Behavioral Psychology /Psicología Conductual

Behav. Psychol. **2026**; 34(4): 52713
https://doi.org/10.31083/BP52713

Systematic Review

A Scoping Review of Tools for Evaluating Family-Caregiver Competence in Managing Behavioral and Psychological Symptoms of Dementia

Jie Dong[1,†], Yayan Yi[1,†], Yimei Hu[1], Yanan Mei[1], Yue Fu[1], Yibo Wu[1,*]

[1]Department of Nursing, The Fourth Affiliated Hospital of School of Medicine, and International School of Medicine, International Institutes of Medicine, Zhejiang University, 322000 Yiwu, Zhejiang, China
*Correspondence: wuyiboism@zju.edu.cn (Yibo Wu)
[†]These authors contributed equally.
Academic Editor: Chung-Ying Lin
Submitted: 8 April 2026 Revised: 9 June 2026 Accepted: 12 June 2026 Published: 28 July 2026

Abstract

Background: Behavioral and psychological symptoms of dementia (BPSD) affect up to 90% of individuals with dementia and represent a primary source of stress for family caregivers. Nevertheless, a systematic synthesis of validated instruments for assessing caregiver competence in BPSD management is currently lacking. This scoping review aimed to identify and evaluate tools used to measure family caregiver competence in managing BPSD. **Methods:** Following established scoping review methodology, we systematically searched six electronic databases: PubMed, CINAHL, Web of Science, PsycINFO, Embase, and the Cochrane Library. Eligible studies were original articles focused on the development, validation, or adaptation of assessment tools for family caregivers aged ≥18 years. Data extraction and quality assessment were performed using the COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) Risk of Bias Checklist and associated guidelines to evaluate methodological quality, measurement properties, and certainty of evidence. **Results:** From 62,892 identified records, 11 studies were included. The majority of studies received a validity rating of "doubtful", with the exception of Competence Scale in Managing Behavioral and Psychological Symptoms of Dementia (CS-MBPSD), which received an "adequate" rating. Cross-cultural validity was not reported for any of the tools. Overall, the quality of evidence was low-to-moderate. All 11 tools received a recommendation grade of "B", indicating potential suitability for use; however, they require further validation. **Conclusions:** The CS-MBPSD may represent the most suitable tool for evaluating family caregiver competence in managing BPSD; however, its measurement properties require further testing. Future tool development should adhere to COSMIN standards, incorporate cross-cultural validation, and prioritize responsiveness and real-world applicability. **Registration:** The study has been registered on https://www.crd.york.ac.uk/PROSPERO/ (registration number: CRD420261364464; registration link: https://www.crd.york.ac.uk/PROSPERO/view/CRD420261364464).

Keywords: Alzheimer's disease; assessment tools; behavioral and psychological symptoms; family caregivers; management competence

1. Introduction

The prevalence of dementia among older adults is increasing at an alarming rate due to evolving global demographics (GBD 2021 Nervous System Disorders Collaborators, 2024). According to the World Health Organization, more than 55 million people were living with dementia in 2019, a number projected to rise to 78 million by 2030 and 139 million by 2050 (World Health Organization, 2021). Dementia is a severe social and public health challenge (Mukherjee et al., 2026), and Alzheimer's disease, the most prevalent cause of dementia, inflicts profound and multifaceted damage. For patients, it robs them of memory, autonomy, and dignity (Al Mokdad et al., 2025). A 2025 survey by the Alzheimer's Society found that 42% of people with dementia felt ashamed or stigmatized, 73% felt anxious or worried about living with their symptoms, and 72% found the diagnosis process stressful (Alzheimer's Society, 2025). For families, it places immense emotional and financial strain on caregivers; and for society, it demands escalating healthcare expenditures and social care resources (Cebolla Sousa et al., 2025; Pozzi et al., 2024).

As the disease progresses, most individuals develop varying degrees of behavioral and psychological symptoms of dementia (BPSD) (Pickering et al., 2024), with up to 90% experiencing at least one symptom (Anantapong et al., 2025). More than half of family caregivers reported that providing home-based care is both difficult and burdensome, and managing BPSD has been consistently identified as a primary source of caregiver stress (Gonçalves-Pereira et al., 2025; Lillekroken et al., 2024). A 2026 study further confirmed that BPSD factors were significantly associated with both caregiver burden and depression symptoms (Lim et al., 2026). In the context of dementia care, family caregivers often assume the role of frontline care providers, navigating complex symptom fluctuations with limited formal support (Ramakrishnan and Malhotra, 2025). Given this central role, the strengthening of caregivers' knowledge, practical skills, and access to adequate support is essential

for improving outcomes for both caregivers and care recipients (Watt et al., 2024).

Managing BPSD requires distinct competencies beyond general caregiving. The fluctuating and often unpredictable nature of BPSD demands that caregivers possess not only foundational dementia-care knowledge but also specialized competencies in symptom recognition, behavioral assessment, de-escalation techniques, and individualized non-pharmacological intervention strategies (Yin et al., 2024). Therefore, accurate assessment of caregiver competence is essential to identifying individual strengths and gaps, informing tailored interventions, and facilitating effective, person-centered support (Aguirre et al., 2025; Biernetzky et al., 2025). Research on the measurement of family caregivers' capacity to manage BPSD has progressed. Early work in this area primarily focused on caregivers' general self-efficacy, rather than the specific competencies required to handle complex and fluctuating BPSD episodes. The Revised Scale for Caregiving Self-Efficacy (RSCSE) (Steffen et al., 2002) assesses caregivers' confidence in responding to disruptive behaviors, but it does not capture the nuanced skills involved in identifying BPSD triggers. Similarly, the Caregiver Efficacy Scale (CES) (Crellin et al., 2014) measures caregivers' confidence in managing mood and behavioral disturbances, but its validation relied primarily on associations with burden and depression rather than on direct testing of observed caregiving behaviors. Recently, the Competence Scale in Managing Behavioral and Psychological Symptoms of Dementia (CSMBPSD) (Cheon et al., 2022) was developed specifically to assess caregivers' competence across six multidimensional domains, representing an important shift from confidence-based to competence-focused measurement.

Existing assessment tools exhibit considerable heterogeneity in their development methods, psychometric properties, and cultural applicability (Hu et al., 2023; Song et al., 2026). In addition, they show insufficient content validity verification during instrument development, a lack of cross-cultural validation and measurement invariance for testing across different linguistic or ethnic groups, and minimal evidence on responsiveness. Clinicians and researchers face increased difficulty in selecting appropriate instruments due to the considerable heterogeneity across tools in item content, scoring procedures, and validation strategies (Lillekroken et al., 2025). In the absence of a systematic review and comparative analysis, both researchers and practitioners are left without evidence-based guidance, making the choice of suitable tools for evaluating caregivers' management ability persistently challenging (Ding et al., 2026). To evaluate the methodological quality and psychometric properties of the included instruments rigorously, this study adopted the COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) framework (Mokkink et al., 2024).

The present study systematically reviewed existing scales for evaluating caregivers' management ability, comprehensively described their basic characteristics, assessed their methodological quality, and synthesized their psychometric evidence in accordance with COSMIN methodological quality standards (Mokkink et al., 2020). We formulated evidence-based recommendations (Category A, B, or C) for instrument selection, directly guiding researchers in choosing the most suitable tool for their specific practice or research context.

