

Review Article

Cite this article: Jo M, Park M, Hwang H, Chung H (2025) Measurement properties of patient-reported outcome measures for advance care planning in older people: A COSMIN systematic review. *Palliative and Supportive Care* 23, e53, 1–10. <https://doi.org/10.1017/S1478951524002062>

Received: 7 June 2024

Revised: 20 September 2024

Accepted: 27 November 2024


Keywords:

Advance care planning;
patient reported outcome measure; aged;
psychometric; review

Corresponding author: Mihyun Park;

Email: mvc90@caholic.ac.kr

Measurement properties of patient-reported outcome measures for advance care planning in older people: A COSMIN systematic review

Minjeong Jo, R.N., PH.D.¹, Mihyun Park, R.N., PH.D.^{1,2} , Hye-lyung Hwang, E.D.M.² and Heejin Chung, R.N., M.S.N.³

¹College of Nursing, The Catholic University of Korea, Republic of Korea; ²Research Institute for Hospice/Palliative Care, The Catholic University of Korea, Republic of Korea and ³Graduate School, The Catholic University of Korea, Republic of Korea

Abstract

Objectives. This review provides an overview of patient-reported outcome measure (PROMs) utilized to assess the impact of advance care planning (ACP) among older adults and evaluates their psychometric properties.

Methods. The inclusion criteria were as follows: (1) studies that targeted older adults; (2) studies using of any type of measurement tools that measure patient-reported ACP program outcomes; and (3) studies published in English or Korean. Following PRISMA guidelines, a systematic review was conducted, encompassing electronic searches across 5 databases including PubMed, EMBASE, Cochrane Library, CINHAL, and PsycINFO and manual searches of umbrella reviews on ACP interventions. General characteristics of the selected measures were extracted, and their methodological quality was assessed using the COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) checklist.

Results. Out of 19,503 studies initially identified, 74 met the inclusion criteria, reporting on a total of 202 measures. These measures were categorized into 4 domains reflecting the targets of ACP interventions: process ($n = 56$), action ($n = 18$), process and action ($n = 16$), quality of care ($n = 63$), and health status ($n = 49$). Despite the breadth of measures identified, none fully met all recommended psychometric properties outlined in the checklist.

Significance of results. While this review aids in the selection of measures for both practical and research purposes, it underscores the necessity for further validation of PROMs in assessing ACP outcomes in older adults, advocating for rigorous psychometric evaluations and adherence to standards like the COSMIN checklist to ensure reliable and valid data. It suggests the need for shortened versions and researcher assistance to address the challenges older adults face with self-reported PROMs and improve participation rates.

Introduction

The aging population is growing worldwide (Li et al. 2019). This phenomenon has been expedited by the development of advanced technologies and breakthroughs in medicine that have prolonged the lifespans of individuals, and it represents an important issue because older adults are anticipated to grapple with protracted periods of managing chronic and age-associated disorders (Woods et al. 2022; World Health Organization 2020). Although healthcare systems strive to support older adults for a long and healthy life, they will inevitably face frailties at some point and require end-of-life care. Nevertheless, a considerable number of elderly individuals nearing the end of life find themselves incapable of articulating their preferences due to incapacitation at the critical decision-making juncture. Family members frequently assume the responsibility of decision-making on behalf of their ailing loved ones in these cases, thus resulting in heightened distress and interpersonal conflicts as they endeavor to navigate the complex situations of selecting optimal choices for the individual (Institute of Medicine 2014). Consequently, many older adults who did not officially express their own decisions regarding their end-of-life in advance experience prolonged death while relying on life-sustaining treatments in hospitals, which are associated with high medical costs at the end-of-life and which deteriorates the quality of both end-of-life and death (The Lancet Respiratory Medicine 2016).

Therefore, it is imperative for older individuals to proactively plan for end-of-life care before they lose the ability to communicate their preferences. The Institute of Medicine (2014) has recommended that individuals express their end-of-life care preferences while they are still in good health and cognitively sound. Since the late 20th century, researchers and politicians in Western

© The Author(s), 2025. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives licence (<http://creativecommons.org/licenses/by-nc-nd/4.0>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided that no alterations are made and the original article is properly cited. The written permission of Cambridge University Press must be obtained prior to any commercial use and/or adaptation of the article.

countries have developed specific programs, interventions, or policies to facilitate end-of-life communication among older adults (Park et al. 2021; Thomas et al. 2018). These are collectively described as advance care planning (ACP).

In the initial phase, ACP was conceptualized as the documentation of advance directives based on provided information (Carr and Luth 2017). While it had until recently been used without a commonly shared definition, ACP has undergone multifaceted developments over time. The Institute of Medicine (2014) attempted to comprehensively define ACP based on its experience in America over the prior decade. Scholars in Europe (Rietjens et al. 2017) and North America (Sudore et al. 2017b) used Delphi consensus processes in attempts to comprehensively define ACP. Those efforts uniformly delineated ACP as a process aimed at identifying one's personal values, goals, and preferences regarding future medical treatments. This process involves engaging in discussions with both family members and healthcare providers, maintaining records of one's end-of-life care preferences, and appointing a proxy decision-maker. Finally, those defining works were linked to identifying ACP outcomes (McMahan et al. 2021; Sudore et al. 2018). Based on the definition of ACP, the outcomes reported in the previous studies were categorized along 4 domains, including *process*, *action*, *quality of care*, and *health status/utilization* (McMahan et al. 2021; Sudore et al. 2018). ACP interventions primarily target a patient's behavior changes like attitude, decision-making, and preferences, potentially leading to direct outcomes in the process or action domain. Outcomes related to the quality of care (QOC) may be associated with patients' evaluation of ACP as care itself. Finally, health status outcomes might reflect the indirect, secondary effects of ACP interventions on a patient's mental well-being or life expectancy.

