers, and websites, including any filters and limits used	6
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	7
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	7
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 (for example, for all measures, time points, analyses), and, if not, the methods used to decide which results to collect	7
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	8
Effect measures	#12	Specify for each outcome the effect measure(s) (such as risk ratio, mean difference) used in the synthesis or presentation of results	N/A
Synthesis methods	#13a	Describe the processes used to decide which studies were eligible for each synthesis (such as tabulating the study intervention characteristics and comparing against the planned groups for each synthesis (item #5))	8
Synthesis methods	#13b	Describe any methods required to prepare the data for presentation or synthesis, such as handling of missing summary statistics or data	N/A

		conversions	
Synthesis methods	#13c	Describe any methods used to tabulate or visually display results of individual studies and syntheses	8
Synthesis methods	#13d	Describe any methods used to synthesise 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	9
Synthesis methods	#13e	Describe any methods used to explore possible causes of heterogeneity among study results (such as subgroup analysis, meta-regression)	N/A
Synthesis methods	#13f	Describe any sensitivity analyses conducted to assess robustness of the synthesised results	N/A
Reporting bias assessment	#14	Describe any methods used to assess risk of bias due to missing results in a synthesis (arising from reporting biases)	N/A
Certainty assessment	#15	Describe any methods used to assess certainty (or confidence) in the body of evidence for an outcome	9
Data items	#10b	List and define all other variables for which data were sought (such as participant and intervention characteristics, funding sources). Describe any assumptions made about any missing or unclear information	7
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 (http://www.prisma-statement.org/PRISMAStatement/FlowDiagram)	N/A
Study selection	#16b	Cite studies that might appear to meet the inclusion criteria, but which were excluded, and explain why they were excluded	N/A
Study characteristics	#17	Cite each included study and present its characteristics	N/A
Risk of bias in studies	#18	Present assessments of risk of bias for each included study	N/A
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	N/A

		precision (such as confidence/credible interval), ideally using structured tables or plots	
Results of syntheses	#20a	For each synthesis, briefly summarise the characteristics and risk of bias among contributing studies	N/A
Results of syntheses	#20b	Present results of all statistical syntheses conducted. If meta-analysis was done, present for each the summary estimate and its precision (such as confidence/credible interval) and measures of statistical heterogeneity. If comparing groups, describe the direction of the effect	N/A
Results of syntheses	#20c	Present results of all investigations of possible causes of heterogeneity among study results	N/A
Results of syntheses	#20d	Present results of all sensitivity analyses conducted to assess the robustness of the synthesised results	N/A
Risk of reporting biases in syntheses	#21	Present assessments of risk of bias due to missing results (arising from reporting biases) for each synthesis assessed	N/A
Certainty of evidence	#22	Present assessments of certainty (or confidence) in the body of evidence for each outcome assessed	N/A
Discussion			
Results in context	#23a	Provide a general interpretation of the results in the context of other evidence	N/A
Limitations of included studies	#23b	Discuss any limitations of the evidence included in the review	N/A
Limitations of the review methods	#23c	Discuss any limitations of the review processes used	N/A
Implications	#23d	Discuss implications of the results for practice, policy, and future research	10
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	6
Registration and protocol	#24b	Indicate where the review protocol can be accessed, or state that a protocol was not prepared	6

Registration and protocol	#24c	Describe and explain any amendments to information provided at registration or in the protocol	N/A
Support	#25	Describe sources of financial or non-financial support for the review, and the role of the funders or sponsors in the review	11
Competing interests	#26	Declare any competing interests of review authors	11
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	11

This checklist was completed on 16. February 2023 using https://www.goodreports.org/, a tool made by the EQUATOR Network in collaboration with Penelope.ai

BMJ Open

Instruments for measuring patient health education competence among nursing personnel: Protocol for a COSMIN-based systematic review

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2023-072905.R1
Article Type:	Protocol
Date Submitted by the Author:	28-Aug-2023
Complete List of Authors:	Wang, Shuyi; Central South University, Xiangya School of Nursing Liu, Ke; Central South University, Xiangya School of Nursing SHI, Zeya; Hunan Provincial People's Hospital, Chen, Qirong; Central South University, Xiangya School of Nursing; Xiangya Center for Evidence-Based Practice & Healthcare Innovation: A Joanna Briggs Institute Affiliated Group Tang, Siyuan; Central South University, Xiangya School of Nursing; XiangyaCenterforEvidence-BasedNursingPractice&HealthcareInnovation:AJBIAffiliatedGroup
Primary Subject Heading:	Nursing
Secondary Subject Heading:	Public health, Nursing
Keywords:	Health Education, PUBLIC HEALTH, Primary Health Care, Nurses

Instruments for measuring patient health education competence among nursing personnel: Protocol for a COSMIN-based systematic review

Shuyi Wang[a], Ke Liu[a], Zeya Shi[b], Qirong Chen[a,c,*], Siyuan Tang[a,c,*]

a. Xiangya School of Nursing, Central South University,

172 Tongzipo Road, Changsha, Hunan, 410013, China.

b. Nursing Department, Hunan Provincial People's Hospital,

61 Jiefang West Road, Changsha 410005, China

c. Xiangya Center for Evidence-Based Practice & Healthcare Innovation: A Joanna Briggs Institute Affiliated Group,

172 Tongzipo Road, Changsha, Hunan, 410013, China.

***Corresponding author:**

Qirong Chen

Address: 172 Tongzipo Road, Changsha, Hunan, 410013, China.

Phone number: (86)15084726637

Email address: qirong.chen@csu.edu.cn

Siyuan Tang

Address: 172 Tongzipo Road, Changsha, Hunan, 410013, China.

Email address: sytang263@csu.edu.cn

Phone number: (86)15874107355

Keywords: Health education; Competency assessment; Nursing personnel; Evaluation; Psychometrics

Word count: 3658

ABSTRACT

Introduction Health education, as a crucial strategic measure of disease prevention and

control in the 21st century, has become an important part of healthcare. As the main deliverers of patient health education, nursing personnel's patient health education competence (PHEC) has received much attention. Instruments for assessing the PHEC of nursing personnel have been developed internationally, but there is a lack of systematic reviews and evaluations of the psychometric properties of these instruments. To effectively select appropriate PHEC assessment instruments in specific contexts, a systematic and comprehensive review and evaluation of these measurement instruments are needed. The goal of this systematic review is to systematically evaluate the psychometric properties of existing PHEC instruments.

Methods and analysis In this study, eight databases will be searched between March 1, 2023, and March 31, 2023, to retrieve studies that include instrument(s) measuring the PHEC of nursing personnel. Two researchers will independently perform literature screening, data extraction, and literature evaluation. In case of disagreement, a third researcher will be involved in the resolution. The measurement properties of PHEC assessment instruments will be systematically reviewed based on the consensus-based standards for the selection of health measurement instruments (COMSIN) methodology and guideline.

Ethics and dissemination Ethical approval is not applicable for this study. We will share the findings from the study at national and/or international conferences and in a peer-reviewed journal in the fields of health education and/or patient education.

PROSPERO registration number CRD42023393293

Strengths and limitations of this study

- The Preferred Reporting Items for Systematic Reviews and Meta-analyses protocols (PRISMA-P) 2015 checklist and the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) 2020 checklist will be used to

guide the reporting of the protocol and systematic review, respectively.

- The consensus-based standards for the selection of health measurement instruments (COSMIN) methodology will be used to evaluate the methodological quality of included studies on measurement properties of the instruments and the quality of included instruments.
- The systematic review may fail to include relevant literature published outside of the searched databases.

1. INTRODUCTION

Health education has been identified by the World Health Organization (WHO) as one of the three crucial strategic measures of disease prevention and control in the 21st century, and it is the most economical and effective measure for improving public health.[1] Health education for patients can improve their understanding of their own health status and disease management measures, which can relieve patients' anxiety and improve their compliance and satisfaction with medical staff, thus improving their health status and quality of life.[2] These better patient outcomes could reduce the burden of disease on patients and society at the economic level.[3,4] As the world's largest group of health professionals and the health professionals who have the closest contact with patients, nursing staff plays an important role in patient health education.[3,5] Nurses often develop profound connections with their patients, rendering them optimal conveyors of health information and proponents of constructive behavioral transformations.[6] Their consistent and sustained patient interactions afford them an intimate grasp of individual needs, preferences, and hurdles, enabling the delivery of tailored patient health education that accommodates these divergent factors.[6,7] This education encompasses instructing patients on health preservation, preventive measures, and autonomous health management. Consequently, patients are

empowered to make enlightened choices and enhance compliance with treatment regimens. Functioning as integral healthcare team members, nurses proficiently facilitate intercommunication among patients, physicians, and allied healthcare professionals.[8] Their adeptness at translating medical jargon and disseminating information empowers patients to comprehend medical language, thereby expediting the formulation and execution of efficacious treatment strategies.[7] Therefore, nurses have an integral and important role in patient health education.

Patient health education competency (PHEC) refers to the specific qualities that health educators should have to conduct effective health education activities with patients.[9,10] PHEC is an essential professional competency for nursing staff and determines the quality of patient education.[11-14] However, in existing studies, the PHEC of clinical nurses is often the lowest-rated area of nursing competency.[15,16] Therefore, the development and strengthening of PHEC for nurses are extremely important to improve the quality of patient education, patient care, patient safety, and the development of nursing careers. In addition, we should pay attention to nursing students' PHEC because they are the primary reserve of the clinical nurse workforce.

Accurate measurement of PHEC is important because it can be used to assess the PHEC status of nursing personnel and to develop targeted strategies based on the nursing personnel's PHEC. Moreover, it can be used in research to assess the effectiveness of relevant PHEC interventions. Currently, relevant measurement instruments have been developed internationally: for example, a scale for measuring the PHEC of registered nurses developed by Lin et al. in 2017,[17] a PHEC competency assessment scale developed by Hwang et al. based on a literature review and the Delphi method,[18] and a Spanish version of the nurse PHEC scale developed by Pueyo-Garrigues et al.[19] Although related instruments are available for assessing PHEC in

nursing personnel, these evaluation instruments have been developed in different settings and their validation varies considerably, with none considered the gold standard.

In this study, we defined PHEC as the specific qualities that must be possessed by health educators to provide health education to patients, including knowledge, skills, beliefs or attitudes, self-concept, personality qualities, and motivation. Although there has been a review of PHEC measurement instruments for nursing staff, this review has some limitations on its rigor.[20] First, this review included not only measurement instruments for PHEC but also systems for evaluating PHEC, which are different from measurement instruments. Second, this review did not systematically evaluate the measurement properties of instruments for measuring PHEC based on related guidelines. However, a systematic and comprehensive review of PHEC measurement instruments is crucial for guiding the selection of instruments and/or guiding the development and refinement of high-quality instruments in the future. The consensus-based standards for the selection of health measurement instruments (COSMIN) methodology provides resources to systematically review measurement instruments and evaluate them in terms of both methodological quality and quality of measurement properties to select instruments that are of high quality for study purposes and provide an evidence-based foundation for future high-level instrument development.[21] Eskolin et al. conducted a review on instruments assessing nurses' competence in the empowerment of patient education.[22] However, in this review, the author did not give a clear and specific definition of ‘empowering patient education competence of nurses’. This may lead to an unclear research boundary. Their investigation encompassed not only instruments appraising nurses' PHEC but also instruments evaluating the quality of patient education provided by healthcare professionals. Furthermore, they included tools for measuring nurses' attitudes toward patient education. Considering the

importance of nursing personnel in patient health education, and to ensure a more distinct scope and targeted content, our study will focus specifically on the PHEC measurement instruments, which are designed specifically for nursing personnel, including both nurses and nursing students. Furthermore, in our review, we will incorporate Chinese databases, unveiling more qualified instruments that align with our stringent criteria. Thus, this study is designed to conduct a comprehensive and rigorous systematic review of PHEC assessment instruments based on the COSMIN methodology, to evaluate the measurement properties of these instruments, provide a reference for nursing personnel and researchers to accurately and effectively assess PHEC, and provide recommendations for researchers to develop and improve PHEC assessment instruments.