2. Materials and Methods

Our review adopted a scoping-review design using the methodological framework of Arksey and O'Malley and JBI guidance (Peters et al., 2015). Conducted in line with the Preferred Reporting Items for Systematic reviews and Meta-Analyses extension for Scoping Reviews (PRISMA-ScR) guidelines, this scoping review had its protocol registered with PROSPERO (CRD420261364464).

2.1 Definition of the Research Question

Our present scoping review addressed the following specific research questions: first, which instruments have been used to assess family caregiver competence in managing BPSD? second, what psychometric properties do these instruments possess?

2.2 Conceptual Framework and Construct Definition

Family caregiver competence in this domain is multidimensional, including cognitive, functional, social, and metacognitive abilities (Jung et al., 2022). Instruments that measure self-efficacy or confidence specifically in the context of BPSD management were therefore included because self-efficacy is an established predictor of actual caregiving behavior and is frequently used as a proxy for competence in clinical research. Excluding such measures would have omitted a substantial portion of the available evidence base and limited the comprehensiveness of the review.

2.3 Search Strategy

We searched PubMed, CINAHL, Web of Science, PsycINFO, Embase, and Cochrane Library using MeSH and free-text terms. The search terms encompassed four conceptual domains: dementia, caregiver, assessment tools (e.g., scale, instrument), and BPSD (e.g., neuropsychiatric symptoms, agitation). The search period extended from database inception to September 20, 2025 (**Supplementary Tables 1–6**).

2.4 Inclusion and Exclusion Criteria

Inclusion criteria for the study were as follows: (1) original articles focusing on the development, application, validation, localization, or revision of an assessment tool; (2) participants were family caregivers (unpaid relatives or friends) who cared for a person with a confirmed dementia

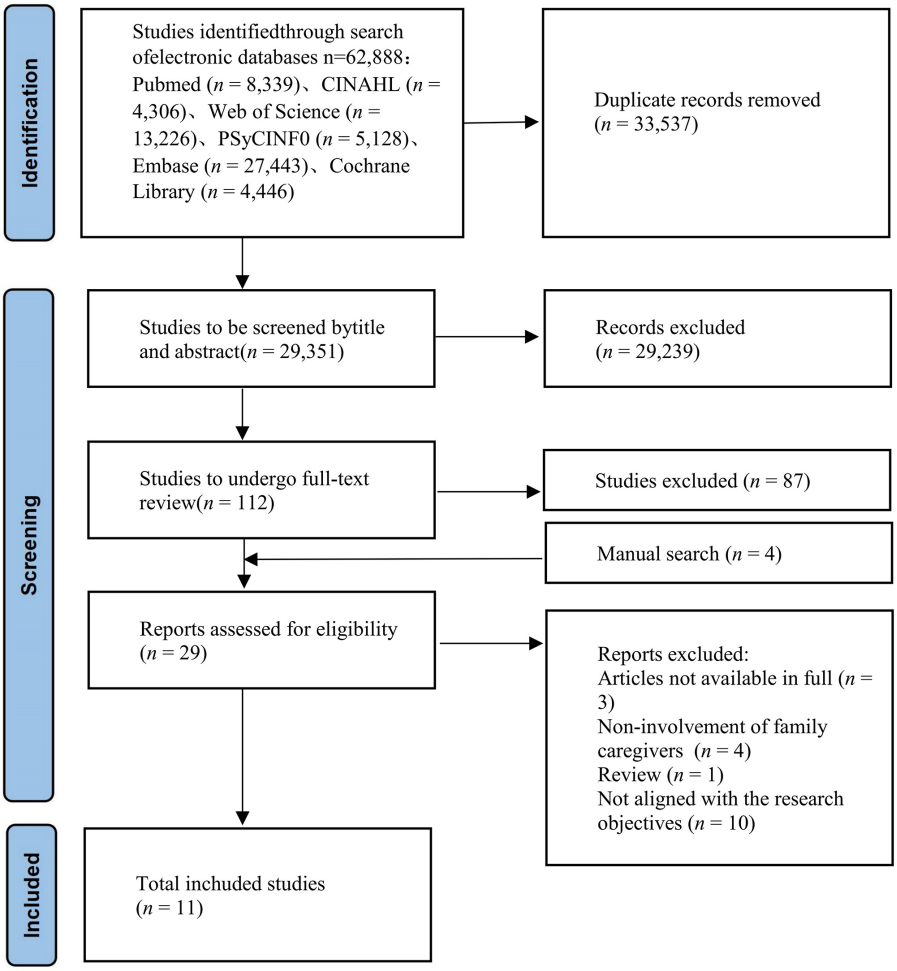

Fig. 1. PRISMA-ScR flow diagram of the study selection process. This Figure illustrates the number of records identified, screened, assessed for eligibility, and finally included in the scoping review. PRISMA-ScR, Preferred Reporting Items for Systematic reviews and Meta-Analyses extension for Scoping Reviews.

diagnosis; (3) participants aged ≥18 years. Studies were excluded if: (1) they were conference abstracts, reviews, or editorial comments; (2) they were duplicate publications; (3) they were non-English publications due to constraints in time, funding, and language resources; (4) full-text articles were inaccessible; (5) they measured general caregiving self-efficacy or competence without any specific reference to BPSD; (6) they assessed caregiver burden, depression, anxiety, or quality of life without measuring perceived ability or competence in managing BPSD.

2.5 Study Selection and Data Extraction

Records were imported into EndNote to remove duplicates. Two reviewers then independently screened titles, abstracts, and full texts against the eligibility criteria, with disagreements resolved through consensus or consultation with a third reviewer. Key data (author, tool name, domains/items, and score range) were extracted from included studies.

2.6 Methodological and Psychometric Evaluation

COSMIN is widely recognized as the international gold standard for systematic reviews of Patient-Reported Outcome Measures, offering a comprehensive and standardized approach to assessing key measurement properties (Mokkink et al., 2024). This framework has been increasingly adopted in recent high-quality reviews across various clinical fields, including goal-setting outcome measures in dementia care (Budgett et al., 2024) and cognitive assessment tools for mild cognitive impairment and Alzheimer's disease (Peng et al., 2024), thereby demonstrating its broad applicability and methodological rigor. This framework was selected because it prioritizes content validity (Cheng et al., 2025), which is essential for evaluating tools designed to capture the complex, multidimensional construct of caregiver competence in BPSD management. Moreover, the COSMIN Risk of Bias Checklist (Mokkink et al., 2018) enables rigorous appraisal of methodological quality in heterogeneous validation studies and synthesizes evidence to guide tool selection (Ren et al., 2025).

Using the COSMIN Risk of Bias Checklist (Mokkink et al., 2018), the methodological quality of the included scales was appraised. Each item within a module was rated for risk of bias as "very good", "adequate", "doubtful", or "inadequate". And the module's overall rating was taken as the lowest of those item scores.

Subsequently, COSMIN quality specifications were used to evaluate the measurement properties, covering nine domains. Each property was rated as sufficient (+), inadequate (–), or indeterminate (?). For internal consistency: Cronbach's $\alpha \geq 0.70$ was "sufficient"; for structural validity: comparative fit index (CFI) >0.95, RMSEA <0.06, or SRMR <0.08 was "sufficient".

The overall rating for a given measurement property of a scale was determined as follows: if all relevant studies consistently reported the property as "sufficient (+)", "inadequate (–)", or "indeterminate (?)", that rating was assigned as the overall conclusion. However, if studies yielded inconsistent ratings (e.g., both "+" and "–") and the inconsistency could not be reasonably explained, the overall rating was designated as "inconsistent (±)".