The growing elderly population underscores the importance of effective ACP interventions for older adults. It is necessary to have a deep understanding of older adults' experiences with ACP as a process to accurately evaluate these interventions. Patient-reported outcome measures (PROMs) offer invaluable insights into the impact that ACP interventions have on this population. While numerous studies have used various PROMs to assess ACP intervention effectiveness, there is a crucial gap regarding the appropriate selection and evaluation of these measures for the specific context of older adults. Prior research may not have adequately considered factors such as the reliability and validity of these PROMs when applied to the elderly population. To address this gap, it is necessary to conduct a systematic review of PROMs specifically focused on ACP for older adults. Such a review could leverage established tools like the CONsensus-based Standards for the selection of health Measurement INSTRuments (COSMIN) to identify high-quality PROMs suitable for both research and clinical settings (Prinsen et al. 2018). Therefore, the purpose of the current review was to provide an overview of the PROMs used to measure the effects of ACP among older adults and to evaluate the psychometric properties of these PROMs using the COSMIN methodology.

Methods

Design

A systematic review of the measurement properties of ACP instruments in older adults was conducted while following the COSMIN methodology for systematic reviews of PROMs (Mokkink et al. 2018; Prinsen et al. 2018).

Search strategy

The search strategy for this study comprised 2 main steps aimed at efficiently managing the extensive body of literature on ACP programs and assessing various PROMs.

Step 1: Identification of studies via umbrella reviews (~2018)

To efficiently manage the extensive body of research on ACP interventions and the various PROMs associated with these interventions, we employed a 2-step search strategy focused on umbrella reviews. Given the large volume of existing studies on the effectiveness of ACP interventions, we initially prioritized umbrella reviews, which synthesize findings from multiple systematic reviews and meta-analyses, to provide a broad and comprehensive spectrum of evidence (Fusar-Poli and Radua 2018). Specifically, we leveraged our previously published umbrella review (Park et al. 2021), which synthesized reviews on ACP interventions for older adults in community-based settings published up to 2019, serving as a foundational reference point. To ensure a current and comprehensive capture of evidence, we expanded our search to include additional umbrella review published up to 2022 that addressed similar topics. We conducted searches across several databases, including PubMed, Cochrane Library, and CINHAL. The search strategy used keywords such as ["Advance Care Planning" OR "ACP"] AND ["Umbrella Review"] AND ["Older Adults" OR "Elderly"], covering studies published up to February 2022. During the initial screening, titles and abstracts were reviewed with inclusion criteria that required systematic reviews specifically addressing ACP interventions, assessing the quality of included studies using a standardized tool, and being published in English or Korean within peer-reviewed journals. A full-text assessment then evaluated the comprehensiveness of the included reviews, their methodological rigor, and their alignment with our study objectives, particularly concerning PROMs in ACP programs. Based on these criteria, 3 umbrella reviews were selected: Jimenez et al. (2018), (Park et al. 2021), and Wendrich-van Dael et al. (2020), as they collectively provided a broad overview of the effectiveness of ACP interventions and the variety of PROMs used. From the selected umbrella review, we included only studies with original data collection and analysis. This approach streamlined the identification and synthesis of relevant studies, allowing us to compile a focused list of research and measures used to evaluate the outcome of ACP interventions in older adult population (Fig. 1).

Step 2: Identification of studies via databases (2018~2022)

We conducted a literature search using electronic databases including PubMed, EMBASE, Cochrane Library, CINHAL, and PsycINFO. These databases were searched for published studies from 1st January 2018 to 28th February 2022. We set the starting point as 2018 because the umbrella reviews that were selected in the first step included studies until 2018. Initially, we established keywords based on prior studies. With the assistance of a professional librarian, we searched each database using the keywords and medical subject headings (MESH) terms related to these keywords. The keywords and MESH terms were combined together with the "OR" and "AND" Boolean operators (Supplementary 1). Given the potential availability of additional relevant studies, we reviewed the references of all the included studies.

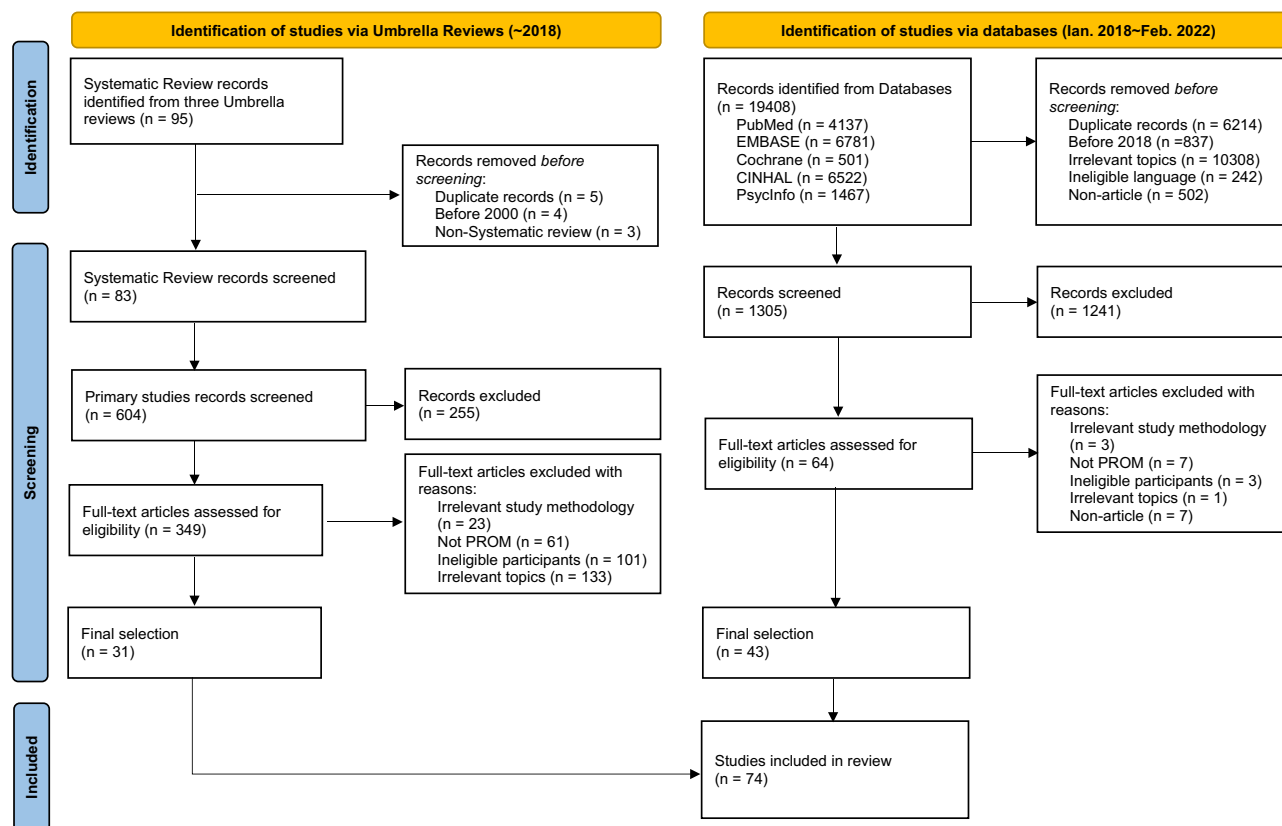


Figure 1. PRISMA flowchart for study screening and selection process.