This systematic review will address the following questions: (1) What instruments are available for assessing the PHEC of nursing personnel? (2) What are the characteristics of these instruments? (3) What is the methodological quality of studies on the measurement properties of these instruments? (4) What are these instruments' measurement properties, interpretability, and feasibility? (5) What are the similarities and differences between these instruments? (6) What are the knowledge and research gaps in the assessment of PHEC of nursing personnel?

2. METHODS

The COSMIN guideline for systematic reviews of PROMs will be used to guide the implementation of the systematic review. PRISMA-P 2015 checklist and PRISMA 2020 checklist will be used to guide the reporting of the protocol and systematic review, respectively.[21,23,24] We registered the protocol in the International Prospective Register of Systematic Reviews (PROSPERO, CRD42023393293).

Inclusion and exclusion criteria of studies

Inclusion criteria

Studies will be included if they (1) address instrument(s) for measuring the PHEC of nurses or nursing students, (2) describe the processes of development and evaluation of one or more measurement properties for eligible instrument(s), (3) discuss instruments designed to measure the PHEC of health professionals (the literature explicitly mentions that it applies to nursing personnel as well), and (4) have full-text availability. If full-text versions of the studies are not available online, the authors of these articles will be contacted, and articles for which valid information was not available after contacting the authors will be excluded.

Exclusion criteria

Studies will be excluded if they are (1) not primary studies (e.g., biographies, addresses, and editorials) or are case studies, (2) reports that used the instruments only for outcome measurements, (3) secondary studies (e.g., reviews and/or systematic reviews), or (4) duplicate published studies.

Search strategy

A systematic search will be performed between March 1, 2023, and March 31, 2023, in six English databases (i.e., CINAHL, EMBASE, Ovid Medline, PubMed, PsycINFO, and Web of Science) and two Chinese databases (i.e., CNKI and WANFANG DATA). We include Chinese databases since the researchers speak Chinese as their native language. We will also search for and screen references of all eligible literature. The search time limit is from the library's creation date to the search date. A literature search will be conducted using a combination of subject terms and free words. The major search concepts will be nursing, health education, competence, instrument, and measurement properties. Related comprehensive and sensitive search strategies developed by other researchers will also be used in this literature search, including (1)

the search filter developed by the University of Oxford for finding PROMs,[25] (2) the sensitive PubMed search filter for measuring attributes developed by Terwee et al., and (3) corresponding search filters applicable to other databases.[26] We will examine results reported by nurses or nursing students, so the first filter will be adjusted appropriately (e.g., we will remove those sections that are relevant to the quality of life and patient-reported outcomes). The search strategy constructed for PubMed is described in Table S1 in the supplementary file. The search strategy for the Chinese databases is shown in Table S2 in the supplementary file.

Study screening

Covidence will be used to manage the references.[27] First, duplicates from the eight databases will be removed with Covidence. After the initial screening, both researchers will independently review and screen titles, abstracts, and full-text articles with the support of Covidence. In case of disagreement, a third researcher will be consulted to screen the literature. The screening processes of this study are shown in Figure 1.

Data extraction

The two researchers will independently extract data from the included papers and resolve their differences through discussion. We will extract the data about the characteristics of the instruments (including instrument name, developer[s]/year developed, construct[s], targeted population, mode of administration, recall period, [sub]scale[s]/[number of items], response options, range of scores/scoring, original language, and available translations; see Table S3 in the supplementary file), the characteristics of the included populations (including sample size, mean of age, gender, setting, country, and language; see Table S4 in the supplementary file,), the results on the psychometric properties (Table S5 in the supplementary file), and information about

the interpretability (Table S6 in the supplementary file) and feasibility (Table S7 in the supplementary file) of the included instruments.

The term 'outcome measure instrument development' will be used instead of the original 'patient-reported outcome measure development' to more accurately reflect the inclusion of studies that examined outcomes reported by nurses or nursing students rather than patients.

Quality appraisal and Data synthesis

Two researchers will independently assess the quality of eligible studies using the COSMIN Risk of Bias checklist, which is divided into three sections: content validity (instrument development and content validity), internal structure (structural validity, internal consistency, and cross-cultural validity/measurement invariance), and other measurement properties (reliability, measurement error, criterion validity, hypothesis testing for construct validity, and responsiveness).[21,23,28] Each measurement property will be evaluated by different items provided by the COSMIN Risk of Bias checklist, and the items will be rated on a five-level score of 'very good', 'adequate', 'doubtful', 'inadequate', or 'not applicable.'[23,28] Based on the 'the worst score counts' principle, each measurement property's overall methodological quality score is expressed by taking the lowest rating of any standard in the box.[23,29] Subsequently, the two researchers will apply the updated criteria for good measurement properties alone to evaluate the reliability and validity of the instruments themselves, and the quality of the evidence will be graded using the Grading of Recommendations Assessment, Development, and Evaluation (GRADE) approach.[23,29] In case of disagreement, a third researcher will be consulted.

We will work according to the following three steps. In the first step, two investigators will apply the COSMIN Risk of Bias checklist to evaluate the

methodological quality of each eligible study individually.[28] The final consensus on the results of the methodological quality will be presented in Tables S8 and S8-1 in the supplementary file. In the second step, the updated criteria for good measurement properties will be applied to evaluate the quality of evidence for each measured property, and the evaluation results will be shown in Tables S5 and S5-1 in the supplementary file.[23,29] This section mainly evaluates the strengths and weaknesses of the measurement properties. Among these, the quality of content validity will be evaluated according to the COSMIN methodology for content validity in three aspects: the relevance, comprehensiveness, and comprehensibility of items, which can be 'sufficient (+)', 'insufficient (-)', 'indeterminate (?)', or 'inconsistent (±)'.[29,30] The quality of the remaining measurement properties (structural validity, internal consistency, cross-cultural validity, measurement invariance, reliability, measurement error, criterion validity, construct validity, and responsiveness) will be evaluated by applying the COSMIN quality criteria, which can be 'sufficient (+)', 'insufficient (-)', and 'indeterminate (?)'.[23] The corresponding results will be reported in the rating columns of Table S5 in the supplementary file, and the results of rating content validity will be presented separately in Table S5-1 in the supplementary file. In the third step, a modified GRADE approach will be used to rate the quality of the above evidence, reflecting the level of confidence in the quality of the evidence. To evaluate the content's validity, three of these factors are applicable: risk of bias, inconsistency, and indirectness.[29] Assuming that the level of evidence quality for each of the remaining measurement properties is high, the quality of the evidence will be downgraded by considering the following factors: risk of bias, inconsistency, imprecision, and indirectness.[23] The quality of evidence will be divided into four levels: 'high', 'moderate', 'low', or 'very low'.[21,23] The corresponding results will be displayed in

Table S9 in the supplementary file. Two investigators will independently grade and cross-check the results. In case of disputes, final decisions will be made in consultation with the third investigator.

Patient and public involvement

Neither patients nor the public will be involved in this study.

Ethics and dissemination

Ethical approval is not applicable for this study. We will share the findings from the study at national and/or international conferences and in a peer-reviewed journal in the fields of health education and/or patient education.

3. DISCUSSION

To our knowledge, this will be the first COSMIN-based systematic review of PHEC assessment instruments for nursing personnel , which will be reported following the Preferred Reporting Items for Systematic Reviews and Meta-analyses protocols (PRISMA-P) 2020 checklist. This systematic review will provide a comprehensive rating of the level of evidence for each measurement property of the PHEC assessment instruments, which will be based on an evaluation of the measurement properties of all included instruments and the methodological quality of the studies. Through this study, we will be able to develop recommendations on the use of existing qualified instruments in clinical practice and research that could assist nursing personnel and researchers in the accurate and valid assessment of PHEC. This review may provide an evidence-based foundation for the development, design, validation, and use of future instruments by identifying problems in instrument development and validation and therefore help researchers to develop and improve these instruments.

Author contributions All authors have read and agreed to the published version of the

manuscript.

Conceptualization: QC, ST, SW;

Methodology: QC, ST, ZS;

Data curation: QC, SW, KL;

Writing—original draft preparation: SW, KL, QC;

Writing—review and editing: QC, ST, ZS;

Supervision: QC and ST.

Funding This work was supported by the National Natural Science Foundation of China (No. 72104250) and the Natural Science Foundation of Hunan Province (No. 2022JJ40642).

Competing interests None declared.

Patient consent Not required.

Data sharing statement No additional data available.

List of Figures:

List of Supplementary File Tables:

Table S8: Quality of studies on measurement properties

Table S8-1: Quality of the instrument development

Table S9: Quality of the evidence for measurement properties of the instruments

References:

1. Organization WH. Health education: theoretical concepts, effective strategies and core competencies: a foundation document to guide capacity development of health educators. 2012. Available from: https://apps.who.int/iris/bitstream/handle/10665/119953/EMRPUB_2012_EN_1362.pdf?sequen
2. Turner A, Anderson JK, Wallace LM, et al. An evaluation of a self-management program for patients with long-term conditions. *Patient Educ Couns* 2015;98(2):213-9. doi: 10.1016/j.pec.2014.08.022.
3. Lin P, Gao C, Su S, et al. Teaching principles: The applications in nursing practice. *Taipei: Farseeing.* 2005.
4. Siegel KR, Ali MK, Zhou X, et al. Cost-effectiveness of interventions to manage diabetes: has the evidence changed since 2008? *Diabetes Care* 2020;43(7):1557-92. doi: 10.2337/dci20-0017.
5. Koutsopoulou S, Papathanassoglou ED, Katapodi MC, et al. Nurses as information providers to cancer patients: Critical review of the evidence. *J Clin Nurs* 2010;19(5-6):749-65. doi: 10.1111/j.1365-2702.2009.02954.x.
6. Feo R, Rasmussen P, Wiechula R, et al. Developing effective and caring nurse-patient relationships. *Nurs Stand* 2017;31(28):54-63. doi: 10.7748/ns.2017.e10735
7. Kwame A, Petrucka PM. A literature-based study of patient-centered care and communication in nurse-patient interactions: barriers, facilitators, and the way

forward. *BMC Nurs* 2021;20(1):158. doi: 10.1186/s12912-021-00684-2

8. Ghiyasvandian S, Zakerimoghadam M, Peyravi H. Nurse as a facilitator to professional communication: a qualitative study. *Glob J Health Sci* 2014;7(2):294-303. doi: 10.5539/gjhs.v7n2p294
9. Hwang H-L, Kuo T-Y. Competency in delivering health education: A concept analysis. *J Interprof Educ Pract* 2018;11:20-5. doi: https://doi.org/10.1016/j.xjep.2018.02.005
10. Karimi Moonaghi H, Emami Zeydi A, Mirhaghi A. Patient education among nurses: bringing evidence into clinical applicability in Iran. *Invest Educ Enferm* 2016;34(1):137-51. doi: 10.17533/udea.iee.v34n1a16.
11. Flinkman M, Leino-Kilpi H, Numminen O, et al. Nurse Competence Scale: a systematic and psychometric review. *J Adv Nurs* 2017;73(5):1035-50. doi: 10.1111/jan.13183.
12. Ko Y, Yu S. Core nursing competency assessment tool for graduates of outcome-based nursing education in South Korea: A validation study. *Jpn J Nurs Sci* 2019;16(2):155-71. doi: 10.1111/jjns.12223.
13. Nilsson J, Engström M, Florin J, et al. A short version of the nurse professional competence scale for measuring nurses' self-reported competence. *Nurse Educ Today* 2018;71:233-9. doi: 10.1016/j.nedt.2018.09.028.
14. Wilkinson CA. Competency assessment tools for registered nurses: An integrative review. *J Contin Educ Nurs* 2013;44(1):31-7. doi: 10.3928/00220124-20121101-53.
15. Cowan DT, Wilson-Barnett DJ, Norman IJ, et al. Measuring nursing competence: Development of a self-assessment tool for general nurses across Europe. *Int J Nurs Stud* 2008;45(6):902-13. doi: 10.1016/j.ijnurstu.2007.03.004.