Finally, based on the graded evidence (Alhazzani and Guyatt, 2018; Guyatt et al., 2025), a formal recommendation for the use of each scale was formulated and assigned a strength category guided by predefined rules: A (recommended); B (potential but requires further validation); or C (not recommended).

3. Results

3.1 Literature Search Results

After screening 62,892 records, 11 studies (Cheon et al., 2022; Crellin et al., 2014; Farran et al., 2011; Fortinsky et al., 2002; Huang et al., 2003; Jansen et al., 2007; Márquez-González et al., 2009; Piggott et al., 2017; Steffen et al., 2002; Zeiss et al., 1999; Zhang et al., 2013) were included. The study selection process is detailed in Fig. 1. The PRISMA extension guidelines checklist is also included as **Supplementary Table 7**.

3.2 Basic Characteristics of the Included Studies

In this review, ten studies were included that focused on the development and validation of assessment tools, and one study (Márquez-González et al., 2009) concerned cross-cultural validation or application. Regarding the chronology of tool development, six studies (Fortinsky et al., 2002; Huang et al., 2003; Jansen et al., 2007; Márquez-González et al., 2009; Steffen et al., 2002; Zeiss et al., 1999) were published between 1990 and 2010, and five (Cheon et al., 2022; Crellin et al., 2014; Farran et al., 2011; Piggott et al., 2017; Zhang et al., 2013) between 2010 and 2025. Finally, 10 scales were included originating from multiple countries and regions, including the United States, Spain, China and Taiwan, the Netherlands, the United Kingdom, and South Korea. Sample sizes ranged from 43 to 460 participants. The number of items per scale varied from 9 to 43, covering 2–6 domains. Scoring methods included continuous 0–100 scales and 4- or 5-point Likert scales. Test-retest intervals ranged from 2 days to 6 weeks, and certain studies did not report scale completion time or retest information. An overview of the basic characteristics of the included studies is presented in **Supplementary Table 8**.

3.3 Methodological Quality and Measurement Properties Quality

The scale's cross-cultural validity and measurement invariance were not assessed in any of the 11 studies. The psychometric properties and methodological quality evaluation of included instruments for assessing caregiver competence in BPSD management are provided in Table 1.

3.3.1 Patient-Reported Outcome Measure (PROM) Development

11 of the included tools were rated as "doubtful" in methodological quality. The limitations of those studies included exclusive use of quantitative methods during development, absence of qualitative methods, and inadequate detail in describing interviews and statistical analyses.

3.3.2 Validity

3.3.2.1 Content Validity. Only the CS-MBPSD (Cheon et al., 2022) received an "adequate" methodological rating for content validity, based on systematic consultation with both experts and patients regarding relevance, comprehensiveness, and comprehensibility, and a "moderate" evidence rating for content validity. The Caregiver Assessment of Behavioral Skill-Self-Report (CAB-SR) (Farran et al., 2011) and the Caregiver Confidence in Medical Sign/Symptom Management scale (CCSM) (Piggott et al., 2017) also involved the consultation of experts and patients. Two studies (Huang et al., 2003; Zeiss et al., 1999) exclusively consulted patients regarding the scale's content validity, and one study (Márquez-González et al., 2009) consulted only experts; most others did not report any structured content validation. However, all of the above studies had limitations in that the descriptions of the study procedures or statistical methods were insufficiently specific or detailed. Therefore, the methodological quality of these studies was rated as "doubtful".

3.3.2.2 Structural Validity. Exploratory Factor Analysis (EFA) is used to explore the underlying factor structure of a scale without imposing a preconceived structure. CFI is a goodness-of-fit index used in structural equation modeling (SEM) and confirmatory factor analysis (CFA) to evaluate how well a hypothesized model reproduces the observed covariance matrix. CFA was used in two of the studies (Cheon et al., 2022; Steffen et al., 2002), both of which reported a CFI below 0.95, indicating insufficient structural validity. Six studies (Crellin et al., 2014; Farran et al., 2011; Fortinsky et al., 2002; Jansen et al., 2007; Piggott et al., 2017;

Table 1. Psychometric properties and methodological quality evaluation of included instruments for assessing caregiver competence in BPSD management (*n* = 11).

PROMs	PROMs development	Content validity	Structural validity	Internal consistency	Reliability	Measurement error	Hypothesis testing	Criterion validity	Responsiveness	Cross-cultural validity
CSES	Doubtful	Doubtful/?	NA/NA	Doubtful/?	Doubtful/?	NA/NA	Doubtful/?	NA/NA	NA/NA	NA/NA
FCSE	Doubtful	Doubtful/?	Adequate/?	Very good/+	NA/NA	NA/NA	Adequate/+	NA/NA	NA/NA	NA/NA
RSCSE	Doubtful	Doubtful/?	Very good/–	Very good/+	Doubtful/?	NA/NA	Very good/+	NA/NA	NA/NA	NA/NA
Spanish Version of the RSCSE	Doubtful	Doubtful/?	NA/NA	Doubtful/?	NA/NA	NA/NA	Adequate/–	NA/NA	NA/NA	NA/NA
AMSS	Doubtful	Doubtful/?	NA/NA	Doubtful/?	NA/NA	NA/NA	NA/NA	NA/NA	Doubtful/?	NA/NA
SCQ	Doubtful	Doubtful/?	Inadequate/–	Doubtful/?	NA/NA	NA/NA	Very good/–	NA/NA	NA/NA	NA/NA
CAB-SR	Doubtful	Doubtful/?	Inadequate/–	Inadequate/–	Doubtful/?	NA/NA	Adequate/–	NA/NA	NA/NA	NA/NA
SEQCFC	Doubtful	Doubtful/?	Adequate/?	Very good/+	Doubtful/?	NA/NA	Very good/+	NA/NA	NA/NA	NA/NA
CES	Doubtful	Doubtful/?	Adequate/?	Very good/+	NA/NA	NA/NA	Very good/+	Adequate/–	NA/NA	NA/NA
CCSM	Doubtful	Doubtful/?	Adequate/?	Very good/+	Doubtful/?	NA/NA	Adequate/–	NA/NA	NA/NA	NA/NA
CS-MBPSD	Doubtful	Adequate/+	Very good/–	Very good/+	Adequate/+	Adequate/–	Adequate/+	NA/NA	NA/NA	NA/NA

BPSD, Behavioral and Psychological Symptoms of Dementia; PROMs, Patient-Reported Outcome Measure; CSES, Caregiver Self-Efficacy Scales; FCSE, Family caregiver self-efficacy; RSCSE, Revised Scale for Caregiving Self-Efficacy; Spanish Version of the RSCSE, Spanish Version of the Revised Scale for Caregiving Self-Efficacy; AMSS, Agitation Management Self-Efficacy Scale; SCQ, Sense of Competence Questionnaire; CAB-SR, Caregiver Assessment of Behavioral Skill-Self-Report; SEQCFC, Self-Efficacy Questionnaire for Chinese Family Caregivers; CES, Caregiver Efficacy Scale; CCSM, Caregiver Confidence in Medical Sign/Symptom Management scale; CS-MBPSD, Competence Scale in Managing Behavioral and Psychological Symptoms of Dementia; +, sufficient; –, insufficient; ?, indeterminate; NA, Not Applicable.

