BMJ Open

Psychometric properties of self-reported measures of active ageing: a systematic review protocol using COSMIN methodology

Shuyu Han ,[1] Mengmeng Ji,[1] Minmin Leng,[1] Jia Zhou,[1] Zhiwen Wang[1,2]

To cite: Han S, Ji M, Leng M, *et al*. Psychometric properties of self-reported measures of active ageing: a systematic review protocol using COSMIN methodology. *BMJ Open* 2022;**12**:e059360. doi:10.1136/bmjopen-2021-059360

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2021-059360).

Received 19 November 2021
Accepted 15 February 2022

[1]School of Nursing, Peking University, Beijing, China
[2]Peking University Health Science Centre for Evidence-Based Nursing A Joanna Briggs Institute Affiliated Group, Beijing, China

Correspondence to
Dr Zhiwen Wang;
wzwjing@sina.com

ABSTRACT

Introduction Evaluation of active ageing is helpful in making public health policies, improving older adults' quality of life and confronting global ageing challenges. However, there remains no systematic review to summarise all active ageing assessment tools and report their psychometric properties. This study aims to apply the COSMIN (COnsensus-based Standards for the selection of health Measurement INstruments) methodology to review the psychometric properties of active ageing assessment tools obtained by multiple validation studies.

Methods and analysis Studies that aim to validate patient-reported outcome measures (PROMs) of active ageing in older adults aged 60 and over and report one or more psychometric properties are eligible for this systematic review. We will consider studies conducted in any country or setting published either in English or Chinese. The following databases will be searched: PubMed, EMBASE, CINAHL, Web of Science, Cochrane Library, ProQuest Dissertations and Theses, CNKI, and Wanfang. Data extraction, assessment of methodological quality, summary of the quality of PROMs and grading of quality of evidence will be conducted according to the COSMIN methodology.

Ethics and dissemination This study will not collect individual data. Therefore, obtaining ethical approval is not applicable. The results will be disseminated through peer-reviewed journals and conferences and will help researchers choose active ageing assessment tools.

PROSPERO registration number CRD42021287395.

Strengths and limitations of this study

- This is the first systematic review that will identify patient-reported outcome measures that assess active ageing in older adults and will provide a comprehensive picture of their psychometric properties.
- The COSMIN (COnsensus-based Standards for the selection of health Measurement INstruments) methodology will be applied in this study to comprehensively report the psychometric properties of each assessment tool.
- This systematic review will follow the most up-to-date Preferred Reporting Items for Systematic Reviews and Meta-Analyses 2020, and the protocol will follow the Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols 2015.
- This systematic review will only include studies published in English and Chinese and this language restriction may bias the results.

INTRODUCTION

Driven by declining fertility and increasing longevity, population ageing has become a significant challenge throughout the world.[1] According to the World Population Prospects 2019, by 2050 the number of persons aged 65 or over is estimated to more than double compared with 2019 and will reach up to 1.5 billion, increasing from 9% (2019) to 16% (2050).[2] In China, as high as 18.7% and 13.5% of the population in 2020 were persons aged 60 or over and 65 or over, respectively.[3] By 2050, persons aged 65 or over may reach 0.38 billion and account for nearly 30% of the whole Chinese population.[4] Population ageing may have implications for nearly all sectors of the society, including labour and financial markets, demand for goods and services such as housing, transportation and social protection, as well as family structures and intergenerational ties.[5] Therefore, more attention and efforts should be put on this global issue.

The WHO defines 'active aging' as 'the process of optimizing opportunities for health, participation and security in order to enhance quality of life as people age'.[6] The word 'active' refers to continuous participation in the social, financial, political, spiritual and cultural life, and not only in the ability to work and be physically active. Older adults who are unable to work and those living with sickness or disabilities can remain active, while contributing to their families, counterparts, communities and nations.[7] Active ageing provides a more positive perspective of

older adults, that is, regarding them as valuable resources instead of a burden to the society.