16. Pueyo-Garrigues M, Whitehead D, Pardavila-Belio MI, et al. Health education: a Rogerian concept analysis. *Int J Nurs Stud* 2019;94:131-8. doi: 10.1016/j.ijnurstu.2019.03.005.
17. Lin LY, Wang RH. Patient Education Competence Scale for Registered Nurses in Taiwan: Scale development and psychometric validation. *Jpn J Nurs Sci* 2017;14(2):117-25. doi: 10.1111/jjns.12141.
18. Hwang HL, Kuo ML, Tu CT. Health education and competency scale: Development and testing. *J Clin Nurs* 2018;27(3-4):e658-e67. doi: 10.1111/jocn.14116.
19. Pueyo-Garrigues M, Pardavila-Belio MI, Whitehead D, et al. Nurses' knowledge, skills and personal attributes for competent health education practice: An instrument development and psychometric validation study. *J Adv Nurs* 2021;77(2):715-28. doi: 10.1111/jan.14632.
20. Mi Y, Wu D, Wei ZZ, et al. Research progress on evaluation tool for health education competency of nursing staff. *Chinese Nursing Research* 2020;34(11):1983-7.
21. Prinsen CA, Mokkink LB, Bouter LM, et al. COSMIN guideline for systematic reviews of patient-reported outcome measures. *Qual Life Res* 2018;27:1147-57. doi: 10.1007/s11136-018-1798-3.
22. Eskolin SE, Inkeroinen S, Leino-Kilpi H, et al. Instruments for measuring empowering patient education competence of nurses: Systematic review. *J Adv Nurs* 2023;79(7):2414-28. doi: 10.1111/jan.15597
23. Mokkink LB, Prinsen C, Patrick DL, et al. COSMIN methodology for systematic reviews of patient-reported outcome measures (PROMs). User manual. 2018;78(1):6-63. Available from: https://www.cosmin.nl/wp-

content/uploads/COSMIN-syst-review-for-PROMs-manual_version-1_feb-2018-1.pdf

24. Page MJ, Moher D, Bossuyt PM, et al. PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. *BMJ* 2021;372. doi: 10.1136/bmj.n160.
25. Mackintosh A, Comabella C, Hadi M, et al. PROM group construct & instrument type filters. *Department of Public Health, University of Oxford, Oxford.* 2010. Available from: https://cosmin.nl/wp-content/uploads/prom-search-filter-oxford-2010.pdf
26. Terwee CB, Jansma EP, Riphagen II, et al. Development of a methodological PubMed search filter for finding studies on measurement properties of measurement instruments. *Qual Life Res* 2009;18:1115-23. doi: 10.1007/s11136-009-9528-5.
27. Babineau J. Product review: Covidence (systematic review software). *Journal of the Canadian Health Libraries Association/Journal de l'Association des bibliothèques de la santé du Canada* 2014;35(2):68-71.
28. Mokkink LB, De Vet HC, Prinsen CA, et al. COSMIN risk of bias checklist for systematic reviews of patient-reported outcome measures. *Qual Life Res* 2018;27:1171-9. doi: 10.1007/s11136-017-1765-4.
29. Terwee CB, Prinsen C, Chiarotto A, et al. COSMIN methodology for assessing the content validity of PROMs–user manual. *Amsterdam: VU University Medical Center.* 2018. Available from: https://cosmin.nl/wp-content/uploads/COSMIN-methodology-for-content-validity-user-manual-v1.pdf
30. Terwee CB, Prinsen CA, Chiarotto A, et al. COSMIN methodology for evaluating

1 the content validity of patient-reported outcome measures: a Delphi study. *Qual*
2 *Life Res* 2018;27:1159-70. doi: 10.1007/s11136-018-1829-0.

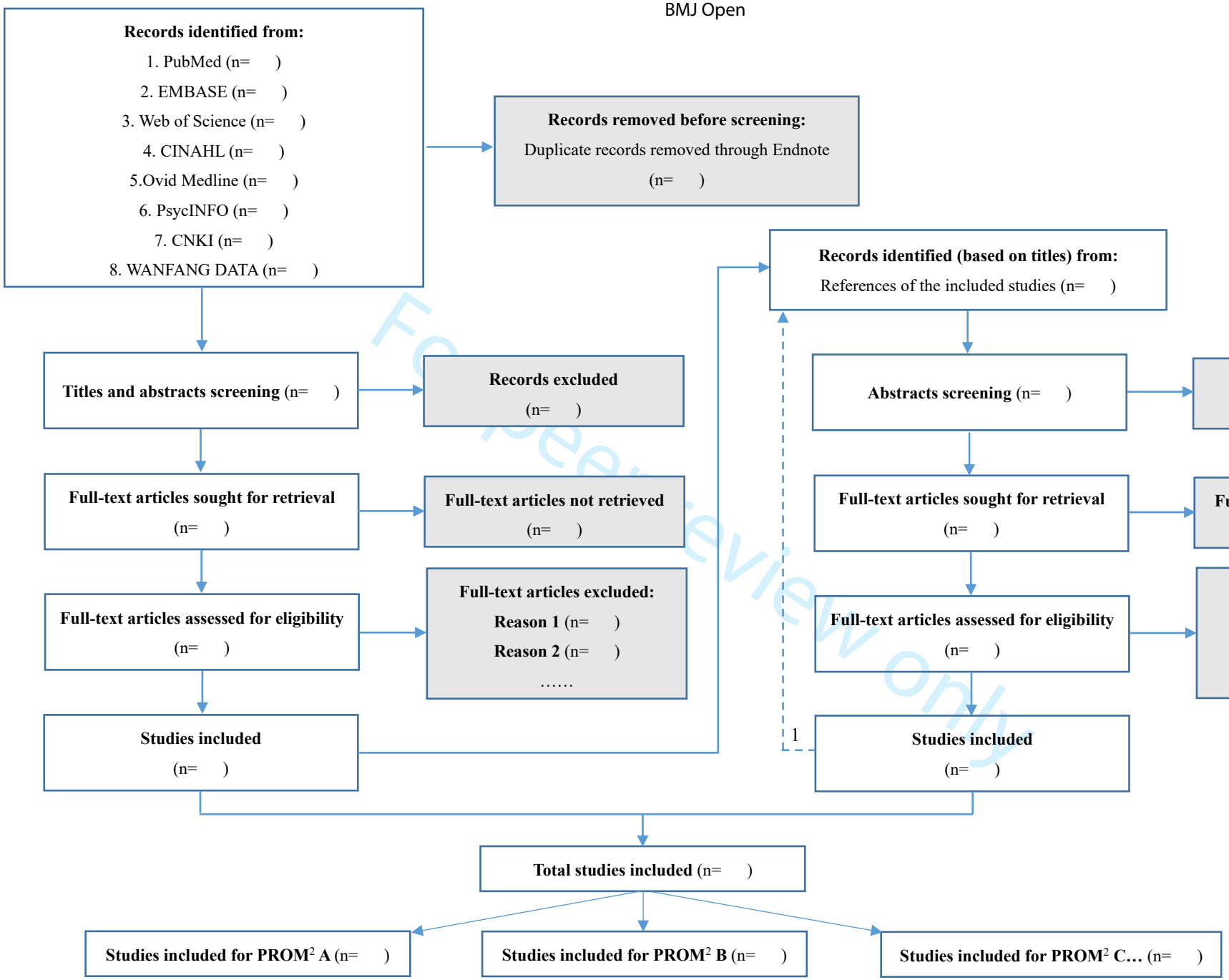


Figure 1. Flowchart of literature selection process

Note: 1. References of all included studies will be manually screened until no eligible studies can be identified.

2. PROM: patient-reported outcome measure.

Table S1. Search strategy for PubMed

1	#1	nurs*
2	#2	((((((((((Health Education[Mesh] OR Patient Education as Topic[Mesh]) OR (health education*[tiab])) OR (education, health[tiab])) OR (patient education*[tiab])) OR (education program*[tiab])) OR (educational program*[tiab])) OR (education, patient*[tiab])) OR (education of patient*[tiab])) OR (nursing education*[tiab])) OR (hospital education*[tiab])) OR (educational activit*[tiab])) OR ((educat*[tiab]) AND (individual patient*[tiab]))
3	#3	(((((((professional competence[MeSH]) OR (professional competence[tiab])) OR (competence, professional[tiab])) OR (generalization of expertise[tiab])) OR (expertise generalization[tiab])) OR (technical expertise[tiab])) OR (expertise, technical[tiab])) OR (competenc*[tiab] OR capabilit*[tiab] OR capacit*[tiab] OR abilit*[tiab])) OR ((skill*[tiab] OR belief[tiab] OR attitude[tiab]) AND (knowledge[tiab]))
4	#4	(report[tiab] OR reported[tiab] OR reporting[tiab] OR rated[tiab] OR rating[tiab] OR ratings[tiab] OR based[tiab] OR assessed[tiab] OR assessment[tiab] OR assessments[tiab] OR disability[tiab] OR function[tiab] OR functional[tiab] OR functions[tiab] OR subjective[tiab] OR utility[tiab] OR utilities[tiab] OR wellbeing[tiab] OR well being[tiab]) AND (index[tiab] OR indices[tiab] OR instrument[tiab] OR instruments[tiab] OR measure[tiab] OR measures[tiab] OR questionnaire[tiab] OR questionnaires[tiab] OR profile[tiab] OR profiles[tiab] OR scale[tiab] OR scales[tiab] OR score[tiab] OR scores[tiab] OR status[tiab] OR survey[tiab] OR surveys[tiab])
5	#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 reproducib*[tiab] OR "discriminant analysis"[MeSH] OR reliab*[tiab] OR unreliab*[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 alphas[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 intertester[tiab] OR inter-tester[tiab] OR intratester[tiab] OR intra-tester[tiab] OR interobserver[tiab] OR inter-observer[tiab] OR intraobserver[tiab] OR intra-observer[tiab] OR intertechnician[tiab] OR inter-technician[tiab] OR intratechnician[tiab] OR intra-technician[tiab] OR interexaminer[tiab] OR inter-examiner[tiab] OR intraexaminer[tiab] OR intra-examiner[tiab] OR interassay[tiab] OR inter-

		assay[tiab] OR intraassay[tiab] OR intra-assay[tiab] OR interindividual[tiab] OR inter-individual[tiab] OR intraindividual[tiab] OR intra-individual[tiab] OR interparticipant[tiab] OR inter-participant[tiab] OR intraparticipant[tiab] OR intra-participant[tiab] OR kappa[tiab] OR kappa's[tiab] OR kappas[tiab] OR repeatab*[tw] OR ((replicab*[tw] OR repeated[tw]) AND (measure[tw] OR measures[tw] OR findings[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] OR "interval variability"[tiab] OR "rate variability"[tiab] OR (variability[tiab] AND (analysis[tiab] OR values[tiab])) OR (uncertainty[tiab] AND (measurement[tiab] OR measuring[tiab])) OR "standard error of measurement"[tiab] OR sensitiv*[tiab] OR responsive*[tiab] OR (limit[tiab] AND detection[tiab]) OR "minimal detectable concentration"[tiab] OR interpretab*[tiab] OR ((minimal[tiab] OR minimally[tiab] OR clinical[tiab] OR clinically[tiab]) AND (important[tiab] OR significant[tiab] OR detectable[tiab]) AND (change[tiab] OR difference[tiab])) OR (small*[tiab] AND (real[tiab] OR detectable[tiab]) AND (change[tiab] OR difference[tiab])) OR "meaningful change"[tiab] OR "ceiling effect"[tiab] OR "floor effect"[tiab] OR "Item response model"[tiab] OR IRT[tiab] OR Rasch[tiab] OR "Differential item functioning"[tiab] OR DIF[tiab] OR "computer adaptive testing"[tiab] OR "item bank"[tiab] OR "cross-cultural equivalence"[tiab])
6	#6	("addresses"[Publication Type] OR "biography"[Publication Type] OR "case reports"[Publication Type] OR "comment"[Publication Type] OR "directory"[Publication Type] OR "editorial"[Publication Type] OR "festschrift"[Publication Type] OR "interview"[Publication Type] OR "lectures"[Publication Type] OR "legal cases"[Publication Type] OR "legislation"[Publication Type] OR "letter"[Publication Type] OR "news"[Publication Type] OR "newspaper article"[Publication Type] OR "patient education handout"[Publication Type] OR "popular works"[Publication Type] OR "congresses"[Publication Type] OR "consensus development conference"[Publication Type] OR "consensus development conference, nih"[Publication Type] OR "practice guideline"[Publication Type]) NOT ("animals"[MeSH Terms] NOT "humans"[MeSH Terms])
7	#7	#1 AND #2 AND #3 AND #4 AND #5 NOT #6

Note: "*" to include all derivatives of that word or concept.