Zhang et al., 2013) used only EFA and did not report CFI. Therefore, the structural validity of the scale was "insufficient" or "doubtful". The remaining three studies (Huang et al., 2003; Márquez-González et al., 2009; Zeiss et al., 1999) did not report the structural validity.

3.3.2.3 Hypothesis Testing. Four studies (Crellin et al., 2014; Jansen et al., 2007; Steffen et al., 2002; Zhang et al., 2013) were rated "very good" in methodological quality. Those studies provided detailed information on hypothesis testing and correlation, and their test results consistently aligned with the hypotheses, resulting in a "sufficient" rating for measurement attributes. Five studies (Cheon et al., 2022; Farran et al., 2011; Fortinsky et al., 2002; Márquez-González et al., 2009; Piggott et al., 2017) received an "adequate" methodological quality rating. The Caregiver Self-Efficacy Scales (CSES) (Zeiss et al., 1999) received a "doubtful" methodological quality rating due to the absence of a description of the handling of missing values. Only the Agitation Management Self-Efficacy Scale (AMSS) (Huang et al., 2003) did not mention hypothesis testing.

3.3.2.4 Criterion Validity. The criterion validity was reported in only one study (Crellin et al., 2014). This evaluation's methodological quality was rated "doubtful", as the comparator instrument did not meet COSMIN gold standard requirements and its measurement quality was "insufficient". No other study reported criterion validity.

3.3.3 Reliability

3.3.3.1 Internal Consistency. All 11 studies reported internal consistency. Six of them (Cheon et al., 2022; Crellin et al., 2014; Fortinsky et al., 2002; Piggott et al., 2017; Steffen et al., 2002; Zhang et al., 2013), had Cronbach's $\alpha \geq 0.7$, resulting in a "very good" methodological quality rating and a "sufficient" measurement quality rating.

However, because EFA revealed that the factor structure for the two sub-scales was not clearly defined and that their items exhibited cross-loadings, the internal consistency of CAB-SR (Farran et al., 2011) was rated as "inadequate". CSES (Zeiss et al., 1999) received a "doubtful" methodological quality rating as it did not describe how to handle missing values. The Spanish versions of Revised Scale for Caregiving Self-Efficacy (RSCSE) (Márquez-González et al., 2009), AMSS (Huang et al., 2003), and Sense of Competence Questionnaire (SCQ) (Jansen et al., 2007) received a "doubtful" methodological quality rating because it was unclear whether the scales were unidimensional.

3.3.3.2 Stability. The instrument's stability was reported in six studies (Cheon et al., 2022; Farran et al., 2011; Piggott et al., 2017; Steffen et al., 2002; Zeiss et al., 1999; Zhang et al., 2013) and was not considered in the remaining five studies. Five studies (Farran et al., 2011; Piggott et al., 2017; Steffen et al., 2002; Zeiss et al., 1999; Zhang et al., 2013) reported Pearson or Spearman correlation coefficients, and CCSM (Piggott et al., 2017) reported a sub-scale inter-class correlation coefficient (ICC) <0.7. Therefore, the methodological quality evaluation was "doubtful", and the stability of the scale was "indeterminate". In contrast, CS-MBPSD (Cheon et al., 2022) had an ICC of 0.781. Accordingly, the methodological quality evaluation was "adequate", and the stability of the scale was "sufficient".

3.3.4 Other Measurement Properties

Measurement Error and Responsiveness. Measurement error was reported only for the CS-MBPSD (Cheon et al., 2022), which provided standard error of measurement values but not the smallest detectable change, leading to "insufficient" measurement quality. Responsiveness was reported only for the AMSS (Huang et al., 2003), but the sample size was insufficient, resulting in "doubtful" methodological quality and "indeterminate" measurement quality.

3.4 Evidence Level Evaluation and Evidence Recommendation

Only CS-MBPSD (Cheon et al., 2022) had "sufficient" content validity; the remaining ten scales were rated "indeterminate". The structural validity of the Family caregiver self-efficacy (FCSE) (Fortinsky et al., 2002), Self-Efficacy Questionnaire for Chinese Family Caregivers (SEQCFC) (Zhang et al., 2013), CES (Crellin et al., 2014), and CCSM (Piggott et al., 2017) was "indeterminate", whereas for the other four scales it was rated as "insufficient". The internal consistency of FCSE (Fortinsky et al., 2002), RSCSE (Steffen et al., 2002), SEQCFC (Zhang et al., 2013), CES (Crellin et al., 2014), CCSM (Piggott et al., 2017), and CS-MBPSD (Cheon et al., 2022) was rated as "sufficient". The internal consistency of the CAB-SR (Farran et al., 2011) was rated as "insufficient", and for the remaining four scales, it was rated as "indeterminate". The reliability of CS-MBPSD (Cheon et al., 2022) was rated as "sufficient", and for the remaining five scales was rated as "indeterminate". The hypothesis testing of FCSE (Fortinsky et al., 2002), RSCSE (Steffen et al., 2002), SEQCFC (Zhang et al., 2013), CES (Crellin et al., 2014), and CS-MBPSD (Cheon et al., 2022) was rated as "sufficient". The internal consistency of the Spanish version of RSCSE (Márquez-González et al., 2009), SCQ (Jansen et al., 2007), CAB-SR (Farran et al., 2011), and CCSM (Piggott et al., 2017) was rated as "insufficient", and for the remaining one scale was rated as "indeterminate".

The overall quality of evidence for the measurement properties of the scales was primarily moderate to low, as indicated by the synthesized evidence across all measurement properties, which was influenced by the risk of bias as a downgrading factor. The recommendation grade for all 11 scales was Category B. This suggests that these scales

Table 2. Evidence synthesis and recommendation categories for each instrument for assessing caregiver competence in BPSD management (*n* = 11).

PROMs	Content validity	Structural validity	Internal consistency	Reliability	Measurement error	Hypothesis testing	Criterion validity	Responsiveness	Cross-cultural validity	Category of the PROM
CSES	Low	NA/NA	Low	Low	NA/NA	Low	NA/NA	NA/NA	NA/NA	B
FCSE	Low	Moderate	High	NA/NA	NA/NA	Moderate	NA/NA	NA/NA	NA/NA	B
RSCSE	Moderate	High	High	Low	NA/NA	High	NA/NA	NA/NA	NA/NA	B
Spanish Version of the RSCSE	Moderate	NA/NA	Low	NA/NA	NA/NA	Moderate	NA/NA	NA/NA	NA/NA	B
AMSS	Low	NA/NA	Very low	NA/NA	NA/NA	NA/NA	NA/NA	Very low	NA/NA	B
SCQ	Very low	Very low	Very low	NA/NA	NA/NA	Low	NA/NA	NA/NA	NA/NA	B
CAB-SR	Low	Very low	Very low	Very low	NA/NA	Moderate	NA/NA	NA/NA	NA/NA	B
SEQCFC	Moderate	High	High	Low	NA/NA	High	NA/NA	NA/NA	NA/NA	B
CES	Low	Moderate	High	NA/NA	NA/NA	High	NA/NA	NA/NA	NA/NA	B
CCSM	Low	Moderate	High	Low	NA/NA	High	NA/NA	NA/NA	NA/NA	B
CS-MBPSD	Moderate	High	High	Moderate	Moderate	Moderate	NA/NA	NA/NA	NA/NA	B

can be recommended for use, but further validation studies are required. Evidence synthesis and recommendation categories for each instrument for assessing caregiver competence in BPSD management are presented in Table 2.