Inclusion and exclusion criteria

The inclusion criteria were as follows: (1) studies that targeted older adults (65 years or older); (2) studies using of any type of measurement tools that measure patient-reported ACP program outcomes; and (3) studies published in English or Korean. In our study, ACP was defined by following Sudore et al. (2017a) who stated that “ACP is a process that supporting adults at any age or stage of health in understanding and sharing their personal values, life goals, and preferences regarding future medical care.” Based on this definition, we have considered ACP programs to be interventions focused on individualized support and preparation for communication and in-the moment decision-making. Programs that involve discussions between healthcare professionals or researchers and individuals regarding personal goal of care, values, etc., as well as the provision of information or education related to end-of-life treatment or medical decision-making, have been considered to be ACP programs. However, the provision of palliative care received by terminally ill patients in hospital or at home for pain management and comfort enhancement has not been regarded as part of ACP programs. The exclusion criteria were as follows: (1) nonexperimental studies; (2) studies that only provide outcomes of family caregivers or surrogates; (3) studies that include dementia patients; (4) studies for which the full text is unavailable; and (5) studies published in non-peer-reviewed journals.

Study screening and selection

We imported the articles (citations, abstracts, and full texts) identified from umbrella reviews and relevant databases into

EndNote X9, a bibliographic manager. After removing duplicates, 4 researchers individually conducted an initial review of the titles and abstracts of the remaining articles for their relevance to the review topic. These 4 researchers were divided into 2 groups: Two of them reviewed articles that were identified through umbrella review search, while the other 2 reviewed articles that had been identified from databases. Studies that potentially or completely met the inclusion criteria were kept, and their full-texts were reviewed to decide whether to keep the record in the review. In cases of confusion or the need for exclusion, comments were noted separately. Upon conclusion of the initial review, researchers from each group proceeded to conduct additional reviews of articles from the opposing group that had not been previously examined. Discrepancies regarding article eligibility were resolved through discussion among researchers to achieve consensus. Subsequently, all researchers reviewed the ultimately included studies to ensure agreement.

Data extraction

Four researchers individually extracted information using 3 standardized information forms. The following information was extracted from each included study: (1) characteristics of included studies including author, year, type of ACP intervention, patient-reported outcomes category, name of PROMs, and presence of intervention effects (Supplementary 2); (2) PROMs characteristics including name, target population (size, age, and setting), number of items, subscales, response options, mode of administration (self-reported/researcher-reported), and original language (Supplementary 3); and (3) psychometric properties of PROMs

including content validity, structural validity, internal consistency, cross-cultural validity, reliability, measurement error, criterion validity, hypotheses testing for construct validity, and responsiveness (Supplementary 3; Table 1). The ACP intervention outcome category in supplementary 2 was based on the standardized outcome framework of ACP by McMahan et al. (2021): process, action, QOC, and health status.

To ensure accurate understanding of each included measure, the original development articles were identified using the reference information provided. Any discrepancies arising during information extraction were then discussed among researchers to reach consensus. Finally, all researchers reviewed the final extracted information from the included studies to confirm agreement.

Appraisal of psychometric properties

Psychometric validation data specific to older people who received an ACP program were searched for each measure by reviewing all studies citing the original development article. The psychometric properties of the included measure, a tool measuring patient-reported ACP program outcomes, were assessed using the COSMIN checklist (Prinsen et al. 2018). Each measure was evaluated against the following COSMIN criteria: content validity, structural validity, internal consistency, cross-cultural validity, reliability, measurement error, criterion validity, hypotheses testing for construct validity, and responsiveness. The qualities of each measure were established according to criteria developed by Prinsen et al. (2018). However, the research team modified the rating system of measurement properties. According to this framework, measurement properties were rated as “Yes” if conducted, “Yes(+)” if provided information is appropriate, and “No” if not conducted or no information was available. For example, internal consistency is rated as “Yes(+)” where Cronbach's alpha is equal to or greater than 0.70; criterion validity is rated as “Yes(+)” where the correlation with the gold standard is equal to or greater 0.70 (Mokkink et al. 2018; Prinsen et al. 2018). In our study, ratings of the psychometric properties of each measure were confirmed by 1 researcher while 4 members of the review team independently rated all measurement properties. Any discrepancies were resolved through team discussion until reaching a consensus.

Results

Literature search

In total, 19,503 articles (95 from the umbrella review and 19,408 from the database) were initially identified. After removing duplicates, completing preliminary screening, and reviewing full texts, we ultimately included 74 articles (31 from the umbrella review and 43 from the database). The study search and selection process are presented in Figure 1.

Study characteristics and ACP intervention outcomes

Supplementary 2 presents the description of included ACP intervention studies. The years of publication ranged from 2000 to 2022. Of 74 studies, ACP interventions were categorized into 3 types: 23 were information-based, aiming to influence patient's thoughts though the provision of ACP information resources; 20 were conversation-based, providing opportunities to discuss the patient's goal of care and value; and the remaining 31 utilized

a combination of different approaches, including education and support for advance directives. In total, 202 PROM outcomes were identified from 74 ACP intervention studies. Approximately 80% ($n = 165$) of the outcomes indicated that the ACP intervention had positive or negative effects, while the remaining 20% did not provide specific information about the outcomes. Among the tools measuring 202 outcomes, 55 (27.2%) were tools developed by researchers either previously or during the course of their studies that were specifically tailored to their research. The remaining 147 (72.4%) tools were previously developed by other researchers.