Table S2. Search strategy for Chinese databases

1	#1	TKA = 护理 OR TKA = 护士 OR TKA = 护生
2	#2	TKA = 健康教育能力 OR TKA = 患者教育能力 OR TKA = 健康教育胜任力 OR TKA = 患者教育胜任力
3	#3	SU = 评估 OR SU = 测量 OR SU = 评价 OR SU = 收集 OR SU = 调查 OR SU = 工具 OR SU = 问卷 OR SU = 量表 OR SU = 仪器 OR SU = 研究
4	#4	TKA = 信度 OR TKA = 效度 OR TKA = 反应度 OR TKA = 内部一致性 OR TKA = 稳定性 OR TKA = 相关系数 OR TKA = 克朗巴赫系数 OR TKA = 探索性因子分析 OR TKA = 验证性因子分析 OR TKA = 探索性因素分析 OR TKA = 验证性因素分析 OR TKA = 检验 OR TKA = 结果
5	#5	#1 AND #2 AND #3 AND #4

Note: TKA = title/abstract; SU = title/abstract/keywords.

Table S3. Characteristics of the included instruments

Instrument name	Developer(s)/ year developed	Construct (s)	Target population	Mode of administration	Recall period	(Sub)scale (s) (number of items)	Response options	Range of scores/scoring	Original language	Available translations

Table S4. Characteristics of the included study populations

		Population			Instrument administration			
Instrument	**Ref**	**N**	**Age Mean (SD, range) yr**	**Gender % female**	**Setting**	**Country**	**Language**	**Response rate**
A	1							
	2							
	3							
B	1							

Table S5. Rating the measurement properties of the instruments

Instrument	Study 1			Study 2			Study 3			OVERALL					
	RATING	RATING	RATING	RATING	RATING	RATING	RATING	RATING	RATING	OVERALL RATING	OVERALL RATING	OVERALL RATING	QUALITY OF EVIDENCE	QUALITY OF EVIDENCE	QUALITY OF EVIDENCE
	+ / - / ?	+ / - / ?	+ / - / ?	+ / - / ?	+ / - / ?	+ / - / ?	+ / - / ?	+ / - / ?	+ / - / ?	+ / - / ?	+ / - / ?	+ / - / ?	High, moderate, low, very low	High, moderate, low, very low	High, moderate, low, very low
	rater 1	rater 2	consensus	rater 1	rater 2	consensus	rater 1	rater 2	consensus	rater 1	rater 2	consensus	rater 1	rater 2	consensus
Structural validity															
internal consistency															
Cross-cultural validity															
Measurement invariance															
Reliability															
Measurement error															
Criterion validity															
Construct validity															
Responsiveness															

Note: “+” = sufficient; “-” = insufficient; “?” = indeterminate.

Table S5-1. Rating of the content validity of instruments

Instrument (Reference – study type/Rating of reviewers)	Content Validity												
	Relevance[1]						Comprehensiveness[1]		Comprehensibility[1]				CONTENT VALIDITY RATING[2]
	1. Are the included items relevant for the construct of interest?	2. Are the included items relevant for the target population of interest?	3. Are the included items relevant for the context of use of interest?	4. Are the response options appropriate?	5. Is the recall period appropriate?	RELEVANCE RATING[2]	6. Are all key concepts included?	COMPREHENSIVENESS RATING[2]	7. Are the PROM instructions understood by the population of interest as intended?	8. Are the PROM items and response options understood by the population of interest as intended?	9. Are the PROM items appropriately worded?	10. Do the response options match the question?	COMPREHENSIBILITY RATING[2]	
A (Ref 1- instrument development study)														
A (Ref 2 - Content validity study)														
A (Ref 3 - Content validity study)														
Rating of reviewers														
B (Ref 1- instrument development study)														
B (Ref 2 - Content validity study)														
Rating of reviewers														

……														

Note:1. Rating for the 10 criteria for relevance, comprehensiveness, comprehensibility can be + / − /±/ ?: ‘ + ’= sufficient, ‘ − ’= insufficient, ‘±’ = inconsistent,‘?’ =indeterminate.

2. The RELEVANCE, COMPREHENSIVENESS, COMPREHESIBILITY, AND CONTENT VALIDITY rating can be + / − /±/ ? : ‘ + ’= sufficient, ‘ − ’= insufficient, ‘±’ = inconsistent, ‘?’ =indeterminate.

Table S6. Information on interpretability of instruments

Instrument (ref)		**Percentage of missing items and percentage of missing total scores**	**Floor and ceiling effects**	**Scores and change scores available for relevant (sub)groups**	**Minimal important change (MIC) or minimal important difference (MID)**	**Information on response shift**
Instrument A (ref 1)						
Instrument A (ref 2)						
Instrument A (ref 3)						
Instrument B (ref 1)						
……						

Table S7. Information on feasibility of instruments

Feasibility aspects	Instrument A	Instrument B	Instrument C	Instrument D
Patient's comprehensibility				
Clinician's comprehensibility				
Type and ease of administration				
Length of the instrument				
Completion time				
Patient's required mental and physical ability level				
Ease of standardization				
Ease of score calculation				
Copyright				
Cost of an instrument				
Required equipment				
Availability in different settings				
Regulatory agency's requirement for approval				

Table S8. Quality[1] of studies on measurement properties

Instrument	Content validity[2]					Structural validity	Internal consistency	Cross-cultural validity	Reliability	Measurement error	Criterion validity	Co
	Asking patients			Asking experts								
	Relevance	Comprehensiveness	Comprehensibility	Relevance	Comprehensiveness							Converg validit
A												
B												
……												

Note: 1. Quality: V = very good, A = adequate, D = doubtful, I = inadequate.
2. Given that the criteria and rating systems for evaluating the content validity of instruments are different from those for other measurement properties, the quality results of content validity are not included in this table but separately shown in following Table S8-1.

Table S8-1. Quality[1] of the instrument development

Instrument	PROM design							Cognitive interview (CI) study[3]				TOTAL PROM DEVELOPMENT
	General design requirements					Concept elicitation[2]	Total PROM design	General design requirements	Comprehensibility	Comprehensiveness	Total CI study	
	Clear construct	Clear origin of construct	Clear target population for which the PROM was developed	Clear context of use	PROM developed in sample representing the target population			CI study performed in sample representing the target population				
A												
B												
……												

Note: 1. Quality: V = very good, A = adequate, D = doubtful, I = inadequate.
2.The concept elicitation will not be further rated if the instrument(s) was not developed in the sample representing the target population.
3. Empty cells indicate that a CI study (or part of it) was not performed.

Table S9. Quality of the evidence for measurement properties of the instruments (Summary of findings)

Instrument	Content validity		Structural validity		Internal consistency		Cross-cultural validity		Reliability		Measurement error		Criterion validity		Hypotheses testing		Responsiveness	
	Overall Rating[1]	Quality of Evidence[3]	Overall Rating[2]	Quality of evidence[3]	Overall rating[2]	Quality of evidence[3]	Overall rating[2]	Quality of evidence[3]	Overall rating[2]	Quality of evidence[3]	Overall rating[2]	Quality of evidence[3]	Overall rating[2]	Quality of evidence[3]	Overall rating[2]	Quality of evidence[3]	Overall rating[2]	Quality of evidence[3]
Instrument A																		
Instrument B																		
Instrument C																		
……																		

Note:1. Overall ratings for the content validity (relevance, comprehensiveness, comprehensibility) can only be + / − /±: ‘ + ’= sufficient, ‘ − ’= insufficient, ‘±’ = inconsistent.

2. Overall ratings for other measurement properties can be + / − /±/ ?: ‘ + ’= sufficient, ‘ − ’= insufficient, ‘±’ = inconsistent, ‘?’ =indeterminate.

3. Ratings for quality of evidence: High, Moderate, Low, Very low.

PRISMA-P (Preferred Reporting Items for Systematic review and Meta-Analysis Protocols) 2015 checklist: recommended items to address in a systematic review protocol*

Section and topic	Item No	Checklist item	
ADMINISTRATIVE INFORMATION			**Pg.**
Title:			
Identification	1a	Identify the report as a protocol of a systematic review	Pg. 1
Update	1b	If the protocol is for an update of a previous systematic review, identify as such	N/A
Registration	2	If registered, provide the name of the registry (such as PROSPERO) and registration number	Pg. 2 and Pg. 6
Authors:			
Contact	3a	Provide name, institutional affiliation, e-mail address of all protocol authors; provide physical mailing address of corresponding author	Pg. 1
Contributions	3b	Describe contributions of protocol authors and identify the guarantor of the review	Pg. 11-12
Amendments	4	If the protocol represents an amendment of a previously completed or published protocol, identify as such and list changes; otherwise, state plan for documenting important protocol amendments	N/A
Support:			
Sources	5a	Indicate sources of financial or other support for the review	Pg. 12
Sponsor	5b	Provide name for the review funder and/or sponsor	N/A
Role of sponsor or funder	5c	Describe roles of funder(s), sponsor(s), and/or institution(s), if any, in developing the protocol	N/A
INTRODUCTION			
Rationale	6	Describe the rationale for the review in the context of what is already known	Pg. 3-6
Objectives	7	Provide an explicit statement of the question(s) the review will address with reference to participants, interventions, comparators, and outcomes (PICO)	Pg. 6
METHODS			
Eligibility criteria	8	Specify the study characteristics (such as PICO, study design, setting, time frame) and report characteristics (such as years considered, language, publication status) to be used as criteria for eligibility for the review	Pg. 7
Information sources	9	Describe all intended information sources (such as electronic databases, contact with study authors, trial registers or other grey literature sources) with planned dates of coverage	Pg. 7
Search strategy	10	Present draft of search strategy to be used for at least one electronic database, including planned limits, such that it could be repeated	Supplemental file
Study records:			
Data management	11a	Describe the mechanism(s) that will be used to manage records and data throughout the review	Pg. 8

Selection process	11b	State the process that will be used for selecting studies (such as two independent reviewers) through each phase of the review (that is, screening, eligibility and inclusion in meta-analysis)	Pg. 8
Data collection process	11c	Describe planned method of extracting data from reports (such as piloting forms, done independently, in duplicate), any processes for obtaining and confirming data from investigators	Pg. 8-9 Supplemental file
Data items	12	List and define all variables for which data will be sought (such as PICO items, funding sources), any pre-planned data assumptions and simplifications	Pg. 8-9 Supplemental file
Outcomes and prioritization	13	List and define all outcomes for which data will be sought, including prioritization of main and additional outcomes, with rationale	Pg. 8-9 Supplemental file
Risk of bias in individual studies	14	Describe anticipated methods for assessing risk of bias of individual studies, including whether this will be done at the outcome or study level, or both; state how this information will be used in data synthesis	Pg. 9
Data synthesis	15a	Describe criteria under which study data will be quantitatively synthesised	N/A
	15b	If data are appropriate for quantitative synthesis, describe planned summary measures, methods of handling data and methods of combining data from studies, including any planned exploration of consistency (such as I^2, Kendall's τ)	N/A
	15c	Describe any proposed additional analyses (such as sensitivity or subgroup analyses, meta-regression)	N/A
	15d	If quantitative synthesis is not appropriate, describe the type of summary planned	Pg. 10
Meta-bias(es)	16	Specify any planned assessment of meta-bias(es) (such as publication bias across studies, selective reporting within studies)	N/A
Confidence in cumulative evidence	17	Describe how the strength of the body of evidence will be assessed (such as GRADE)	Pg. 9-10

*** It is strongly recommended that this checklist be read in conjunction with the PRISMA-P Explanation and Elaboration (cite when available) for important clarification on the items. Amendments to a review protocol should be tracked and dated. The copyright for PRISMA-P (including checklist) is held by the PRISMA-P Group and is distributed under a Creative Commons Attribution Licence 4.0.**

From: Shamseer L, Moher D, Clarke M, Ghersi D, Liberati A, Petticrew M, Shekelle P, Stewart L, PRISMA-P Group. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. BMJ. 2015 Jan 2;349(jan02 1):g7647.