Active ageing has become a global goal under the current ageing situation.[8] Evaluation of active ageing will be helpful in making public health policies, improving older adults' quality of life and confronting global ageing challenges. Since accurate and reproducible assessment tools are prerequisite to robust and reliable results, it is significant to choose an acceptable patient-reported outcome measure (PROM) with strong psychometric properties to assess active ageing. Several assessment tools have been validated to assess active ageing, such as the Active Aging Scale for Thai Adults,[8] the Active Aging Index[9] and the University of Jyvaskyla Active Aging Scale.[10] However, these assessment tools were developed in different situations. Their validation also varied significantly and none of them is considered the gold standard. There remains no systematic review that summarises all active ageing assessment tools and reports their psychometric properties. Therefore, this study aims to adopt the COSMIN (COnsensus-based Standards for the selection of health Measurement INstruments) approach[11] to comprehensively report the psychometric properties of active ageing assessment tools obtained by multiple validation studies. Our attempt will be helpful in selecting PROMs of active ageing both in research and in clinical practice.

METHODS AND ANALYSIS

This protocol is reported according to the Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols 2015.[12] This systematic review will follow the COSMIN methodology for conducting systematic reviews of psychometric properties[11] and the Preferred Reporting Items for Systematic Reviews and Meta-Analyses 2020 statement.[13] We also have submitted registration materials prior to the literature search in the International Prospective Register of Systematic Reviews.[14]

Search strategy

We will follow the Peer Review of Electronic Search Strategies (PRESS)[15] to develop the search strategy. First, two researchers (MJ and SH) will independently conduct the primary search in PubMed using both free terms and Medical Subject Headings (MeSH) terms to develop the search words and will fill out the updated PRESS 2015 Guideline Assessment Form.[15] Second, a third researcher (ML) will revise the two forms and assess the inter-rater reliability using the PRESS 2015 Guideline Evidence-Based Checklist.[1] Third, the identified search strategy will be confirmed through a discussion among three researchers. Fourth, a researcher (MJ) will conduct literature search using the identified search strategy in all included databases. The entire process will also be checked by another researcher (SH). Fifth, a researcher (MJ) will review the references of all the included studies to identify eligible literature not found using the search strategies.

We will search the following databases: PubMed, EMBASE, CINAHL, Web of Science, Cochrane Library, ProQuest Dissertations and Theses, CNKI, and Wanfang. The COSMIN filter[16] will be applied in feasible databases. The detailed search strategies for PubMed are available in online supplemental appendix 1.

Inclusion criteria

Population

All adults aged 60 years and older in any country or setting are eligible populations for this systematic review. To provide a comprehensive description of PROMs of active ageing, we will not limit any possible issues that could explain inconsistencies between results, including but not limited to older adults living with or without any disease, institutionalised or non-institutionalised older adults, etc.

Instruments

We will include any type of measurement tools, including but not limited to questionnaires, checklists and scales. They can be self-report, interview-based and proxy report. PROMs that measure active ageing either as a whole or as a subscale will be considered.

Construct

We will apply the definition of 'active ageing' released by the WHO in 2002, which is the process of optimising opportunities for health, participation and security in order to enhance quality of life as people age.[6] Obviously, health, participation and security are three important domains of active ageing. There are several terms similar to active ageing, such as healthy ageing, productive ageing, ageing well, optimal ageing, positive ageing and successful ageing. They may be used interchangeably with active ageing.[17] Although active ageing is the only construct in this study, we will still expand our search strategy according to these similar terms, which is consistent with Kenbubpha *et al*'s study.[18]

Outcomes

We will involve any study that evaluated one or more psychometric properties of a PROM for active ageing according to the COSMIN methodology. Common psychometric property indicators include content validity (relevance, comprehensiveness and comprehensibility), structural validity (comparative fit index, Tucker-Lewis index, root mean square error of approximation or standardised root mean residuals), internal consistency (Cronbach's alpha coefficients), reliability (intraclass correlation coefficient (ICC) or weighted kappa statistics), measurement error (smallest detectable change, limits of agreement or minimal important change), hypothesis testing for construct validity (whether the result is in accordance with the hypothesis), cross-cultural validity/measurement invariance (differential item functioning), criterion validity (correlation with gold standard

Table 1 PROM characteristics

PROM	Target population	Mode of administration (self-report, interview-based or proxy report)	Recall period	Subscale and item number	Range of scores	Original language	Theory	Available translations

PROM, patient-reported outcome measure.

or area under the curve) and responsiveness (area under the curve).