The identified 202 outcomes were categorized into 4 domains: process, action, QOC, and health status (McMahan et al. 2021; Sudore et al. 2018). The outcomes of process category specify how an effect of ACP intervention occurs. In our study, the included outcome variables were ACP behavior's change, self-efficacy, attitude, knowledge (including Cardiopulmonary resuscitation [CPR] and Advance Directive [AD]), autonomy, and readiness (intention). ACP-specific action outcomes indicate an individual's completion of specific components of ACP such as discussing or documenting goal of care or treatment preferences. Some outcomes were measured in conjunction with whether actions were taken, as well as ACP behavior such as attitudes and self-efficacy. Therefore, we added the category of “process/action” to the existing 4 domains. The outcomes of QOC category address the quality of ACP such as perceived satisfaction with ACP, communication, and decision making, and congruence between patients and surrogates. Meanwhile, the outcomes of health status category were the impact of ACP on health outcomes, such as mental status (anxiety, depression, and stress) or quality of life. Further among the 202 outcomes of ACP interventions, those corresponding to the QOC category were the largest group ($n = 63$), followed by process ($n = 56$), health status ($n = 49$), action ($n = 18$), and process and action ($n = 16$). The outcomes in the categories of QOC, process, and health status accounted for more than 80% of identified outcomes of ACP intervention.

Characteristics of PROMs

To assess the properties of the PROMs of ACP interventions, a selection process was conducted among the 202 tools measuring outcomes. For the selection, tools measuring outcomes classified under the health status domain were excluded ($n = 49$). Based on the definition of ACP (Sudore et al. 2017b), as a process supporting and facilitating decision-making that reflects an individual's values and treatment preferences, variables related to the health status category were deemed to be unlikely to be directly influenced by ACP intervention. Consequently, they may not be the most suitable outcome indicators for evaluating the specific effects of ACP interventions. Moreover, tools were excluded if their names were provided but their measurement properties were not reported, or if it was not possible to find references regarding the tools ($n = 29$). After removing duplicate tools ($n = 38$), 86 measures were ultimately selected for psychometric appraisal of PROMs in this study.

The characteristics of the 86 PROMs are presented in Supplementary 3. Approximately 71% of the PROMs (61 out of 86) measured the process and action categories either individually or in combination. The PROMs that have been used in studies either targeting older people or involving more than half of the participants as the elderly were analyzed. The range of years in which the tools were developed spanned from 1979 to 2021. This review identified

Table 1. Summary of psychometric properties of PROMs scoring over 4 points on the COSMIN checklist

Category	Measurement [No. for OA]*	Version (Language used in studies)	Content validity	Structural validity	Internal consistency	Cross- cultural validity	Reliability	Measurement error	Criterion validity	Construct validity	Responsiveness	COSMIN score	No. for IA*
P	ACP knowledge [IA]	8-item version (Taiwanese)	Y	N	Y(+)	N	N	N	N	Y(+)	Y	4	33
P	ADAS [13, 64]	16-item version (Modified Korean)	Y	N	Y(+)	N	N	N	N	Y(+)	Y	4	74
P	ADAS [13, 43]	16-item version (Modified Chinese)	Y	N	Y(+)	N	N	N	N	Y(+)	Y(+)	4	18, 47
P	ADAS [13]	12-item version (Modified Spanish)	Y	Y	Y(+)	N	N	N	N	Y	N	4	16
P	Behavioral intention [IA]	4-item version (Taiwanese)	Y	N	Y(+)	N	N	N	N	Y(+)	Y	4	33
P	CPR-related knowledge	5-item version (Taiwanese)	Y	N	Y(+)	N	N	N	N	Y	Y(+)	4	8
P	Decisional bal- ance, beliefs, processes of change survey in ACP [53]	34-item version (English)	Y	Y	Y(+)	N	N	N	N	Y	Y	5	60
P	Navigational barriers to ACP communication [IA]	6-item version (Spanish/English)	N	Y	Y(+)	N	N	N	N	Y	Y	4	41
P	ODCNF scale [26]	8-item version (Dutch)	Y	Y	Y(+)	Y	Y	N	N	Y	Y	7	25
P	Readiness to engage in ACP discussion	9-item version (Spanish/English)	N	Y	Y(+)	N	N	N	N	Y	Y	4	41
P/A	ACP Engagement Survey [22]	49-item version (English)	Y	N	Y(+)	N	Y	N	Y	Y	N	5	19, 60
P/A	ACP Engagement Survey [23]	82-item version (English)	Y	Y	Y(+)	N	Y	N	Y	Y	Y(+)	7	34, 59
		55-item version (English)	Y	Y	Y(+)	N	Y	N	Y	Y	Y(+)	7	22
		34-item version (English)	Y	Y	Y(+)	N	Y	N	Y	Y	Y(+)	7	64
		15-item version (English)	Y	Y	Y(+)	N	Y	N	Y	Y	Y(+)	7	23, 72, 73

(Continued)

Table 1. (Continued.)

Category	Measurement [No. for OA]*	Version (Language used in studies)	Content validity	Structural validity	Internal consistency	Cross- cultural validity	Reliability	Measurement error	Criterion validity	Construct validity	Responsiveness	COSMIN score	No. for IA*
		9-item version (English)	Y	Y	Y(+)	N	Y	N	Y	Y	Y(+)	7	36
		4-item version (English)	Y	Y	Y(+)	N	Y	N	Y	Y	Y(+)	7	35, 45, 49, 63
P/A	ACP Engagement Survey [23, 38]	15-item version (Japanese)	Y	Y	Y(+)	N	Y	N	Y	Y	Y(+)	7	43
QOC	CANHELP [35, 36]	38-item version (English/French)	Y	Y	Y(+)	Y	N	N	Y	Y(+)	Y(+)	7	40
QOC	CSQ-8 [19]	8-item version (English)	Y	Y	Y(+)	Y	N	N	N	Y(+)	N	5	17
QOC	DCS [6]	16-item version (English)	Y	N	Y(+)	Y	Y	N	N	Y(+)	Y(+)	6	4, 11, 12, 20, 32, 38, 55, 56, 57, 58, 63, 65, 66, 71
QOC	DCS [6, 24]	16-item version (Korean)	Y	Y	Y(+)	N	Y	N	N	Y(+)	N	5	24
QOC	DCS; SURE test [8, 34]	4-item version (English/French)	Y	Y	Y(+)	Y	N	N	Y	Y(+)	Y(+)	7	6, 9, 40
QOC	EORTC IN- PATSAT32 [4]	32-item version (English)	Y	N	Y(+)	N	Y	N	Y	Y(+)	Y(+)	6	3, 31
QOC	FACIT-TS-G[10]	26-item version (English)	Y	Y	Y(+)	N	Y	N	Y	Y(+)	N	6	12
QOC	LSPQ [9]	36-item version (English)	Y	N	Y(+)	N	N	N	N	Y(+)	Y(+)	4	6, 10, 24, 27, 28
QOC	SHARED [37]	10-item version (English)	Y	Y	Y(+)	Y	N	N	N	Y	Y	6	42
QOC	Quality of com- munication [3]	13-item version (English)	Y	Y	Y(+)	Y	N	N	Y(+)	Y(+)	Y(+)	7	2, 4, 14, 21, 56, 57
QOC	Quality of patient-clinician communication about end-of-life care [12]	4-item version (English)	Y	N	Y(+)	N	Y	N	N	Y(+)	N	4	14