4. Discussion

4.1 Misalignment Between Assessment Focus and Care Needs

This study synthesized findings from 11 studies that utilized assessment scales to evaluate family caregivers' competence in managing BPSD. These scales were developed for dementia caregivers. However, the psychometric evidence supporting them was insufficient to address the unique complexities of dementia care.

Through a comprehensive and systematic evaluation of the included studies, we found that most instruments focused on assessing caregivers' general "self-efficacy" or "confidence" rather than directly measuring their specific competence in specifically managing BPSD. Similarly, a systematic review of nonpharmacological interventions for BPSD management emphasized that although caregiver confidence was crucial for intervention success, validated tools specifically targeting BPSD management competence remain scarce (Lyons and Champion, 2022). Only a few scales, including CS-MBPSD (Cheon et al., 2022) and AMSS (Huang et al., 2003), were specifically designed to target BPSD management skills.

The disease trajectory of dementia is characterized by progressive cognitive decline and complex, fluctuating behavioral symptoms. These symptoms impose highly specific skill requirements on caregivers, including managing wandering, aggression, and sundowning. The extensive prevalence of "doubtful" content validity in existing scales suggested that their items inadequately covered real, concrete, and challenging scenarios in dementia care. Consequently, the validity of using scales based on the self-efficacy construct to predict or evaluate the actual caregiving competence and outcomes of dementia caregivers remains uncertain, as these tools may fail to capture the nuanced, symptom-specific skills required for effective BPSD management.

4.2 The Methodological Quality of Assessment Tools of Family-Caregiver Competence for Managing BPSD Needs to Be Improved

All included scales received only a "doubtful" rating for their development methodology, highlighting a fundamental limitation in the transparency and rigor of the underlying evidence. Rigorous PROM development is the cornerstone for ensuring content relevance and conceptual appropriateness. The prevalent deficiencies highlighted here suggested that future development or revision of scales must strictly adhere to established guidelines, thereby improving transparency in reporting the conceptual framework, qualitative interviews, Delphi expert consultations, and cognitive debriefing processes (Mokkink et al., 2024; Suo et al., 2026).

According to COSMIN guidelines, content validity is the most fundamental measurement property of an assessment tool (Mokkink et al., 2018). Guidelines recommend that both experts and target patients assess scale items for relevance, comprehensiveness, and comprehensibility. However, most included studies obtained input from only one of these groups opinions. CAB-SR (Farran et al., 2011), CCSM (Piggott et al., 2017), and CS-MBPSD (Cheon et al., 2022) consulted both experts and patients. In the future, to improve the content validity of the scale, studies should use in-depth or focus-group interviews to solicit participants' understanding, views, and attitudes toward the scale items (Mokkink et al., 2025). Empirical studies on dementia care have shown that cognitive interviews and qualitative feedback from caregivers can effectively identify unclear wording, inappropriate assumptions, and missing content, thereby supporting scale refinement and strengthening content validity evidence (Mokkink et al., 2024; Zhou et al., 2025).

The included studies primarily used factor analysis to determine the scales' structural validity. However, certain studies were limited to EFA and did not proceed to CFA. According to COSMIN guidelines, CFA is preferred over EFA. Consequently, it is advisable to apply EFA to clarify the internal structure of scales that lack a strong theoretical foundation, and subsequently apply CFA to more thoroughly investigate the relationships between items and factors (Li et al., 2025). Researchers should formulate hypotheses about scale correlations when conducting structural validity testing to avoid biased conclusions.

Regarding scale reliability, internal consistency refers to the degree of interrelation among items within a unidimensional scale or subscale. All included studies calculated Cronbach's α. However, the Spanish versions of RSCSE (Márquez-González et al., 2009), AMSS (Huang et al., 2003), and SCQ (Jansen et al., 2007) received a "doubtful" methodological quality rating because it was unclear whether the scales were unidimensional. Consequently, ensuring the scale's unidimensionality is a primary requirement. To ensure that the full scale and its individual dimensions exhibit satisfactory reliability, researchers should compute and report Cronbach's α for each subscale, in addition to calculating the overall internal consistency coefficient for the full scale after the scale's unidimensionality has been established.

The stability of the included tools was rated as "doubtful" or "inadequate". According to COSMIN guidelines, a two week retest interval is recommended, and evidence should confirm that participants' life circumstances and the measured constructs remain stable between the two assessments (Mokkink et al., 2024). Stability of the tool was reported in six studies (Cheon et al., 2022; Farran et al., 2011; Piggott et al., 2017; Steffen et al., 2002; Zeiss et al.,

1999; Zhang et al., 2013), but only CCSM (Piggott et al., 2017) and CS-MBPSD (Cheon et al., 2022) reported the ICC. The retest time of CCSM (Piggott et al., 2017) was only two days, whereas for CAB-SR (Farran et al., 2011) it was 4–6 weeks. Given that patient requirements are dynamic and open to change, frequently varying with environmental changes, evaluators of scales should clearly justify the selected interval between repeated measurements. Currently, retest time is frequently determined by subjective judgment based on experience, which can, to some extent, introduce bias into the evaluation of test-retest reliability. Future studies should justify the retest interval and provide evidence that the sample conditions remained stable, thereby supporting the validity of the stability estimates.

COSMIN guidelines specify that a gold standard is not usually available for Patient-Reported Outcome Measure (PROMs). The only recognized exception is when developing a simplified tool, in which the original instrument can serve as the standard. However, CES (Crellin et al., 2014) was compared with RSCSE subscales, which did not meet the criteria. Consequently, using a widely used scale as a benchmark for a new scale in research should be considered a hypothesis-testing exercise rather than a test of criterion validity (Terwee et al., 2012).

None of the 11 scales included in this review reported cross-cultural validity. Those findings indicated that if existing tools were applied to caregivers from different ethnic or linguistic backgrounds without cross-cultural validation, measurement bias may occur, compromising the accuracy of the assessments and the validity of cross-group comparisons. To fill the gap in cross-cultural-validity evidence, future research should systematically undertake cross-cultural adaptation and validation (Yu et al., 2025). A practical framework grounded in ecological validity theory and implementation science, organized the cultural adaptation process into five core dimensions. That framework has been applied to systematically adapt an evidence-based intervention for Chinese- and Korean-American dementia caregivers (O'Donald and Calia, 2025).

4.3 Selection of Assessment Tools for Family Caregiver Competence for Managing BPSD

All 11 tools received a B-grade recommendation. Notably, COSMIN considers content validity the most important measurement property. CS-MBPSD (Cheon et al., 2022) received an "adequate" methodological rating and a moderate evidence rating for content validity. Furthermore, as an assessment tool specifically developed to evaluate family caregivers' competence in managing BPSD, the CS-MBPSD (Cheon et al., 2022) provided a comprehensive assessment of caregivers' abilities across six dimensions, enabling precise identification of their specific weaknesses in areas such as knowledge, skills, attitudes, and self-care. Based on the scores obtained in each dimension, the tool can guide the development of differentiated support strategies. To date, the CS-MBPSD has been validated for use with both family caregivers and professional care workers, providing a scientific instrument to improve the quality of life of both caregivers and care recipients (Song et al., 2026). Among the 11 included instruments, CS-MBPSD demonstrated the strongest content validity evidence. Therefore, in accordance with COSMIN guidance for selecting among multiple 'B' category instruments, CS-MBPSD is provisionally recommended for clinical and research use, pending further validation of its responsiveness and cross-cultural validity.