BMJ Open

Instruments for measuring patient health education competence among nursing personnel: Protocol for a COSMIN-based systematic review

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2023-072905.R2
Article Type:	Protocol
Date Submitted by the Author:	11-Sep-2023
Complete List of Authors:	Wang, Shuyi; Central South University, Xiangya School of Nursing Liu, Ke; Central South University, Xiangya School of Nursing SHI, Zeya; Hunan Provincial People's Hospital, Chen, Qirong; Central South University, Xiangya School of Nursing; Xiangya Center for Evidence-Based Practice & Healthcare Innovation: A Joanna Briggs Institute Affiliated Group Tang, Siyuan; Central South University, Xiangya School of Nursing; XiangyaCenterforEvidence-BasedNursingPractice&HealthcareInnovation:AJBIAffiliatedGroup
Primary Subject Heading:	Nursing
Secondary Subject Heading:	Public health, Nursing
Keywords:	Health Education, PUBLIC HEALTH, Primary Health Care, Nurses

Instruments for measuring patient health education competence among nursing personnel: Protocol for a COSMIN-based systematic review

Shuyi Wang[a], Ke Liu[a], Zeya Shi[b], Qirong Chen[a,c,*], Siyuan Tang[a,c,*]

a. Xiangya School of Nursing, Central South University,

172 Tongzipo Road, Changsha, Hunan, 410013, China.

b. Nursing Department, Hunan Provincial People's Hospital,

61 Jiefang West Road, Changsha 410005, China

c. Xiangya Center for Evidence-Based Practice & Healthcare Innovation: A Joanna Briggs Institute Affiliated Group,

172 Tongzipo Road, Changsha, Hunan, 410013, China.

***Corresponding author:**

Qirong Chen

Address: 172 Tongzipo Road, Changsha, Hunan, 410013, China.

Phone number: (86)15084726637

Email address: qirong.chen@csu.edu.cn

Siyuan Tang

Address: 172 Tongzipo Road, Changsha, Hunan, 410013, China.

Email address: sytang263@csu.edu.cn

Phone number: (86)15874107355

Keywords: Health education; Competency assessment; Nursing personnel; Evaluation; Psychometrics

Word count: 3658

ABSTRACT

Introduction Health education, as a crucial strategic measure of disease prevention and

control in the 21st century, has become an important part of health care. As the main deliverers of patient health education, nursing personnel's patient health education competence (PHEC) has received much attention. Instruments for assessing the PHEC of nursing personnel have been developed internationally, but there is a lack of systematic reviews and evaluations of the psychometric properties of these instruments. To effectively select appropriate PHEC assessment instruments in specific contexts, a systematic and comprehensive review and evaluation of these measurement instruments are needed. The goal of this systematic review is to systematically evaluate the psychometric properties of existing PHEC instruments.

Methods and analysis In this study, eight databases will be searched between March 1, 2023, and March 31, 2023, to retrieve studies that include instrument(s) measuring the PHEC of nursing personnel. Two researchers will independently perform literature screening, data extraction, and literature evaluation. In case of disagreement, a third researcher will be involved in the resolution. The measurement properties of PHEC assessment instruments will be systematically reviewed based on the consensus-based standards for the selection of health measurement instruments (COMSIN) methodology and guideline.

Ethics and dissemination Ethical approval is not applicable for this study. We will share the findings from the study at national and/or international conferences and in a peer-reviewed journal in the fields of health education and/or patient education.

PROSPERO registration number CRD42023393293

Strengths and limitations of this study

- The Preferred Reporting Items for Systematic Reviews and Meta-analyses protocols (PRISMA-P) 2015 checklist and the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) 2020 checklist will be used to

guide the reporting of the protocol and systematic review, respectively.

- The consensus-based standards for the selection of health measurement instruments (COSMIN) methodology will be used to evaluate the methodological quality of included studies on measurement properties of the instruments and the quality of included instruments.
- The systematic review may fail to include relevant literature published outside of the searched databases.

1. INTRODUCTION

Health education has been identified by the World Health Organization (WHO) as one of the three crucial strategic measures of disease prevention and control in the 21st century, and it is the most economical and effective measure for improving public health.[1] Health education for patients can improve their understanding of their own health status and disease management measures, which can relieve patients' anxiety and improve their compliance and satisfaction with medical staff, thus improving their health status and quality of life.[2] These better patient outcomes could reduce the burden of disease on patients and society at the economic level.[3,4] As the world's largest group of health professionals and the health professionals who have the closest contact with patients, nursing staff plays an important role in patient health education.[3,5] Nurses often develop profound connections with their patients, rendering them optimal conveyors of health information and proponents of constructive behavioral transformations.[6] Their consistent and sustained patient interactions afford them an intimate grasp of individual needs, preferences, and hurdles, enabling the delivery of tailored patient health education that accommodates these divergent factors.[6,7] This education encompasses instructing patients on health preservation, preventive measures, and autonomous health management. Consequently, patients are

empowered to make enlightened choices and enhance compliance with treatment regimens. Functioning as integral health care team members, nurses proficiently facilitate intercommunication among patients, physicians, and allied health care professionals.[8] Their adeptness at translating medical jargon and disseminating information empowers patients to comprehend medical language, thereby expediting the formulation and execution of efficacious treatment strategies.[7] Therefore, nurses have an integral and important role in patient health education.

Patient health education competency (PHEC) refers to the specific qualities that health educators should have to conduct effective health education activities with patients.[9,10] PHEC is an essential professional competency for nursing staff and determines the quality of patient education.[11-14] However, in existing studies, the PHEC of clinical nurses is often the lowest-rated area of nursing competency.[15,16] Therefore, the development and strengthening of PHEC for nurses are extremely important to improve the quality of patient education, patient care, patient safety, and the development of nursing careers. In addition, we should pay attention to nursing students' PHEC because they are the primary reserve of the clinical nurse workforce.

Accurate measurement of PHEC is important because it can be used to assess the PHEC status of nursing personnel and to develop targeted strategies based on the nursing personnel's PHEC. Moreover, it can be used in research to assess the effectiveness of relevant PHEC interventions. Currently, relevant measurement instruments have been developed internationally: for example, a scale for measuring the PHEC of registered nurses developed by Lin et al. in 2017,[17] a PHEC competency assessment scale developed by Hwang et al. based on a literature review and the Delphi method,[18] and a Spanish version of the nurse PHEC scale developed by Pueyo-Garrigues et al.[19] Although related instruments are available for assessing PHEC in

nursing personnel, these evaluation instruments have been developed in different settings and their validation varies considerably, with none considered the gold standard.

In this study, we defined PHEC as the specific qualities that must be possessed by nursing personnel to provide health education to patients, including knowledge, skills, beliefs or attitudes, self-concept, personality qualities, and motivation. Although there has been a review of PHEC measurement instruments for nursing staff, this review has some limitations on its rigor.[20] First, this review included not only measurement instruments for PHEC but also systems for evaluating PHEC, which are different from measurement instruments. Second, this review did not systematically evaluate the measurement properties of instruments for measuring PHEC based on related guidelines. However, a systematic and comprehensive review of PHEC measurement instruments is crucial for guiding the selection of instruments and/or guiding the development and refinement of high-quality instruments in the future. The consensus-based standards for the selection of health measurement instruments (COSMIN) methodology provides resources to systematically review measurement instruments and evaluate them in terms of both methodological quality and quality of measurement properties to select instruments that are of high quality for study purposes and provide an evidence-based foundation for future high-level instrument development.[21] Eskolin et al. conducted a review on instruments assessing nurses' competence in the empowerment of patient education.[22] However, in this review, the author did not give a clear and specific definition of ‘empowering patient education competence of nurses’. This may lead to an unclear research boundary. Their investigation encompassed not only instruments appraising nurses’ PHEC but also instruments evaluating the quality of patient education provided by health care professionals. Furthermore, they included tools for measuring nurses’ attitudes toward patient education. Considering the

importance of nursing personnel in patient health education, and to ensure a more distinct scope and targeted content, our study will focus specifically on the PHEC measurement instruments, which are designed specifically for nursing personnel, including both nurses and nursing students. Furthermore, in our review, we will incorporate Chinese databases, unveiling more qualified instruments that align with our stringent criteria. Thus, this study is designed to conduct a comprehensive and rigorous systematic review of PHEC assessment instruments based on the COSMIN methodology, to evaluate the measurement properties of these instruments, provide a reference for nursing personnel and researchers to accurately and effectively assess PHEC, and provide recommendations for researchers to develop and improve PHEC assessment instruments.

This systematic review will address the following questions: (1) What instruments are available for assessing the PHEC of nursing personnel? (2) What are the characteristics of these instruments? (3) What is the methodological quality of studies on the measurement properties of these instruments? (4) What are these instruments' measurement properties, interpretability, and feasibility? (5) What are the similarities and differences between these instruments? (6) What are the knowledge and research gaps in the assessment of PHEC of nursing personnel?

2. METHODS

The COSMIN guideline for systematic reviews of PROMs will be used to guide the implementation of the systematic review. PRISMA-P 2015 checklist and PRISMA 2020 checklist will be used to guide the reporting of the protocol and systematic review, respectively.[21,23,24] We registered the protocol in the International Prospective Register of Systematic Reviews (PROSPERO, CRD42023393293). The inconsistency between this protocol and that registered on PROSPERO and the reasons for this are

shown in Table S1.

Inclusion and exclusion criteria of studies

Inclusion criteria

Studies will be included if they (1) address instrument(s) for measuring the PHEC of nurses or nursing students, (2) describe the processes of development and evaluation of one or more measurement properties for eligible instrument(s), (3) discuss instruments designed to measure the PHEC of health professionals (the literature explicitly mentions that it applies to nursing personnel as well), and (4) have full-text availability. If full-text versions of the studies are not available online, the authors of these articles will be contacted, and articles for which valid information was not available after contacting the authors will be excluded. We will limit the included studies to those written in English and Chinese.

Exclusion criteria

Studies will be excluded if they are (1) not primary studies (e.g., biographies, addresses, and editorials) or are case studies, (2) reports that used the instruments only for outcome measurements, (3) secondary studies (e.g., reviews and/or systematic reviews), or (4) duplicate published studies.