Types of studies

Any original study, such as cross-sectional and longitudinal studies, that developed or translated PROMs and reported psychometric property indicators using their raw data will be included. We only included literature and grey literature in English and Chinese after 2002 (the year WHO launched the definition of active ageing). The specific dates of coverage were January 2002 and February 2022.

Study selection

We will import all the references searched from the databases and remove duplicates in NoteExpress. Two researchers (MJ and SH) who have been trained in evidence-based methodologies will independently filter the references, first through reading the title and abstract and then through review of full text. Every excluded study will be recorded reasons according to inclusion criteria. A third researcher (ZW) will resolve disagreements in the whole study selection process.

Data extraction

We will extract data on the characteristics of PROMs as well as the study characteristics. As shown in table 1, the characteristics of PROMs will include title, target population, mode of administration (self-report, interview-based or proxy report), recall period, subscale and item number, range of scores, original language, theory, and available translations. We also have designed table 2 to present the study characteristics, which include author (year), PROM, country, PROM language, study design, sample size and participants, and year of development/validation. Two researchers (MJ and SH) will independently extract data and information using tables 1 and 2. A third researcher (ZW) will be invited to discuss any inconsistency and disagreement.

Assessment of methodological quality

Two researchers (MJ and SH) will independently assess the methodological quality of each involved study using the COSMIN risk of bias checklist.[19] A third researcher (ZW) will be invited to discuss any inconsistency and disagreement. The COSMIN risk of bias checklist has 10 domains and 116 items. It is used to assess methodological quality in terms of PROM development, content validity, structural validity, internal consistency, cross-cultural validity/measurement invariance, reliability, measurement error, criterion validity, hypotheses testing for construct validity and responsiveness. Each item has five options, namely 'very good', 'adequate', 'doubtful', 'inadequate' and 'not applicable'. The 'worst score counts' principle is used to determine the overall quality of relative domains.

Summarising the quality of PROMs

Two researchers (MJ and SH) will independently summarise the quality of psychometric properties for each PROM according to the COSMIN criteria.[20] A third researcher (ZW) will be invited to discuss any inconsistency and disagreement. The COSMIN criteria rates the psychometric properties of PROMs, including structural validity, internal consistency, cross-cultural validity/measurement invariance, reliability, measurement error, criterion validity, hypotheses testing for construct validity and responsiveness, as sufficient (+), insufficient (−) or indeterminate (?). The specific criteria for good measurement properties are shown in table 3. For instance, reliability will be rated as '+' if the ICC or weighted kappa is ≥0.70, '−' if the ICC or weighted kappa is <0.70, and '?' if the ICC or weighted kappa is not reported. We will first rate each single study on psychometric properties. Then we will synthesise the results and come to an overall conclusion on the quality of the PROM's psychometric properties as a whole according to the specific situation. If the ratings for each measure are consistent, the results from different studies on one psychometric property will be qualitatively summarised or pooled through

Table 2 Study characteristics

Author (year)	PROM	Country	PROM language	Study design	Sample size and participants	Year of development/validation

PROM, patient-reported outcome measure.