Notes: ADAS = Advance Directive Attitude Survey; CANHELP = Canadian Healthcare Evaluation Project questionnaire; CSQ-8 = Client Satisfaction Questionnaire; DCS = Decisional Conflict Scale; EORTC IN-PATSAT32 = European Organization for Research and Treatment of Cancer IN-PATIENT Satisfaction with care questionnaire 32; FACIT-TS-G = Functional Assessment of Chronic Illness Therapy Treatment Satisfaction General; IA = Included Article; LSPQ = Life-Support Preferences Questionnaire; ODCNF = Openness to Discuss Cancer in the Nuclear Family scale; OA = Original Article; P = Process; PA = Process and Action; QOC = quality of care; SHARED = patient experience of SHARED decision making questionnaire. Y = measurement properties were rated as "y" if conducted, Yes(+) = if the information was appropriate, and N = No, if not conducted or information was not provided.

*The lists (number) of original and included articles are provided in Supplementary 2.

15 tools that were used repeatedly across the included articles, with usage frequencies ranging from 2 to 14 times. The original languages of 86 PROMs were English ($n = 61$), Taiwanese ($n = 12$), Japanese ($n = 3$), Korean ($n = 3$), Chinese ($n = 2$), and Dutch ($n = 1$). Four PROMs were developed not only in English but also in French (Heyland et al. 2013; Légaré et al. 2010) and Chinese (Hinderer and Lee 2019; Lee et al. 2015). Most of the PROMs were self-administered assessment tools ($n = 67$ out of 86). The selected PROMs were used for older people from a diverse setting, including both inpatients and outpatients from hospitals and clinics, as well as from community centers. The participants had an age distribution ranging from an average of 39–87.5 years. The PROMs had a diverse range of item numbers, ranging from 1 to 82 questions. Response scales included options such as multiple choices, yes/no, true/false, agree/disagree, numeric rating scale, and Likert scales (4, 5, 6, and 11 points). Moreover, a variety of response scales were used to indicate preferences, such as who the decision-maker wishes to be, the extent of medical treatment desired, the amount of desired information, and the level of preparedness for ACP communication.

Appraisal of PROMs psychometric properties

The 9 psychometric properties (content validity, structural validity, internal consistency, cross-cultural validity, reliability, measurement error, criterion validity, hypotheses testing for construct validity, and responsiveness) of the selected 86 PROMs were assessed according to the COSMIN criteria. Supplementary 3 presents an overview of the psychometric properties of the 86 PROMs. It was found that approximately 78% ($n = 67$) of the 86 PROMs did not meet half or more of the 9 COSMIN criteria for measurement properties; 14 measures had no psychometric validation data documented. Therefore, we selected 29 PROMs that met 4 or more of the 9 COSMIN criteria, and their properties are specified in Table 1. The most commonly addressed COSMIN criteria were evidence for internal consistency ($n = 29$ measures), hypotheses testing construct validity ($n = 29$ measures), and content validity ($n = 27$ measures). By contrast, there was limited evidence regarding measurement error, cross-cultural validity, and reliability. Out of the 29 tools, none of them presented evidence of measurement error. Data on cross-cultural validity was unavailable for 22 tools, and data on reliability was absent for 15 tools. Among the 29 PROMs, Decisional Conflict Scale (DCS) were originally in English (O'Connor 1995) and had translated versions in languages such as French (Ferron Parayre et al. 2014; Légaré et al. 2010) and Korean (Kim et al. 2017). Advance Directive Attitude Survey (ADAS) were also originally in English (Nolan and Bruder 1997) and had translated versions in languages Spanish (Nolan and Bruder 1997), Chinese (Hinderer and Lee 2019), and Korean (Joung 2012). DCS had evidence of cross-cultural validity for the English and French versions, whereas there was no evidence for ADAS in this regard. Evidence for cross-cultural validity was present for 7 (Bekker et al. 2012; Engelberg et al. 2006; Ferron Parayre et al. 2014; Fried et al. 2010; Heyland et al. 2010, 2013; Larsen et al. 1979; Légaré et al. 2010; O'Connor 1995) out of 29 PROMs. None had evidence for all 9 psychometric properties, while 11 PROMs had evidence for 7 of the 9 properties. Eleven PROMs include an ACP Engagement Survey (with 4, 9, 15, 34, 55, and 82 items) (Sudore et al. 2017a), the Japanese version of ACP Engagement Survey (with 15 items) (Okada et al. 2021), Canadian Health Care Evaluation Project Questionnaire (CANHELP) (Heyland et al.

2010, 2013), Openness to Discuss Cancer in the Nuclear Family scale (ODCNF) (Mesters et al. 1997), quality of communication (QOC) (Engelberg et al. 2006), and SURE test (Légaré et al. 2010).