This review identified a critical issue: all included instruments share fundamental limitations. These include a lack of transparency in the development process, insufficient validation of content validity, and a general absence of evidence for responsiveness and cross-cultural validity. Consequently, all reviewed tools were classified as Category B recommendations, indicating they should be implemented with caution in clinical or research settings and that further high-quality validation studies are urgently required. Future research should prioritize the systematic development of scales with rigorous content validity testing and conduct extensive evaluations of their measurement properties across diverse cultural contexts, with a particular focus on responsiveness. The development or optimization of caregiver competence assessment tools specifically focused on dementia is an urgent requirement for future research. This process should adhere strictly to COSMIN methodological standards. During the scale-development phase, systematic incorporation of the perspectives of multiple stakeholders is essential, including dementia caregivers, geriatric psychiatric nurses, and geriatric medicine specialists. That ensures that the instrument items comprehensively and accurately represent the full spectrum of challenges across caregiving contexts from the early to late stages of the disease. Furthermore, conducting systematic psychometric comparative studies across various cultural and healthcare systems is essential (Lillekroken et al., 2025).

4.4 Translating Caregiver-Competence Assessment Tools Into Practice

Currently, most caregiver-competence assessment tools remain at the stage of providing dimensional profiles without embedded intervention guidance. The key limitations are as follows: the absence of standardized protocols for translating subscale scores into actionable care plans; inadequate integration of assessment findings with evidence-based intervention resources; and minimal application of digital technologies to automate the assessment-to-intervention pathway (Biernetzky et al., 2025; Rosen et al., 2025). These shortcomings have substantially constrained the practical utility of existing instruments in clinical and community settings. The team led by Malhotra developed a theory-driven, multi-component mobile

application, "CareBuddy", specifically designed for caregivers of individuals with dementia. The main functions of the mobile application included personalized assessment and tailored solutions, an artificial intelligence-driven chatbot, GPS location monitoring, peer support, helplines, telemedicine, caregiver self-care resources, and integration with healthcare providers (Malhotra et al., 2025).

Future efforts should focus on developing integrated assessment-intervention tools that embed not only assessment items but also score-based actionable recommendations. For example, when a score on a particular dimension falls below a threshold, the tool could automatically deliver relevant micro-courses and skills videos. Such a design would enable integration of assessment and intervention, reducing the costs of translating findings into practice (Malhotra et al., 2025).

4.5 Limitations

This study had certain limitations: (1) The assessment of some COSMIN criteria involved a degree of subjectivity, which could have introduced bias into the evaluation results. (2) Most studies did not report measurement error or responsiveness, which limited the ability to assess whether the scales could precisely measure subtle changes in the competence of dementia caregivers. (3) The included scales were primarily applied to dementia samples; however, the high heterogeneity in the samples of certain studies could have diluted information specific to the care requirements of dementia. Future studies should explicitly recruit underrepresented populations to enhance the generalizability of existing tools. (4) The use of only English-language literature limited the generalizability of the results and may have introduced language bias. Although English is the predominant language of scientific communication in dementia care research, important validation studies and culturally adapted tools may have been published in other languages. Consequently, this review may have missed instruments that were reported exclusively in non-English journals. However, we noted that the two most pertinent instruments developed in non-English settings, the Korean CS-MBPSD and the Chinese SEQCFC, were included in this review because both were published in English-language journals and thus were captured by our search. To fully address language bias, future reviews should include non-English databases such as CNKI, WanFang Data, and RISS, and non-English reports should be considered for inclusion using available translation resources.

5. Conclusions

In this study, the measurement characteristics of family caregiver competence in managing BPSD were comprehensively evaluated using COSMIN guidelines. Of 11 instruments studied, CS-MBPSD was methodologically rated as "adequate" and received a moderate evidence rating for its content validity. The scale is suitable for provisional use due to its specific focus on BPSD management. While the instrument may be used for screening or exploratory work, researchers are advised to conduct concurrent validation with established measures and, where possible, collect local psychometric data to confirm the factor structure and reliability within their specific sample. Future research should both adhere to COSMIN guidelines and integrate the target country's cultural characteristics and healthcare context to develop a scale supported by high-quality evidence for clinical use.

Abbreviations

BPSD, behavioral and psychological symptoms of dementia; COSMIN, COnsensus-based Standards for the selection of health Measurement Instruments; PROMs, Patient-Reported Outcome Measures; CFA, confirmatory factor analysis; EFA, exploratory factor analysis; CSES, Caregiver Self-Efficacy Scales; FCSE, Family caregiver self-efficacy; RSCSE, Revised Scale for Caregiving Self-Efficacy; Spanish Version of the RSCSE, Spanish Version of the Revised Scale for Caregiving Self-Efficacy; AMSS, Agitation Management Self-Efficacy Scale; SCQ, Sense of Competence Questionnaire; CAB-SR, Caregiver Assessment of Behavioral Skill-Self-Report; SEQCFC, Self-Efficacy Questionnaire for Chinese Family Caregivers; CES, Caregiver Efficacy Scale; CCSM, Caregiver Confidence in Medical Sign/Symptom Management scale; CS-MBPSD, Competence Scale in Managing Behavioral and Psychological Symptoms of Dementia; Spanish Version of the RSCSE, Spanish Version of the Revised Scale for Caregiving Self-Efficacy; SEM, structural equation modeling.

Availability of Data and Materials

All data generated or analysed during this study are included in this published article.

Author Contributions

JD, YW, YY, YH, YM, and YF: conceptualization, methodology, validation, data analysis, and writing – review and editing. JD and YW: data curation. YY: supervision and project administration. All authors contributed to the critical revision of the manuscript for important intellectual content and approved the final version to be submitted. All authors have participated sufficiently in the work and agreed to be accountable for all aspects of the work.

Ethics Approval and Consent to Participate

Not applicable.

Acknowledgment

The authors thank all participants that took part in this study.

Funding

This study was internally funded by Medical Scientific Research Foundation of Zhejiang Province, China (Grant Number 2024KY125).

Conflicts of Interest

The authors declare no conflicts of interest.

Declaration of AI and AI-Assisted Technologies in the Writing Process

During the preparation of this work, the authors used DeepSeek-V4 in order to check spelling and grammar. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

Supplementary Material

Supplementary material associated with this article can be found, in the online version, at https://doi.org/10.31083/BP52713.

References

Aguirre A, Benge J, Rovianne R. Assessing the Individual Needs and Goals of Dementia Caregivers: Development of the COPE-ING Measure. Innovation in Aging. 2025; 9: igaf122.1568. https://doi.org/10.1093/geroni/igaf122.1568

Al Mokdad Z, Msheik A, Tarabey L, Abou-Mrad F, Fadel P. Ethical dilemmas in the care of patients with Alzheimer's disease and related dementias unable to give informed consent: positioning Lebanon within the Global North-South context. BMC Medical Ethics. 2025; 26: 152. https://doi.org/10.1186/s12910-025-01277-3

Alhazzani W, Guyatt G. An overview of the GRADE approach and a peek at the future. The Medical Journal of Australia. 2018; 209: 291–292. https://doi.org/10.5694/mja18.00012

Alzheimer's Society. Survey report: Lived experiences of dementia 2025. 2025. Available at: https://www.alzheimers.org.uk/about-us/policy-and-influencing/reports/survey-lived-experiences-dementia (Accessed: 6 June 2026).