Table 3 Criteria for good measurement properties

Measurement property	Rating	Criteria
Structural validity	+	*CTT* CFA: CFI or TLI or comparable measure >0.95 *OR* RMSEA <0.06 *OR* SRMR <0.08*. *IRT/Rasch* No violation of unidimensionality†: CFI or TLI or comparable measure >0.95 *OR* RMSEA <0.06 *OR* SRMR <0.08. *AND* no violation of local independence: residual correlations among the items after controlling for the dominant factor <0.20 *OR* Q3 <0.37. *AND* no violation of monotonicity: adequate looking graphs *OR* item scalability >0.30. *AND* adequate model fit. IRT: χ^2 >0.001. Rasch: infit and outfit mean squares ≥0.5 and ≤1.5 *OR* Z-standardised values >–2 and <2.
	?	CTT: not all information for '+' reported. IRT/Rasch: model fit not reported.
	–	Criteria for '+' not met.
Internal consistency	+	At least low evidence‡ for sufficient structural validity§ *AND* Cronbach's alpha ≥0.70 for each unidimensional scale or subscale¶.
	?	Criteria for 'At least low evidence‡ for sufficient structural validity§' not met.
	–	At least low evidence‡ for sufficient structural validity§ *AND* Cronbach's alpha <0.70 for each unidimensional scale or subscale¶.
Reliability	+	ICC or weighted kappa ≥0.70.
	?	ICC or weighted kappa not reported.
	–	ICC or weighted kappa <0.70.
Measurement error	+	SDC or LoA <MIC§.
	?	MIC not defined.
	–	SDC or LoA >MIC§.
Hypotheses testing for construct validity	+	The result is in accordance with the hypothesis**.
	?	No hypothesis defined (by the review team).
	–	The result is not in accordance with the hypothesis**.
Cross-cultural validity/ measurement invariance	+	No important differences found between group factors (such as age, gender, language) in multiple group factor analysis *OR* no important DIF for group factors (McFadden's R^2 <0.02).
	?	No multiple group factor analysis *OR* DIF analysis performed.
	–	Important differences between group factors *OR* DIF were found.
Criterion validity	+	Correlation with gold standard ≥0.70 *OR* AUC ≥0.70.
	?	Not all information for '+' reported.
	–	Correlation with gold standard <0.70 *OR* AUC <0.70.
Responsiveness	+	The result is in accordance with the hypothesis** *OR* AUC ≥0.70.

Continued

Table 3 Continued

Measurement property	Rating	Criteria
	?	No hypothesis defined (by the review team).
	–	The result is not in accordance with the hypothesis** *OR* AUC <0.70.

From Prinsen *et al*.[11]
'+', sufficient; '–', insufficient; '?', indeterminate.
*To rate the quality of the summary score, the factor structures should be equal across studies.
†Unidimensionality refers to a factor analysis per subscale, while structural validity refers to a factor analysis of a (multidimensional) PROM.
‡As defined by grading the evidence according to the GRADE approach.
§This evidence may come from different studies.
¶The criteria 'Cronbach's alpha <0.95' was deleted as this is relevant in the development phase of a PROM and not when evaluating an existing PROM.
**The results of all studies should be taken together and it should then be decided if 75% of the results are in accordance with the hypotheses.
AUC, area under the curve; CFA, confirmatory factor analysis; CFI, comparative fit index; CTT, classical test theory; DIF, differential item functioning; GRADE, Grading of Recommendations, Assessment, Development and Evaluation; ICC, intraclass correlation coefficient; IRT, item response theory; LoA, limits of agreement; MIC, minimal important change; PROM, patient-reported outcome measure; RMSEA, root mean square error of approximation; SDC, smallest detectable change; SRMR, standardised root mean residuals; TLI, Tucker-Lewis index.

Table 4 Amendment notes

Date	Section	Original protocol	Revised protocol	Rationale

meta-analysis and finally be rated as '+' or '–'. A meta-analysis will be conducted according to the availability of quantitative data of psychometric properties. For instance, ICC between different studies that assess the same PROM may be pooled by calculating the weighted means (according to the sample size in each study) and 95% CI. We will not apply systematic review data management software considering we may not have much quantitative data synthesis work. On the other hand, if the ratings are inconsistent, we will explore explanations for the inconsistency between studies according to the context of different studies. Ratings will be provided in subgroups, such as different languages of the PROM, different participant characteristics (such as older adults living with or without any disease), etc, if the explanations are reasonable. If there is not enough information to explain the inconsistency, the overall rating will be rated as '?'.

Grading the quality of evidence

Four researchers (MJ, SH, ML and JZ) will independently grade the quality of evidence, that is, the confidence that the pooled or summarised result is trustworthy, according to the modified Grading of Recommendations, Assessment, Development and Evaluation system.[21] A fifth researcher (ZW) will be invited to discuss any inconsistency and disagreement. Using four factors to determine the quality of evidence (risk of bias, inconsistency, indirectness and imprecision), each psychometric property of PROM is graded as high, moderate, low or very low evidence.

Plans in case of possible amendments

If we need to amend this protocol, we will give the date of each amendment, describe the change and provide the rationale in this section. Changes will not be incorporated into the protocol. A researcher (SH) will ultimately be responsible for reporting any necessary amendments in a tabular format (table 4).