Search strategy

A systematic search will be performed between March 1, 2023, and March 31, 2023, in six English databases (i.e., CINAHL, EMBASE, Ovid Medline, PubMed, PsycINFO, and Web of Science) and two Chinese databases (i.e., CNKI and WANFANG DATA). We include Chinese databases since the researchers speak Chinese as their native language. We will also search for and screen references of all eligible literature. The search time limit is from the library's creation date to March 31, 2023. A literature search will be conducted using a combination of subject terms and free words. The

major search concepts will be nursing, health education, competence, instrument, and measurement properties. Related comprehensive and sensitive search strategies developed by other researchers will also be used in this literature search, including (1) the search filter developed by the University of Oxford for finding PROMs,[25] (2) the sensitive PubMed search filter for measuring attributes developed by Terwee et al., and (3) corresponding search filters applicable to other databases.[26] We will examine results reported by nurses or nursing students, so the first filter will be adjusted appropriately (e.g., we will remove those sections that are relevant to the quality of life and patient-reported outcomes). The search strategy constructed for PubMed is described in Table S2 in the supplementary file. The search strategy for the Chinese databases is shown in Table S3 in the supplementary file.

Study screening

Covidence will be used to manage the references.[27] First, duplicates from the eight databases will be removed with Covidence. After the initial screening, both researchers will independently review and screen titles, abstracts, and full-text articles with the support of Covidence. In case of disagreement, a third researcher will be consulted to screen the literature. The screening processes of this study are shown in Figure 1.

Data extraction

The two researchers will independently extract data from the included papers and resolve their differences through discussion. We will extract the data about the characteristics of the instruments (including instrument name, developer[s]/year developed, construct[s], targeted population, mode of administration, recall period, [sub]scale[s]/[number of items], response options, range of scores/scoring, original language, and available translations; see Table S4 in the supplementary file), the

characteristics of the included populations (including sample size, mean of age, gender, setting, country, and language; see Table S5 in the supplementary file,), the results on the psychometric properties (Table S6 in the supplementary file), and information about the interpretability (Table S7 in the supplementary file) and feasibility (Table S8 in the supplementary file) of the included instruments.

The term 'outcome measure instrument development' will be used instead of the original 'patient-reported outcome measure development' to more accurately reflect the inclusion of studies that examined outcomes reported by nurses or nursing students rather than patients.

Quality appraisal and Data synthesis

Two researchers will independently assess the quality of eligible studies using the COSMIN Risk of Bias checklist, which is divided into three sections: content validity (instrument development and content validity), internal structure (structural validity, internal consistency, and cross-cultural validity/measurement invariance), and other measurement properties (reliability, measurement error, criterion validity, hypothesis testing for construct validity, and responsiveness).[21,23,28] Each measurement property will be evaluated by different items provided by the COSMIN Risk of Bias checklist, and the items will be rated on a five-level score of 'very good', 'adequate', 'doubtful', 'inadequate', or 'not applicable.'[23,28] Based on the 'the worst score counts' principle, each measurement property's overall methodological quality score is expressed by taking the lowest rating of any standard in the box.[23,29] Subsequently, the two researchers will apply the updated criteria for good measurement properties alone to evaluate the reliability and validity of the instruments themselves, and the quality of the evidence will be graded using the Grading of Recommendations Assessment, Development, and Evaluation (GRADE) approach.[23,29] In case of

disagreement, a third researcher will be consulted.

We will work using the following three steps. In the first step, two investigators will apply the COSMIN Risk of Bias checklist to evaluate the methodological quality of each eligible study individually.[28] The final consensus on the results of the methodological quality will be presented in Tables S9 and S9-1 in the supplementary file. In the second step, the updated criteria for good measurement properties will be applied to evaluate the quality of evidence for each measured property, and the evaluation results will be shown in Tables S6 and S6-1 in the supplementary file.[23,29] This section mainly evaluates the strengths and weaknesses of the measurement properties. Among these, the quality of content validity will be evaluated according to the COSMIN methodology for content validity in three aspects: the relevance, comprehensiveness, and comprehensibility of items, which can be 'sufficient (+)', 'insufficient (-)', 'indeterminate (?)', or 'inconsistent (±)'.[29,30] The quality of the remaining measurement properties (structural validity, internal consistency, cross-cultural validity, measurement invariance, reliability, measurement error, criterion validity, construct validity, and responsiveness) will be evaluated by applying the COSMIN quality criteria, which can be 'sufficient (+)', 'insufficient (-)', and 'indeterminate (?)'.[23] The corresponding results will be reported in the rating columns of Table S6 in the supplementary file, and the results of rating content validity will be presented separately in Table S6-1 in the supplementary file. In the third step, a modified GRADE approach will be used to rate the quality of the above evidence, reflecting the level of confidence in the quality of the evidence. To evaluate the content's validity, three of these factors are applicable: risk of bias, inconsistency, and indirectness.[29] Assuming that the level of evidence quality for each of the remaining measurement properties is high, the quality of the evidence will be downgraded by

considering the following factors: risk of bias, inconsistency, imprecision, and indirectness.[23] The quality of evidence will be divided into four levels: 'high', 'moderate', 'low', or 'very low'.[21,23] The corresponding results will be displayed in Table S10 in the supplementary file. Two investigators will independently grade and cross-check the results. In case of disputes, final decisions will be made in consultation with the third investigator.

Patient and public involvement

Neither patients nor the public will be involved in this study.

Ethics and dissemination

Ethical approval is not applicable for this study. We will share the findings from the study at national and/or international conferences and in a peer-reviewed journal in the fields of health education and/or patient education.

3. DISCUSSION

To our knowledge, this will be the first COSMIN-based systematic review of PHEC assessment instruments for nursing personnel , which will be reported following the Preferred Reporting Items for Systematic Reviews and Meta-analyses protocols (PRISMA-P) 2020 checklist. This systematic review will provide a comprehensive rating of the level of evidence for each measurement property of the PHEC assessment instruments, which will be based on an evaluation of the measurement properties of all included instruments and the methodological quality of the studies. Through this study, we will be able to develop recommendations on the use of existing qualified instruments in clinical practice and research that could assist nursing personnel and researchers in the accurate and valid assessment of PHEC. This review may provide an evidence-based foundation for the development, design, validation, and use of future instruments by identifying problems in instrument development and validation and therefore help

researchers to develop and improve these instruments.

Author contributions All authors have read and agreed to the published version of the manuscript.

Conceptualization: QC, ST, SW;

Methodology: QC, ST, ZS;

Data curation: QC, SW, KL;

Writing—original draft preparation: SW, KL, QC;

Writing—review and editing: QC, ST, ZS;

Supervision: QC and ST.

Funding This work was supported by the National Natural Science Foundation of China (No. 72104250) and the Natural Science Foundation of Hunan Province (No. 2022JJ40642).

Competing interests None declared.

Patient consent Not required.

Data sharing statement No additional data available.

List of Figures:

List of Supplementary File Tables:

Table S6: Rating the measurement properties of the instruments

Table S6-1: Rating of the content validity of instruments

Table S7: Information on interpretability of instruments

Table S8: Information on feasibility of instruments

Table S9: Quality of studies on measurement properties

Table S9-1: Quality of the instrument development

Table S10: Quality of the evidence for measurement properties of the instruments

References:

1. Organization WH. Health education: theoretical concepts, effective strategies and core competencies: a foundation document to guide capacity development of health educators. 2012. Available from: https://apps.who.int/iris/bitstream/handle/10665/119953/EMRPUB_2012_EN_1362.pdf?sequen accessed September 11 2023.
2. Turner A, Anderson JK, Wallace LM, et al. An evaluation of a self-management program for patients with long-term conditions. *Patient Educ Couns* 2015;98(2):213-9. doi: 10.1016/j.pec.2014.08.022
3. Lin P, Gao C, Su S, et al. Teaching principles: The applications in nursing practice. *Taipei: Farseeing.* 2005.
4. Siegel KR, Ali MK, Zhou X, et al. Cost-effectiveness of interventions to manage diabetes: has the evidence changed since 2008? *Diabetes Care* 2020;43(7):1557-92. doi: 10.2337/dci20-0017
5. Koutsopoulou S, Papathanassoglou ED, Katapodi MC, et al. Nurses as information providers to cancer patients: Critical review of the evidence. *J Clin Nurs* 2010;19(5-6):749-65. doi: 10.1111/j.1365-2702.2009.02954.x

6. Feo R, Rasmussen P, Wiechula R, et al. Developing effective and caring nurse-patient relationships. *Nurs Stand* 2017;31(28):54-63. doi: 10.7748/ns.2017.e10735
7. Kwame A, Petrucka PM. A literature-based study of patient-centered care and communication in nurse-patient interactions: barriers, facilitators, and the way forward. *BMC Nurs* 2021;20(1):158. doi: 10.1186/s12912-021-00684-2
8. Ghiyasvandian S, Zakerimoghadam M, Peyravi H. Nurse as a facilitator to professional communication: a qualitative study. *Glob J Health Sci* 2014;7(2):294-303. doi: 10.5539/gjhs.v7n2p294
9. Hwang H-L, Kuo T-Y. Competency in delivering health education: A concept analysis. *J Interprof Educ Pract* 2018;11:20-5. doi: https://doi.org/10.1016/j.xjep.2018.02.005
10. Karimi Moonaghi H, Emami Zeydi A, Mirhaghi A. Patient education among nurses: bringing evidence into clinical applicability in Iran. *Invest Educ Enferm* 2016;34(1):137-51. doi: 10.17533/udea.iee.v34n1a16
11. Flinkman M, Leino-Kilpi H, Numminen O, et al. Nurse Competence Scale: a systematic and psychometric review. *J Adv Nurs* 2017;73(5):1035-50. doi: 10.1111/jan.13183
12. Ko Y, Yu S. Core nursing competency assessment tool for graduates of outcome-based nursing education in South Korea: A validation study. *Jpn J Nurs Sci* 2019;16(2):155-71. doi: 10.1111/jjns.12223
13. Nilsson J, Engström M, Florin J, et al. A short version of the nurse professional competence scale for measuring nurses' self-reported competence. *Nurse Educ Today* 2018;71:233-9. doi: 10.1016/j.nedt.2018.09.028

14. Wilkinson CA. Competency assessment tools for registered nurses: An integrative review. *J Contin Educ Nurs* 2013;44(1):31-7. doi: 10.3928/00220124-20121101-53
15. Cowan DT, Wilson-Barnett DJ, Norman IJ, et al. Measuring nursing competence: Development of a self-assessment tool for general nurses across Europe. *Int J Nurs Stud* 2008;45(6):902-13. doi: 10.1016/j.ijnurstu.2007.03.004
16. Pueyo-Garrigues M, Whitehead D, Pardavila-Belio MI, et al. Health education: a Rogerian concept analysis. *Int J Nurs Stud* 2019;94:131-8. doi: 10.1016/j.ijnurstu.2019.03.005
17. Lin LY, Wang RH. Patient Education Competence Scale for Registered Nurses in Taiwan: Scale development and psychometric validation. *Jpn J Nurs Sci* 2017;14(2):117-25. doi: 10.1111/jjns.12141
18. Hwang HL, Kuo ML, Tu CT. Health education and competency scale: Development and testing. *J Clin Nurs* 2018;27(3-4):e658-e67. doi: 10.1111/jocn.14116
19. Pueyo-Garrigues M, Pardavila-Belio MI, Whitehead D, et al. Nurses' knowledge, skills and personal attributes for competent health education practice: An instrument development and psychometric validation study. *J Adv Nurs* 2021;77(2):715-28. doi: 10.1111/jan.14632
20. Mi Y, Wu D, Wei ZZ, et al. Research progress on evaluation tool for health education competency of nursing staff. *Chinese Nursing Research* 2020;34(11):1983-7. doi: 10.12102/j.issn.1009-6493.2020.11.023
21. Prinsen CA, Mokkink LB, Bouter LM, et al. COSMIN guideline for systematic reviews of patient-reported outcome measures. *Qual Life Res* 2018;27:1147-57. doi: 10.1007/s11136-018-1798-3

22. Eskolin SE, Inkeroinen S, Leino-Kilpi H, et al. Instruments for measuring empowering patient education competence of nurses: Systematic review. *J Adv Nurs* 2023;79(7):2414-28. doi: 10.1111/jan.15597

23. Mokkink LB, Prinsen C, Patrick DL, et al. COSMIN methodology for systematic reviews of patient-reported outcome measures (PROMs). User manual. 2018;78(1):6-63. Available from: https://www.cosmin.nl/wp-content/uploads/COSMIN-syst-review-for-PROMs-manual_version-1_feb-2018-1.pdf accessed September 11 2023.