Discussion

This review study has provided an overview of the PROMs used to assess the impact of ACP interventions on older adults. While the identified PROMs categorized into 4 domains aligned with previous research (McMahan et al. 2021; Sudore et al. 2018), a wide range of instruments were employed, thus suggesting a lack of consensus on the most suitable PROM for this area and the existence of potential challenges in integrating ACP outcomes. Several factors may contribute to this heterogeneity: First, despite the fact that studies on ACP have been emerging since the 1990s, a consistent definition of the concept only recently gained consensus among scholars (Rietjens et al. 2017; Sudore et al. 2017b). Second, some studies utilized unspecified measures for ACP, employing generic PROMs instead. Notably, most health status/utilization outcomes relied on generic PROMs, causing the present review to exclude them from further analyses of psychometric properties to focus on ACP-specific instruments.

This review also assessed the psychometric properties of instruments used to measure the outcomes of ACP interventions. None of the ACP-specific PROMs fully met the COSMIN quality standards for measurement development. No measurement error was reported in any PROM, and evidence for several other key properties was limited. While the COSMIN checklist could not be strictly applied to all included PROMs, less than half met at least 50% of its criteria. This suggests that many PROMs used in ACP research might not be optimal, which could potentially lead to inaccurate conclusions about intervention effectiveness due to insufficient data quality. Therefore, for the evaluation of the effectiveness of ACP interventions in future research, it is imperative to first assess the appropriateness of the psychometric properties of the PROMs to measure such outcomes.

However, it must be acknowledged that the repeated use of certain PROMs across diverse populations, timeframes, and studies can contribute to accumulating evidence for their validity, even in the absence of complete formal psychometric evaluations. This ongoing application in various contexts over time strengthens the validity argument.

Eighteen PROMs, specifically those focusing on the process domain or both process and action domains, exceeded 4 points out of 9 on the COSMIN checklist. Among these, the ACP Engagement Survey (Sudore et al. 2017a) achieved higher COSMIN scores compared to others. Notably, the ACP Engagement Survey demonstrates a broader perspective by not only capturing the conceptualization of ACP as a process but also examining participants' action changes following interventions. Furthermore, translations and adaptations into various languages including Chinese (Wei et al. 2022), Dutch (van der Smitten et al. 2021), Japanese (Okada et al. 2021), and Spanish (Sudore et al. 2017a) highlight its potential for diverse applications across populations. The review's identification of the 4-item version as frequently used further underscores its potential value for concise and efficient data collection in research and clinical settings.

Among the PROMs evaluated in this review for QOC assessment, 11 achieved scores exceeding 4 points out of 9 on the COSMIN checklist. Interestingly, 8 of these PROMs were generic satisfaction measures. Four of the PROMs demonstrated higher

utilization frequencies: the QOC (Engelberg *et al.* 2006), the DCS (O'Connor 1995) including the SURE test (Légaré *et al.* 2010), and the Life-Support Preferences Questionnaire (LSPQ) (Coppola *et al.* 1999). The QOC assesses patient satisfaction with communication regarding end-of-life care with clinicians and demonstrated evidence for 78% (7 points) of the COSMIN criteria. Both the DCS and SURE test assess patient satisfaction with their own value-based decision-making, with the SURE test being a shorter version of the DCS. While the DCS was initially developed as a generic measure, its frequent use in ACP research reflects its applicability in evaluating patient decision-making based on personal values (O'Connor 1993; 2010), *i.e.*, the key outcome of ACP. Both measures demonstrated evidence for 67% (6 points) of the COSMIN criteria in this review. This review included the Korean version of DCS (Kim *et al.* 2017), and several language versions for the DCS have been developed (O'Connor 1993; 2010). The LSPQ (Coppola *et al.* 1999), which is specifically designed to measure the outcome of ACP interventions, is frequently used to assess the congruence of treatment preferences between patients and their surrogates (Chan *et al.* 2018; Ditto *et al.* 2001; Jo *et al.* 2021; Ke *et al.* 2021; Song *et al.* 2010, 2009). However, it exhibited low evidence for meeting the COSMIN criteria. PROMs employing hypothetical vignettes, such as the LSPQ, may not fully align with the COSMIN criteria for psychometric properties like structural validity, cross-cultural validity, measurement error, or criterion validity.

Of the 86 PROMs included in this review, approximately 51% demonstrated limited or no evidence for meeting COSMIN criteria (*i.e.*, scoring less than 4 points), which encompassed all PROMs within the action domain and lacked consistent usage across studies. Notably, the PROMs in this domain primarily serve as tools that patients can use to express preferences or make end-of-life care decisions. In ACP research, action outcomes are typically assessed through the presence of documented forms (*e.g.*, goal of care, advance directives, physician orders for life-sustaining treatment) in patient charts, rather than relying on PROMs (Jimenez *et al.* 2018; Park *et al.* 2021). This suggests that the PROMs reviewed within the action domain may be perceived more as documents, which could potentially explain the scarcity of validation studies confirming their psychometric properties. Consequently, to evaluate the action outcomes of ACP interventions, utilizing existing standardized advance directive forms specific to each region might be more effective than developing new measurement tools.

This review offers valuable insights into selecting appropriate PROMs for assessing ACP outcomes in older adults based on their psychometric properties. The review also identifies key considerations for researchers and clinicians to ensure effective data collection in this population. First, the self-reported nature of PROMs, while advantageous in capturing patient perspectives, may pose challenges for older adults. This review highlights potential challenges involved in the self-reported format. The finding that 22% of included studies ($n = 19$ out of 86) involved researcher assistance, often through the form of interviews, suggests that self-reporting may be difficult for some older adults. Moreover, the frequent use of shortened versions of PROMs further supports this notion. These challenges likely stem from broader functional decline, which can manifest as fatigue or difficulty reading and comprehending lengthy questionnaires. Therefore, researchers and clinicians could consider implementing shortened versions of PROMs and providing assistance during data collection as potential interventions to mitigate these difficulties and enhance participation rates among

older adults. Second, this review underscores the need for further validation of most included PROMs in the context of assessing ACP intervention outcomes in older adults. This highlights the importance of using rigorous psychometric evaluation methods to ensure the reliability and validity of collected data. Researchers are encouraged to adhere to established standards such as the COSMIN checklist (Mokkink *et al.* 2010) when selecting PROMs for use in ACP research, which will ultimately lead to more robust and meaningful data for evaluating intervention effectiveness in the populations.