Contributors Study design: MJ. Search strategy, study selection, data extraction, assessment of methodological quality, data analysis: MJ, SH, ZW. Grading the quality of evidence: MJ, SH, ML, JZ, ZW. Supervision: ZW. Writing - original draft preparation: SH. Writing - review and editing: ML, JZ, ZW. Funding acquisition: ZW. All authors have read and agreed to the published version of the manuscript.

Funding This study was funded by the major projects of the National Social Science Foundation (21ZDA107). This funding provides scholarship for coauthors. The funder will have no input on protocol development, review conduct, data analysis and interpretation, and dissemination of the final report.

Competing interests None declared.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been

ORCID iD

Shuyu Han http://orcid.org/0000-0002-2506-6888

REFERENCES

1 Steenman M, Lande G. Cardiac aging and heart disease in humans. *Biophys Rev* 2017;9:131–7.
2 United Nations. World population prospects 2019,, 2019. Available: https://population.un.org/wpp/Publications/Files/WPP2019_Highlights.pdf [Accessed 16 Nov 2021].
3 National Bureau of Statistics.. Report of China's Seveth Population Census, 2021.. Available: http://www.stats.gov.cn/tjsj/tjgb/rkpcgb/ [Accessed 16 Nov 2021].
4 China Development Research Foundation. China Development Report 2020: Development Trends and Policies of China's Population Aging, 2020.. Available: https://www.cdrf.org.cn/laolinghua/index.htm [Accessed 16 Nov 2021].
5 United Nations. Ageing, 2020.. Available: https://www.un.org/en/global-issues/ageing [Accessed 16 Nov 2021].
6 World Health Organization. Active ageing : a policy framework, 2002.. Available: https://apps.who.int/iris/handle/10665/67215
7 Dionyssiotis Y, Ageing A. Active ageing. *J Frailty Sarcopenia Falls* 2018;03:125–7.
8 Haque MN. Active ageing level of older persons: regional comparison in Thailand. *J Aging Res* 2016;2016:1–9.
9 MdN H, Soonthorndhada K, Hunchangsith P. Active ageing level in Thailand: a comparison between female and male elderly. *Journal of Health Research* 2016;30:99–107.
10 Rantanen T, Portegijs E, Kokko K, *et al*. Developing an assessment method of active aging: University of Jyvaskyla active aging scale. *J Aging Health* 2019;31:1002–24.
11 Prinsen CAC, Mokkink LB, Bouter LM, *et al*. COSMIN guideline for systematic reviews of patient-reported outcome measures. *Qual Life Res* 2018;27:1147–57.
12 Shamseer L, Moher D, Clarke M, *et al*. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. *BMJ* 2015;349:g7647.
13 Page MJ, McKenzie JE, Bossuyt PM, *et al*. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 2021;372:n71.
14 National Institute for Health Research (NIHR).. Prospero International prospective register of systematic reviews, 2021.. Available: https://www.crd.york.ac.uk/prospero/ [Accessed 16 Nov 2021].
15 McGowan J, Sampson M, Salzwedel DM, *et al*. PRESS Peer Review of Electronic Search Strategies: 2015 Guideline Statement. *J Clin Epidemiol* 2016;75:40–6.
16 Consensus-dased standards for the selection of health measurement instructures (COSMIN).. Guideline for systematic reviews of outcome measurement instruments, 2021.. Available: https://www.cosmin.nl/tools/guideline-conducting-systematic-review-outcome-measures/ [Accessed 16 Nov 2021].
17 Udo D. Active ageing: a concept analysis. *Caribbean Journal of Nursing* 2016;3:59–79. Available: https://www.researchgate.net/publication/303784510_Active_Ageing_A_Concept_Analysis
18 Kenbubpha K, Higgins I, Chan SW-C, *et al*. Promoting active ageing in older people with mental disorders living in the community: an integrative review. *Int J Nurs Pract* 2018;24:e12624.
19 Mokkink LB, de Vet HCW, Prinsen CAC, *et al*. COSMIN risk of bias checklist for systematic reviews of patient-reported outcome measures. *Qual Life Res* 2018;27:1171–9.
20 Prinsen CAC, Vohra S, Rose MR, *et al*. How to select outcome measurement instruments for outcomes included in a "Core Outcome Set" - a practical guideline. *Trials* 2016;17:449.
21 Guyatt GH, Oxman AD, Vist GE, *et al*. Grade: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 2008;336:924–6.