24. Page MJ, Moher D, Bossuyt PM, et al. PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. *BMJ* 2021;372. doi: 10.1136/bmj.n160

25. Mackintosh A, Comabella C, Hadi M, et al. PROM group construct & instrument type filters. *Department of Public Health, University of Oxford, Oxford.* 2010. Available from: https://cosmin.nl/wp-content/uploads/prom-search-filter-oxford-2010.pdf accessed September 11 2023.

26. Terwee CB, Jansma EP, Riphagen II, et al. Development of a methodological PubMed search filter for finding studies on measurement properties of measurement instruments. *Qual Life Res* 2009;18:1115-23. doi: 10.1007/s11136-009-9528-5

27. Babineau J. Product review: Covidence (systematic review software). *Journal of the Canadian Health Libraries Association/Journal de l'Association des bibliothèques de la santé du Canada* 2014;35(2):68-71.

28. Mokkink LB, De Vet HC, Prinsen CA, et al. COSMIN risk of bias checklist for systematic reviews of patient-reported outcome measures. *Qual Life Res* 2018;27:1171-9. doi: 10.1007/s11136-017-1765-4

29. Terwee CB, Prinsen C, Chiarotto A, et al. COSMIN methodology for assessing the content validity of PROMs–user manual. *Amsterdam: VU University Medical Center.* 2018. Available from: https://cosmin.nl/wp-content/uploads/COSMIN-methodology-for-content-validity-user-manual-v1.pdf accessed September 11 2023.

30. Terwee CB, Prinsen CA, Chiarotto A, et al. COSMIN methodology for evaluating the content validity of patient-reported outcome measures: a Delphi study. *Qual Life Res* 2018;27:1159-70. doi: 10.1007/s11136-018-1829-0

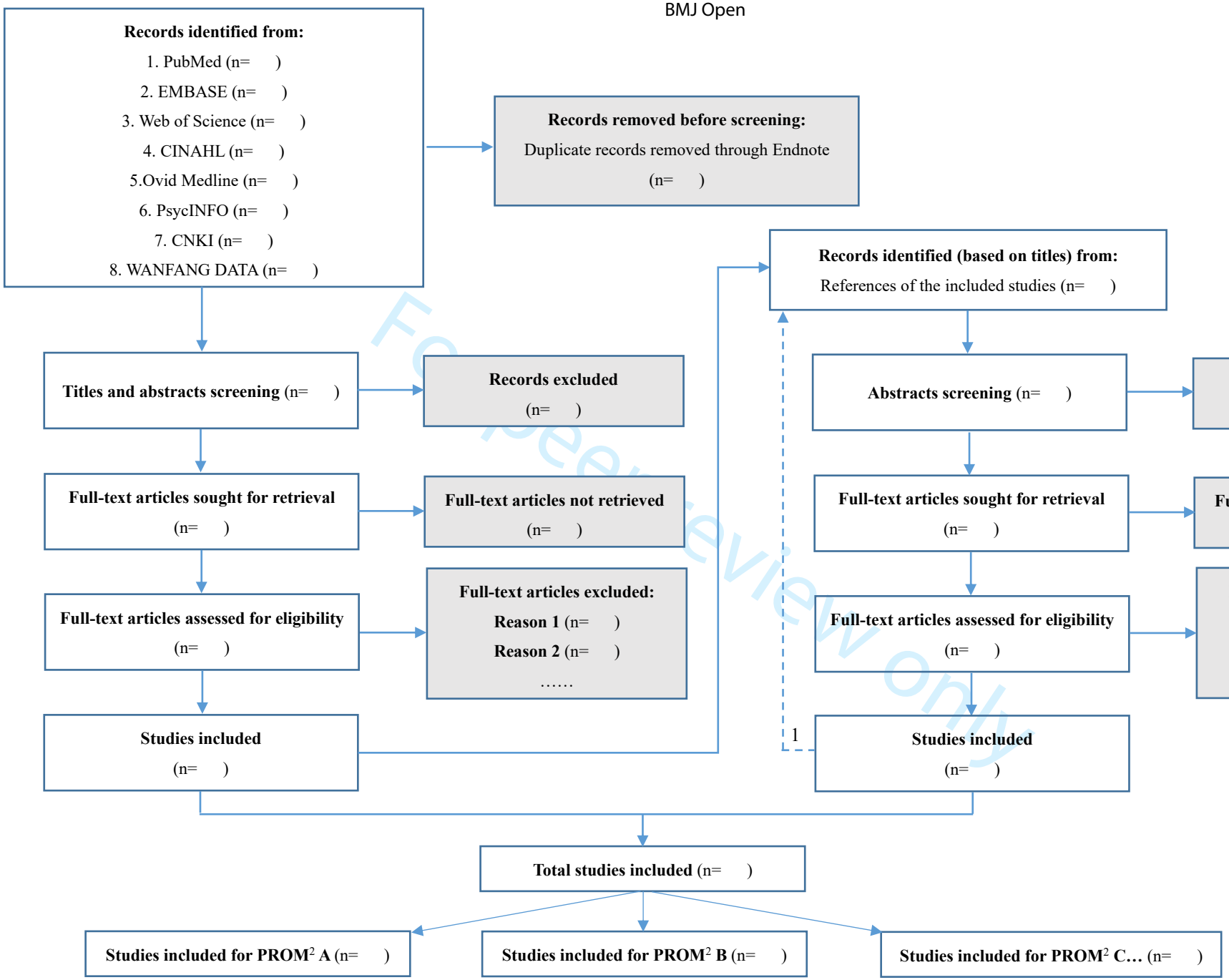


Figure 1. Flowchart of literature selection process

Note: 1. References of all included studies will be manually screened until no eligible studies can be identified.

2. PROM: patient-reported outcome measure.

Table S1. Reasons for inconsistencies

Revised content	Reason
1. Review title	Taking into consideration the inclusive scope of the article, encompassing both nursing staff and nursing students, a deliberate revision has been undertaken in the title adjustment, transitioning from the initial term "nursing staff" to the more comprehensive descriptor "nursing personnel".
2. Funding	In the PROSPERO registration, the Funding sources/sponsors were initially documented as "None." However, it is important to note that during the course of the study, we secured pertinent funding support. Consequently, the Funding section in the protocol was subsequently revised to accurately reflect this development.

Table S2. Search strategy for PubMed

1	#1	nurs*
2	#2	((((((((((Health Education[Mesh] OR Patient Education as Topic[Mesh]) OR (health education*[tiab])) OR (education, health[tiab])) OR (patient education*[tiab])) OR (education program*[tiab])) OR (educational program*[tiab])) OR (education, patient*[tiab])) OR (education of patient*[tiab])) OR (nursing education*[tiab])) OR (hospital education*[tiab])) OR (educational activit*[tiab])) OR ((educat*[tiab]) AND (individual patient*[tiab]))
3	#3	(((((((professional competence[MeSH]) OR (professional competence[tiab])) OR (competence, professional[tiab])) OR (generalization of expertise[tiab])) OR (expertise generalization[tiab])) OR (technical expertise[tiab])) OR (expertise, technical[tiab])) OR (competenc*[tiab] OR capabilit*[tiab] OR capacit*[tiab] OR abilit*[tiab])) OR ((skill*[tiab] OR belief[tiab] OR attitude[tiab]) AND (knowledge[tiab]))
4	#4	(report[tiab] OR reported[tiab] OR reporting[tiab] OR rated[tiab] OR rating[tiab] OR ratings[tiab] OR based[tiab] OR assessed[tiab] OR assessment[tiab] OR assessments[tiab] OR disability[tiab] OR function[tiab] OR functional[tiab] OR functions[tiab] OR subjective[tiab] OR utility[tiab] OR utilities[tiab] OR wellbeing[tiab] OR well being[tiab]) AND (index[tiab] OR indices[tiab] OR instrument[tiab] OR instruments[tiab] OR measure[tiab] OR measures[tiab] OR questionnaire[tiab] OR questionnaires[tiab] OR profile[tiab] OR profiles[tiab] OR scale[tiab] OR scales[tiab] OR score[tiab] OR scores[tiab] OR status[tiab] OR survey[tiab] OR surveys[tiab])

5	#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 reproducib*[tiab] OR "discriminant analysis"[MeSH] OR reliab*[tiab] OR unreliab*[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 alphas[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 intertester[tiab] OR inter-tester[tiab] OR intratester[tiab] OR intra-tester[tiab] OR interobserver[tiab] OR inter-observer[tiab] OR intraobserver[tiab] OR intra-observer[tiab] OR intertechnician[tiab] OR inter-technician[tiab] OR intratechnician[tiab] OR intra-technician[tiab] OR interexaminer[tiab] OR inter-examiner[tiab] OR intraexaminer[tiab] OR intra-examiner[tiab] OR interassay[tiab] OR inter-assay[tiab] OR intraassay[tiab] OR intra-assay[tiab] OR interindividual[tiab] OR inter-individual[tiab] OR intraindividual[tiab] OR intra-individual[tiab] OR interparticipant[tiab] OR inter-participant[tiab] OR intraparticipant[tiab] OR intra-participant[tiab] OR kappa[tiab] OR kappa's[tiab] OR kappas[tiab] OR repeatab*[tw] OR ((replicab*[tw] OR repeated[tw]) AND (measure[tw] OR measures[tw] OR findings[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] OR "interval variability"[tiab] OR "rate variability"[tiab] OR (variability[tiab] AND (analysis[tiab] OR values[tiab])) OR (uncertainty[tiab] AND (measurement[tiab] OR measuring[tiab])) OR "standard error of measurement"[tiab] OR sensitiv*[tiab] OR responsive*[tiab] OR (limit[tiab] AND detection[tiab]) OR "minimal detectable concentration"[tiab] OR interpretab*[tiab] OR ((minimal[tiab] OR minimally[tiab] OR clinical[tiab] OR clinically[tiab]) AND (important[tiab] OR significant[tiab] OR detectable[tiab]) AND (change[tiab] OR difference[tiab])) OR (small*[tiab] AND (real[tiab] OR detectable[tiab]) AND (change[tiab] OR difference[tiab])) OR "meaningful change"[tiab] OR "ceiling effect"[tiab] OR "floor effect"[tiab] OR "Item response model"[tiab] OR IRT[tiab] OR Rasch[tiab] OR "Differential item

		functioning"[tiab] OR DIF[tiab] OR "computer adaptive testing"[tiab] OR "item bank"[tiab] OR "cross-cultural equivalence"[tiab])
6	#6	("addresses"[Publication Type] OR "biography"[Publication Type] OR "case reports"[Publication Type] OR "comment"[Publication Type] OR "directory"[Publication Type] OR "editorial"[Publication Type] OR "festschrift"[Publication Type] OR "interview"[Publication Type] OR "lectures"[Publication Type] OR "legal cases"[Publication Type] OR "legislation"[Publication Type] OR "letter"[Publication Type] OR "news"[Publication Type] OR "newspaper article"[Publication Type] OR "patient education handout"[Publication Type] OR "popular works"[Publication Type] OR "congresses"[Publication Type] OR "consensus development conference"[Publication Type] OR "consensus development conference, nih"[Publication Type] OR "practice guideline"[Publication Type]) NOT ("animals"[MeSH Terms] NOT "humans"[MeSH Terms])
7	#7	#1 AND #2 AND #3 AND #4 AND #5 NOT #6

Note: "*" to include all derivatives of that word or concept.