This review has certain limitations that could potentially have impacted its findings. First, the review's reliance on references provided by included articles may have excluded relevant validation studies, thus leading to a potential underestimation of the psychometric properties of some measurements. Second, the focus on the elderly population was compromised by the use of average age across studies, which potentially included subjects outside the target demographic. This raises concerns about the suitability of some measurements as specific tools for older adults. Despite these limitations, this review provides valuable insights into PROMs for assessing ACP outcomes in older adults. Notably, it identifies several adequately validated PROMs based on the COSMIN checklist, including the ACP Engagement Survey, QOC, and DCS. However, the review also highlights the need for further validation studies to be conducted on many previously used PROMs while specifically considering the unique conditions faced by older adult participants.

Supplementary material. The supplementary material for this article can be found at <https://doi.org/10.1017/S1478951524002062>.

Acknowledgments. We would like to acknowledge NaJin Kim, the professional librarian at the Medical Library of the Catholic University of Korea, for her expert support in literature searching, which was essential for the successful data extraction in this review. We also wish to acknowledge the financial support of the Catholic Medical Center Research Foundation made in the program year of 2022.

Competing interests. The authors declared no competing interests.

Ethical approval. This study was granted exemption by the Institutional Review Board of the Catholic University of Korea (No: MC22ZISI0083).

References

- Bekker H, Légaré F, Walker E, *et al.* (2012) *SHARED–Patient Experience of Shared Decision Making Questionnaire*. UK: Bekker at University of Leeds.
- Carr D and Luth EA (2017) Advance care planning: contemporary issues and future directions. *Innovation in Aging* 1(1), 1–10. doi:10.1093/geron/i/igx012
- Chan HYL, Ng JSC, Chan KS, *et al.* (2018) Effects of a nurse-led post-discharge advance care planning programme for community-dwelling patients nearing the end of life and their family members: A randomised controlled trial. *International Journal of Nursing Studies* 87, 26–33. doi:10.1016/j.ijnurstu.2018.07.008
- Coppola KM, Bookwala J, Ditto PH, *et al.* (1999) Elderly adults' preferences for life-sustaining treatments: The role of impairment, prognosis, and pain. *Death Studies* 23, 617–634. doi:10.1080/074811899200803
- Ditto PH, Danks JH, Smucker WD, *et al.* (2001) Advance directives as acts of communication: A randomized controlled trial. *Archives of Internal Medicine* 161(3), 421–430. doi:10.1001/archinte.161.3.421
- Engelberg R, Downey L and Curtis JR (2006) Psychometric characteristics of a quality of communication questionnaire assessing communication about