Anantapong K, Jiraphan A, Aunjitsakul W, Sathaporn K, Werachattawan N, Teetharatkul T, et al. Behavioural and psychological symptoms of people with dementia in acute hospital settings: a systematic review and meta-analysis. Age and Ageing. 2025; 54: afaf013. https://doi.org/10.1093/ageing/afaf013

Biernetzky OA, Thyrian JR, Boekholt M, Berndt M, Hoffmann W, Teipel SJ, et al. Identifying Unmet Needs of Informal Dementia Caregivers in Clinical Practice: User-Centered Development of a Digital Assessment Tool. JMIR Aging. 2025; 8: e59942. https://doi.org/10.2196/59942

Budgett J, Sommerlad A, Kupeli N, Zabihi S, Olsen A, Cooper C. Setting individualised goals for people living with dementia and their family carers: A systematic review of goal-setting outcome measures and their psychometric properties. Dementia (London, England). 2024; 23: 312–340. https://doi.org/10.1177/14713012231222309

Cebolla Sousa A, Greenfield G, Nair P, Aldakhil R, Udoyeh J, Karki M, et al. Effectiveness of Telephone Interventions for the Management of Behavioral and Psychological Symptoms of Dementia in the Community: Systematic Review. Journal of Medical Internet Research. 2025; 27: e77233. https://doi.org/10.2196/77233

Cheng Y, Bai R, Shan S, Zhao X, Xia C. Childbirth experience assessment tools based on COSMIN guidelines: A systematic review. International Journal of Nursing Sciences. 2025; 12: 89–95. https://doi.org/10.1016/j.ijnss.2024.12.001

Cheon H, Song JA, Kim J, Jung S. The Competence Scale in Managing Behavioral and Psychological Symptoms of Dementia (CS-MBPSD) for family caregivers: Instrument development and cross-sectional validation study. International Journal of Nursing Studies. 2022; 129: 104205. https://doi.org/10.1016/j.ijnurstu.2022.104205

Crellin N, Charlesworth G, Orrell M. Measuring family caregiver efficacy for managing behavioral and psychological symptoms in dementia: a psychometric evaluation. International Psychogeriatrics. 2014; 26: 93–103. https://doi.org/10.1017/S1041610213001646

Ding X, Alavi Toussi R, Dal Pizzol FLF, Grewal A, Hyde A, Parmar J, et al. Mapping Caregiver Needs' Assessment Tools for Family and Friend Caregivers: A Rapid Scoping Review. International Journal of Environmental Research and Public Health. 2026; 23: 300. https://doi.org/10.3390/ijerph23030300

Farran CJ, Fogg LG, McCann JJ, Etkin C, Dong X, Barnes LL. Assessing family caregiver skill in managing behavioral symptoms of Alzheimer's disease. Aging & Mental Health. 2011; 15: 510–521. https://doi.org/10.1080/13607863.2010.536140

Fortinsky RH, Kercher K, Burant CJ. Measurement and correlates of family caregiver self-efficacy for managing dementia. Aging & Mental Health. 2002; 6: 153–160. https://doi.org/10.1080/13607860220126763

GBD 2021 Nervous System Disorders Collaborators. Global, regional, and national burden of disorders affecting the nervous system, 1990-2021: a systematic analysis for the Global Burden of Disease Study 2021. The Lancet. Neurology. 2024; 23: 344–381. https://doi.org/10.1016/S1474-4422(24)00038-3

Gonçalves-Pereira M, Marques MJ, Alves RF, Jelley H, Wolfs C, Meyer G, et al. Sense of coherence, subjective burden, and anxiety and depression symptoms in caregivers of people with dementia: Causal dynamics unveiled by a longitudinal cohort study in Europe. Journal of Affective Disorders. 2025; 373: 1–11. https://doi.org/10.1016/j.jad.2024.12.078

Guyatt G, Agoritsas T, Brignardello-Petersen R, Mustafa RA, Rylance J, Foroutan F, et al. Core GRADE 1: overview of the Core GRADE approach. BMJ (Clinical Research Ed.). 2025; 389: e081903. https://doi.org/10.1136/bmj-2024-081903

Hu R, Zhang F, Zhang Y, Lai B, Ma W, Xu X, et al. Development and Validation of a Knowledge Scale About the Behav-

ioral and Psychological Symptoms of Dementia (KS-BPSD) Among Chinese Formal Caregivers. Clinical Gerontologist. 2023; 46: 599–607. https://doi.org/10.1080/07317115.2022.2111290

Huang HL, Shyu YIL, Chen MC, Chen ST, Lin LC. A pilot study on a home-based caregiver training program for improving caregiver self-efficacy and decreasing the behavioral problems of elders with dementia in Taiwan. International Journal of Geriatric Psychiatry. 2003; 18: 337–345. https://doi.org/10.1002/gps.835

Jansen APD, van Hout HPJ, van Marwijk HWJ, Nijpels G, Gundy C, Vernooij-Dassen MJFJ, et al. Sense of competence questionnaire among informal caregivers of older adults with dementia symptoms: a psychometric evaluation. Clinical Practice and Epidemiology in Mental Health: CP & EMH. 2007; 3: 11. https://doi.org/10.1186/1745-0179-3-11

Jung S, Song JA, Kim J, Cheon H, Kim J. Family caregiver competence in managing behavioral and psychological symptoms of dementia: A concept synthesis. Japan Journal of Nursing Science: JJNS. 2022; 19: e12462. https://doi.org/10.1111/jjns.12462

Li Y, Ouyang H, Lin G, Peng Y, Yao J, Chen Y. Evaluation of the measurement properties of online health information quality assessment tools: A systematic review. International Journal of Nursing Sciences. 2025; 12: 130–136. https://doi.org/10.1016/j.ijnss.2025.02.015

Lillekroken D, Bjørge H, Halvorsrud L, Lidal IB. Assessing quality of life-a scoping review of studies presenting quality of life instruments for informal caregivers of persons with dementia. BMC Geriatrics. 2025; 25: 976. https://doi.org/10.1186/s12877-025-06455-x

Lillekroken D, Halvorsrud L, Bjørge H, Gandhi S, Sivakumar PT, Goyal AR. Caregivers' experiences, challenges, and needs in caring for people with dementia in India: a scoping review. BMC Health Services Research. 2024; 24: 1661. https://doi.org/10.1186/s12913-024-12146-x

Lim J, Andel R, Puga F, Aranda MP, Rivera-Hernandez M, Dávila-Roman AL, et al. Burden and well-being among dementia caregivers in Puerto Rico: the role of behavioral and psychological symptoms of dementia. The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences. 2026; 81: gbag014. https://doi.org/10.1093/geronb/gbag014

Lyons TL, Champion JD. Nonpharmacological Interventions for Management of Behavioral and Psychological Symptoms of Dementia in Long-Term Care Facilities by Direct Caregivers: A Systematic Review. Journal of Gerontological Nursing. 2022; 48: 18–23. https://doi.org/10.3928/00989134-20220606-03

Malhotra C, Yue Y, Ramakrishnan C, Shah S, Chen W, Yap P, et al. Supporting Dementia Caregiving With a Mobile Care Ecosystem: Development and Mixed Methods Study. JMIR Aging. 2025; 8: e78759. https://doi.org/10.2196/78759

Mokkink LB, Boers M, van der Vleuten CPM, Bouter LM, Alonso J, Patrick DL, et al. COSMIN Risk of Bias tool to assess the quality of studies on reliability or measurement error of outcome measurement instruments: a Delphi study. BMC Medical Research Methodology. 2020; 20: 293. https://doi.org/10.1186/s12874-020-01179-5