Table S3. Search strategy for Chinese databases

1	#1	TKA = 护理 OR TKA = 护士 OR TKA = 护生
2	#2	TKA = 健康教育能力 OR TKA = 患者教育能力 OR TKA = 健康教育胜任力 OR TKA = 患者教育胜任力
3	#3	SU = 评估 OR SU = 测量 OR SU = 评价 OR SU = 收集 OR SU = 调查 OR SU = 工具 OR SU = 问卷 OR SU = 量表 OR SU = 仪器 OR SU = 研究
4	#4	TKA = 信度 OR TKA = 效度 OR TKA = 反应度 OR TKA = 内部一致性 OR TKA = 稳定性 OR TKA = 相关系数 OR TKA = 克朗巴赫系数 OR TKA = 探索性因子分析 OR TKA = 验证性因子分析 OR TKA = 探索性因素分析 OR TKA = 验证性因素分析 OR TKA = 检验 OR TKA = 结果
5	#5	#1 AND #2 AND #3 AND #4

Note: TKA = title/abstract; SU = title/abstract/keywords.

Table S4. Characteristics of the included instruments

Instrument name	Developer(s)/ year developed	Construct (s)	Target population	Mode of administration	Recall period	(Sub)scale (s) (number of items)	Response options	Range of scores/scoring	Original language	Available translations

Table S5. Characteristics of the included study populations

		Population			Instrument administration			
Instrument	**Ref**	**N**	**Age Mean (SD, range) yr**	**Gender % female**	**Setting**	**Country**	**Language**	**Response rate**
A	1							
	2							
	3							
B	1							

Table S6. Rating the measurement properties of the instruments

Instrument	Study 1			Study 2			Study 3			OVERALL					
	RATING	RATING	RATING	RATING	RATING	RATING	RATING	RATING	RATING	OVERALL RATING	OVERALL RATING	OVERALL RATING	QUALITY OF EVIDENCE	QUALITY OF EVIDENCE	QUALITY OF EVIDENCE
	+ / - / ?	+ / - / ?	+ / - / ?	+ / - / ?	+ / - / ?	+ / - / ?	+ / - / ?	+ / - / ?	+ / - / ?	+ / - / ?	+ / - / ?	+ / - / ?	High, moderate, low, very low	High, moderate, low, very low	High, moderate, low, very low
	rater 1	rater 2	consensus	rater 1	rater 2	consensus	rater 1	rater 2	consensus	rater 1	rater 2	consensus	rater 1	rater 2	consensus
Structural validity															
internal consistency															
Cross-cultural validity															
Measurement invariance															
Reliability															
Measurement error															
Criterion validity															
Construct validity															
Responsiveness															

Note: “+” = sufficient; “-” = insufficient; “?” = indeterminate.

Table S6-1. Rating of the content validity of instruments

Instrument (Reference – study type/Rating of reviewers)	Content Validity												
	Relevance[1]						Comprehensiveness[1]		Comprehensibility[1]				CONTENT VALIDITY RATING[2]
	1. Are the included items relevant for the construct of interest?	2. Are the included items relevant for the target population of interest?	3. Are the included items relevant for the context of use of interest?	4. Are the response options appropriate?	5. Is the recall period appropriate?	RELEVANCE RATING[2]	6. Are all key concepts included?	COMPREHENSIVENESS RATING[2]	7. Are the PROM instructions understood by the population of interest as intended?	8. Are the PROM items and response options understood by the population of interest as intended?	9. Are the PROM items appropriately worded?	10. Do the response options match the question?	COMPREHENSIBILITY RATING[2]	
A (Ref 1- instrument development study)														
A (Ref 2 - Content validity study)														
A (Ref 3 - Content validity study)														
Rating of reviewers														
B (Ref 1- instrument development study)														
B (Ref 2 - Content validity study)														
Rating of reviewers														

……														

Note:1. Rating for the 10 criteria for relevance, comprehensiveness, comprehensibility can be + / − /±/ ?: ‘ + ’= sufficient, ‘ − ’= insufficient, ‘±’ = inconsistent,‘?’ =indeterminate.

2. The RELEVANCE, COMPREHENSIVENESS, COMPREHESIBILITY, AND CONTENT VALIDITY rating can be + / − /±/ ? : ‘ + ’= sufficient, ‘ − ’= insufficient, ‘±’ = inconsistent, ‘?’ =indeterminate.

Table S7. Information on interpretability of instruments

Instrument (ref)		**Percentage of missing items and percentage of missing total scores**	**Floor and ceiling effects**	**Scores and change scores available for relevant (sub)groups**	**Minimal important change (MIC) or minimal important difference (MID)**	**Information on response shift**
Instrument A (ref 1)						
Instrument A (ref 2)						
Instrument A (ref 3)						
Instrument B (ref 1)						
……						

Table S8. Information on feasibility of instruments

Feasibility aspects	Instrument A	Instrument B	Instrument C	Instrument D
Patient's comprehensibility				
Clinician's comprehensibility				
Type and ease of administration				
Length of the instrument				
Completion time				
Patient's required mental and physical ability level				
Ease of standardization				
Ease of score calculation				
Copyright				
Cost of an instrument				
Required equipment				
Availability in different settings				
Regulatory agency's requirement for approval				

Table S9. Quality[1] of studies on measurement properties

Instrument	Content validity[2]					Structural validity	Internal consistency	Cross-cultural validity	Reliability	Measurement error	Criterion validity	Co
	Asking patients			Asking experts								
	Relevance	Comprehensiveness	Comprehensibility	Relevance	Comprehensiveness							Converg validit
A												
B												
……												

Note: 1. Quality: V = very good, A = adequate, D = doubtful, I = inadequate.
2. Given that the criteria and rating systems for evaluating the content validity of instruments are different from those for other measurement properties, the quality results of content validity are not included in this table but separately shown in following Table S8-1.

Table S9-1. Quality[1] of the instrument development

Instrument	PROM design							Cognitive interview (CI) study[3]				TOTAL PROM DEVELOPMENT
	General design requirements					Concept elicitation[2]	Total PROM design	General design requirements	Comprehensibility	Comprehensiveness	Total CI study	
	Clear construct	Clear origin of construct	Clear target population for which the PROM was developed	Clear context of use	PROM developed in sample representing the target population			CI study performed in sample representing the target population				
A												
B												
……												

Note: 1. Quality: V = very good, A = adequate, D = doubtful, I = inadequate.
2.The concept elicitation will not be further rated if the instrument(s) was not developed in the sample representing the target population.
3. Empty cells indicate that a CI study (or part of it) was not performed.

Table S10. Quality of the evidence for measurement properties of the instruments (Summary of findings)

Instrument	Content validity		Structural validity		Internal consistency		Cross-cultural validity		Reliability		Measurement error		Criterion validity		Hypotheses testing		Responsiveness	
	Overall Rating[1]	Quality of Evidence[3]	Overall Rating[2]	Quality of evidence[3]	Overall rating[2]	Quality of evidence[3]	Overall rating[2]	Quality of evidence[3]	Overall rating[2]	Quality of evidence[3]	Overall rating[2]	Quality of evidence[3]	Overall rating[2]	Quality of evidence[3]	Overall rating[2]	Quality of evidence[3]	Overall rating[2]	Quality of evidence[3]
Instrument A																		
Instrument B																		
Instrument C																		
……																		

Note:1. Overall ratings for the content validity (relevance, comprehensiveness, comprehensibility) can only be + / − /±: ‘ + ’= sufficient, ‘ − ’= insufficient, ‘±’ = inconsistent.

2. Overall ratings for other measurement properties can be + / − /±/ ?: ‘ + ’= sufficient, ‘ − ’= insufficient, ‘±’ = inconsistent, ‘?’ =indeterminate.

3. Ratings for quality of evidence: High, Moderate, Low, Very low.

PRISMA-P (Preferred Reporting Items for Systematic review and Meta-Analysis Protocols) 2015 checklist: recommended items to address in a systematic review protocol*

Section and topic	Item No	Checklist item	Pg.
ADMINISTRATIVE INFORMATION			
Title:			
Identification	1a	Identify the report as a protocol of a systematic review	Pg. 1
Update	1b	If the protocol is for an update of a previous systematic review, identify as such	N/A
Registration	2	If registered, provide the name of the registry (such as PROSPERO) and registration number	Pg. 2 and Pg. 6
Authors:			
Contact	3a	Provide name, institutional affiliation, e-mail address of all protocol authors; provide physical mailing address of corresponding author	Pg. 1
Contributions	3b	Describe contributions of protocol authors and identify the guarantor of the review	Pg. 11-12
Amendments	4	If the protocol represents an amendment of a previously completed or published protocol, identify as such and list changes; otherwise, state plan for documenting important protocol amendments	N/A
Support:			
Sources	5a	Indicate sources of financial or other support for the review	Pg. 12
Sponsor	5b	Provide name for the review funder and/or sponsor	N/A
Role of sponsor or funder	5c	Describe roles of funder(s), sponsor(s), and/or institution(s), if any, in developing the protocol	N/A
INTRODUCTION			
Rationale	6	Describe the rationale for the review in the context of what is already known	Pg. 3-6
Objectives	7	Provide an explicit statement of the question(s) the review will address with reference to participants, interventions, comparators, and outcomes (PICO)	Pg. 6
METHODS			
Eligibility criteria	8	Specify the study characteristics (such as PICO, study design, setting, time frame) and report characteristics (such as years considered, language, publication status) to be used as criteria for eligibility for the review	Pg. 7
Information sources	9	Describe all intended information sources (such as electronic databases, contact with study authors, trial registers or other grey literature sources) with planned dates of coverage	Pg. 7
Search strategy	10	Present draft of search strategy to be used for at least one electronic database, including planned limits, such that it could be repeated	Supplemental file
Study records:			
Data management	11a	Describe the mechanism(s) that will be used to manage records and data throughout the review	Pg. 8

Selection process	11b	State the process that will be used for selecting studies (such as two independent reviewers) through each phase of the review (that is, screening, eligibility and inclusion in meta-analysis)	Pg. 8
Data collection process	11c	Describe planned method of extracting data from reports (such as piloting forms, done independently, in duplicate), any processes for obtaining and confirming data from investigators	Pg. 8-9 Supplemental file
Data items	12	List and define all variables for which data will be sought (such as PICO items, funding sources), any pre-planned data assumptions and simplifications	Pg. 8-9 Supplemental file
Outcomes and prioritization	13	List and define all outcomes for which data will be sought, including prioritization of main and additional outcomes, with rationale	Pg. 8-9 Supplemental file
Risk of bias in individual studies	14	Describe anticipated methods for assessing risk of bias of individual studies, including whether this will be done at the outcome or study level, or both; state how this information will be used in data synthesis	Pg. 9
Data synthesis	15a	Describe criteria under which study data will be quantitatively synthesised	N/A
	15b	If data are appropriate for quantitative synthesis, describe planned summary measures, methods of handling data and methods of combining data from studies, including any planned exploration of consistency (such as I^2, Kendall's τ)	N/A
	15c	Describe any proposed additional analyses (such as sensitivity or subgroup analyses, meta-regression)	N/A
	15d	If quantitative synthesis is not appropriate, describe the type of summary planned	Pg. 10
Meta-bias(es)	16	Specify any planned assessment of meta-bias(es) (such as publication bias across studies, selective reporting within studies)	N/A
Confidence in cumulative evidence	17	Describe how the strength of the body of evidence will be assessed (such as GRADE)	Pg. 9-10

*** It is strongly recommended that this checklist be read in conjunction with the PRISMA-P Explanation and Elaboration (cite when available) for important clarification on the items. Amendments to a review protocol should be tracked and dated. The copyright for PRISMA-P (including checklist) is held by the PRISMA-P Group and is distributed under a Creative Commons Attribution Licence 4.0.**

From: Shamseer L, Moher D, Clarke M, Ghersi D, Liberati A, Petticrew M, Shekelle P, Stewart L, PRISMA-P Group. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. BMJ. 2015 Jan 2;349(jan02 1):g7647.