- end-of-life care. *Journal of Palliative Medicine* 9(5), 1086–1098. doi:10.1089/jpm.2006.9.1086
- Ferron Parayre A, Labrecque M, Rousseau M, *et al.* (2014) Validation of SURE, a four-item clinical checklist for detecting decisional conflict in patients. *Medical Decision Making* 34(1), 54–62. doi:10.1177/0272989X13491463
- Fried TR, Redding CA, Robbins ML, *et al.* (2010) Stages of change for the component behaviors of advance care planning. *Journal of the American Geriatrics Society* 58(12), 2329–2336. doi:10.1111/j.1532-5415.2010.03184.x
- Fusar-Poli P and Radua J (2018) Ten simple rules for conducting umbrella reviews. *BMJ Mental Health* 21(3), 95–100. doi:10.1136/ebmental-2018-300014
- Heyland DK, Cook DJ, Rocker GM, *et al.* and Network CRatEoL (2010) The development and validation of a novel questionnaire to measure patient and family satisfaction with end-of-life care: The Canadian Health Care Evaluation Project (CANHELP) Questionnaire. *Palliative Medicine* 24(7), 682–695. doi:10.1177/0269216310373168
- Heyland DK, Jiang X, Day AG, *et al.* and Network CRatEoL (2013) The development and validation of a shorter version of the Canadian Health Care Evaluation Project Questionnaire (CANHELP Lite): A novel tool to measure patient and family satisfaction with end-of-life care. *Journal of Pain and Symptom Management* 46(2), 289–297. doi:10.1016/j.jpainsymman.2012.07.012
- Hinderer KA and Lee MC (2019) Chinese Americans' attitudes toward advance directives: An assessment of outcomes based on a nursing-led intervention. *Applied Nursing Research* 49, 91–96. doi:10.1016/j.apnr.2019.04.003
- Institute of Medicine (2014) *Dying in America: improving Quality and Honoring Individual Preferences near the End of Life*. Washington DC: The National Academics Press.
- Jimenez G, Tan WS, Virk AK, *et al.* (2018) Overview of systematic reviews of advance care planning: Summary of evidence and global lessons. *Journal of Pain and Symptom Management* 56(3), 436–459.e425. doi:10.1016/j.jpainsymman.2018.05.016
- Jo M, Park M, Park EJ, *et al.* (2021) The acceptability and effect of a communication-based advance care planning program for older adults. *Journal of Hospice & Palliative Nursing* 23(4), 375–385. doi:10.1097/NJH.0000000000000762
- Joung H (2012) *Effects of a Self-decision-making Enrichment Program on Senior Citizen's Advance Directives*. Kangwon National University. Chuncheon: South Korea.
- Ke LS, Hu WY, Chen CY, *et al.* (2021) A quasi-experimental evaluation of advance care planning improves consistency between elderly individuals and their surrogates regarding end-of-life care preferences: Development and application of a decision aid with cartoon pictures. *Patient Education and Counseling* 104(4), 815–825. doi:10.1016/j.pec.2020.09.027
- Kim J, Kim S, Hong SW, *et al.* (2017) Validation of the decisional conflict scale for evaluating advance-dwelling older adults. *Asian Nursing Research* 11(4), 297–303. doi:10.1016/j.anr.2017.11.004
- Larsen DL, Attkisson CC, Hargreaves WA, *et al.* (1979) Assessment of client/patient satisfaction: Development of a general scale. *Evaluation & Program Planning* 2(3), 197–207. doi:10.1016/0149-7189(79)90094-6
- Lee MC, Hinderer KA and Friedmann E (2015) Engaging Chinese American adults in advance care planning: A community-based, culturally sensitive seminar. *Journal of Gerontological Nursing* 41(8), 17–21. doi:10.3928/00989134-20150406-01
- Légaré F, Kearing S, Clay K, *et al.* (2010) Are you SURE?: Assessing patient decisional conflict with a 4-item screening test. *Canadian Family Physician* 56(8), e308–314. <https://www.cfp.ca/content/56/8/e308.long>
- Li J, Han X, Zhang X, *et al.* (2019) Spatiotemporal evolution of global population ageing from 1960 to 2017. *BMC Public Health* 19(1), 127. doi:10.1186/s12889-019-6465-2
- McMahan RD, Tellez I and Sudore RL (2021) Deconstructing the complexities of advance care planning outcomes: What do we know and where do we go? A scoping review. *Journal of the American Geriatrics Society* 69(1), 234–244. doi:10.1111/jgs.16801
- Mesters I, Van Den Borne H, McCormick L, *et al.* (1997) Openness to discuss cancer in the nuclear family: Scale, development, and validation. *Psychosomatic Medicine* 59(3), 269–279. doi:10.1097/00006842-199705000-00010
- Mokkink LB, Prinsen CA, Patrick DL, *et al.* (2018). *COSMIN methodology for systematic reviews of patient-reported outcome measures (PROMs) user manual*. <https://www.cosmin.nl/tools/guideline-conducting-systematic-review-outcome-measures/> (accessed 8 March 2024)
- Mokkink LB, Terwee CB, Patrick DL, *et al.* (2010) The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: An international Delphi study. *Quality of Life Research* 19(4), 539–549. doi:10.1007/s11136-010-9606-8
- Nolan MT and Bruder M (1997) Patients' attitudes toward advance directives and end-of-life treatment decisions. *Nursing Outlook* 45(5), 204–208. doi:10.1016/s0029-6554(97)90066-x
- O'Connor A (1993; 2010) *User Manual - Decisional Conflict Scale*. Ottawa, Canada: Ottawa Hospital Research Institute. https://decisionaid.ohri.ca/eval_dcs.html (accessed 1 December 2022)
- O'Connor AM (1995) Validation of a decisional conflict scale. *Medical Decision Making* 15(1), 25–30. doi:10.1177/0272989x9501500105
- Okada H, Takenouchi S, Okuhara T, *et al.* (2021) Development of a Japanese version of the advance care planning engagement survey: Examination of its reliability and validity. *Palliative and Supportive Care* 19(3), 341–347. doi:10.1017/S1478951520001108
- Park EJ, Jo M, Park M, *et al.* (2021) Advance care planning for older adults in community-based settings: An umbrella review. *International Journal of Older People Nursing* 16(5), e12397. doi:10.1111/opn.12397
- Prinsen CAC, Mokkink LB, Bouter LM, *et al.* (2018) COSMIN guideline for systematic reviews of patient-reported outcome measures. *Quality of Life Research* 27(5), 1147–1157. doi:10.1007/s11136-018-1798-3
- Rietjens JAC, Sudore RL, Connolly M, *et al.* and European Association for Palliative Care (2017) Definition and recommendations for advance care planning: An international consensus supported by the European Association for Palliative Care. *The Lancet Oncology* 18(9), e543–e551. doi:10.1016/S1473-0451(17)30582-X
- Song M-K, Donovan HS, Piraino BM, *et al.* (2010) Effects of an intervention to improve communication about end-of-life care among African Americans with chronic kidney disease. *Applied Nursing Research* 23(2), 65–72. doi:10.1016/j.apnr.2008.05.002
- Song MK, Ward SE, Happ MB, *et al.* (2009) Randomized controlled trial of SPIRIT: An effective approach to preparing African-American dialysis patients and families for end of life. *Research in Nursing and Health* 32(3), 260–273. doi:10.1002/nur.20320
- Sudore RL, Heyland DK, Barnes DE, *et al.* (2017a) Measuring advance care planning: optimizing the advance care planning engagement survey. *Journal of Pain and Symptom Management* 53(4), 669–681.e668. doi:10.1016/j.jpainsymman.2016.10.367
- Sudore RL, Heyland DK, Lum HD, *et al.* (2018) Outcomes that define successful advance care planning: A Delphi panel consensus. *Journal of Pain and Symptom Management* 55(2), 245–255.e248. doi:10.1016/j.jpainsymman.2017.08.025
- Sudore RL, Lum HD, You JJ, *et al.* (2017b) Defining advance care planning for adults: A consensus definition from a multidisciplinary Delphi panel. *Journal of Pain and Symptom Management* 53(5), 821–832.e821. doi:10.1016/j.jpainsymman.2016.12.331
- The Lancet Respiratory Medicine (2016) Prolonging life at all costs: Quantity versus quality. *The Lancet Respiratory Medicine* 4(3), 165. doi:10.1016/S2213-2600(16)00059-X
- Thomas K, Lobo B and Detering K (2018) *Advance Care Planning in End of Life Care*, Second edition. edn. New York, NY: Oxford University Press.
- van der Smitten D, van der Heide A, Sudore RL, *et al.* (2021) Measuring advance care planning behavior in Dutch adults: Translation, cultural adaptation and validation of the Advance Care Planning Engagement Survey. *BMC Medical Research Methodology* 21(1), 194. doi:10.1186/s12874-021-01389-5
- Wei FC, Hsu CK, Wu YL, *et al.* (2022) Reliability and validity of the traditional Chinese Version of the advance care planning engagement survey: A

- pilot evaluation in Taiwanese outpatients. *Journal of Palliative Care* 37(3), 273–279. doi:10.1177/08258597211051208
- Wendrich-van Dael A, Bunn F, Lynch J, et al.** (2020) Advance care planning for people living with dementia: An umbrella review of effectiveness and experiences. *International Journal of Nursing Studies* 107, 103576. doi:10.1016/j.ijnurstu.2020.103576
- Woods T, Palmarini N, Corner L, et al.** (2022) Quantum Healthy Longevity for healthy people, planet, and growth. *The Lancet Healthy Longevity* 3(12), e811–e813. doi:10.1016/S2666-7568(22)00267-7
- World Health Organization.** (2020). *Situation and trends*. World Health Organization. <https://www.who.int/data/gho/data/themes/mortality-and-global-health-estimates/ghe-life-expectancy-and-healthy-life-expectancy> (accessed 15 December 2023)