Mokkink LB, de Vet HCW, Prinsen CAC, Patrick DL, Alonso J, Bouter LM, et al. COSMIN Risk of Bias checklist for systematic reviews of Patient-Reported Outcome Measures. Quality of Life Research: an International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation. 2018; 27: 1171–1179. https://doi.org/10.1007/s11136-017-1765-4

Mokkink LB, Elsman EBM, Terwee CB. COSMIN guideline for systematic reviews of patient-reported outcome measures version 2.0. Quality of Life Research: an International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation. 2024; 33: 2929–2939. https://doi.org/10.1007/s11136-024-03761-6

Mokkink LB, Herbelet S, Tuinman PR, Terwee CB. Content validity: judging the relevance, comprehensiveness, and comprehensibility of an outcome measurement instrument - a COSMIN perspective. Journal of Clinical Epidemiology. 2025; 185: 111879. https://doi.org/10.1016/j.jclinepi.2025.111879

Márquez-González M, Losada A, López J, Peñacoba C. Reliability and Validity of the Spanish Version of the Revised Scale for Caregiving Self-Efficacy. Clinical Gerontologist. 2009; 32: 347–357. https://doi.org/10.1080/07317110903110419

Mukherjee U, Brownell M, Reddy PH. Aging, dementia, and care models: Global perspectives with insights from India. Ageing Research Reviews. 2026; 113: 102941. https://doi.org/10.1016/j.arr.2025.102941

O'Donald F, Calia C. The Process of Translation and Cross-Cultural Adaptation of Functional Assessment Tools for Dementia: A Systematized Review. Health Science Reports. 2025; 8: e70289. https://doi.org/10.1002/hsr2.70289

Peng X, Li RC, Liang T, Xu SF, Cao Y. Critical evaluation of COSMIN scores in scales for mild cognitive impairment and Alzheimer's disease: A comprehensive review. Journal of Alzheimer's Disease Reports. 2024; 8: 1596–1610. https://doi.org/10.1177/25424823241299023

Peters MDJ, Godfrey CM, Khalil H, McInerney P, Parker D, Soares CB. Guidance for conducting systematic scoping reviews. International Journal of Evidence-based Healthcare. 2015; 13: 141–146. https://doi.org/10.1097/XEB.0000000000000050

Pickering CEZ, Winstead V, Yildiz M, Wang D, Yefimova M, Pickering AM. Subsyndromes and symptom clusters: Multilevel factor analysis of behavioral and psychological symptoms of dementia with intensive longitudinal data. Alzheimer's & Dementia: the Journal of the Alzheimer's Association. 2024; 20: 6699–6708. https://doi.org/10.1002/alz.14075

Piggott CA, Zimmerman S, Reed D, Sloane PD. Development and Testing of a Measure of Caregiver Confidence in Medical Sign/Symptom Management. American Journal of Alzheimer's Disease and other Dementias. 2017; 32: 373–381. https://doi.org/10.1177/1533317517711247

Pozzi FE, D'Antonio F, Zuffi M, Pelati O, Vernè D, Panigutti

M, et al. Italian standardization of the BPSD-SINDEM scale for the assessment of neuropsychiatric symptoms in persons with dementia. Frontiers in Neurology. 2024; 15: 1455787. https://doi.org/10.3389/fneur.2024.1455787

Ramakrishnan C, Malhotra C. Challenges to accessing community dementia care services: a qualitative study. BMC Health Services Research. 2025; 25: 747. https://doi.org/10.1186/s12913-025-12895-3

Ren X, Dong X, Tao Y, Chen X, Lu X, Bai D, et al. Instruments for measuring clinical reasoning competence in current and future nurses based on COSMIN: A systematic review. Nurse Education in Practice. 2025; 87: 104474. https://doi.org/10.1016/j.nepr.2025.104474

Rosen T, Shaw A, Elman A, Baek D, Gottesman E, Park S, et al. Developing a screening tool and intervention strategy for elder neglect in persons with dementia in primary care: protocol to use a multistep process. BMJ Open. 2025; 15: e085172. https://doi.org/10.1136/bmjopen-2024-085172

Song JA, Cheon H, Yang Y, Kim J, Lee Y, Oh S, et al. Psychometric Evaluation of the Competence Scale in Managing Behavioural and Psychological Symptoms of Dementia (CSMBPSD) for Care Workers. International Journal of Mental Health Nursing. 2026; 35: e70227. https://doi.org/10.1111/inm.70227

Steffen AM, McKibbin C, Zeiss AM, Gallagher-Thompson D, Bandura A. The revised scale for caregiving self-efficacy: reliability and validity studies. The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences. 2002; 57: P74–P86. https://doi.org/10.1093/geronb/57.1.p74

Suo Y, Hou Z, Xu H, Zhou H, Wang X. Evaluation of self-efficacy of nurses in symptom management practices: a systematic review of psychometric properties. BMC Nursing. 2026; 25: 141. https://doi.org/10.1186/s12912-026-04308-5

Terwee CB, Mokkink LB, Knol DL, Ostelo RWJG, Bouter LM, de Vet HCW. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. Quality of Life Research: an International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation. 2012; 21: 651–657. https://doi.org/10.1007/s11136-011-9960-1

Watt JA, Porter J, Tavilsup P, Chowdhury M, Hatch S, Ismail Z, et al. Guideline Recommendations on Behavioral and Psychological Symptoms of Dementia: A Systematic Review. Journal of the American Medical Directors Association. 2024; 25: 837–846.e21. https://doi.org/10.1016/j.jamda.2024.03.007

World Health Organization. World failing to address dementia challenge. 2021. Available at: https://www.who.int/news/item/02-09-2021-world-failing-to-address-dementia-challenge (Accessed: 16 May 2026).

Yin Z, Li Y, Bao Q, Zhang X, Xia M, Zhong W, et al. Comparative efficacy of multiple non-pharmacological interventions for behavioural and psychological symptoms of dementia: A network meta-analysis of randomised controlled trials. International Journal of Mental Health Nursing. 2024; 33: 487–504. https://doi.org/10.1111/inm.13254

Yu K, Ding X, Wang X, Li H, Zhai J, Shi Y. Instruments to Assess Dementia Fear Among Older Adults: A Systematic Review of Measurement Properties. Psychogeriatrics: the Official Journal of the Japanese Psychogeriatric Society. 2025; 25: e70064. https://doi.org/10.1111/psyg.70064

Zeiss AM, Gallagher-Thompson D, Lovett S, Rose J, McKibbin C. Self-Efficacy as a Mediator of Caregiver Coping: Development and Testing of an Assessment Model. Journal of Clinical Geropsychology. 1999; 5: 221–230. https://doi.org/10.1023/a:1022955817074

Zhang SY, Edwards H, Yates P, Ruth E, Guo Q. Preliminary reliability and validity testing of a self-efficacy questionnaire for Chinese family caregivers. Aging & Mental Health. 2013; 17: 630–637. https://doi.org/10.1080/13607863.2013.771615

Zhou Y, Zeng X, Kilcrease A, Meah N, Choi S, Phelan E. Measuring Dementia Care Partners' Fall Risk Management Behaviors: Development and Content Validation. Innovation in Aging. 2025; 9: igaf122.1570. https://doi.org/10.1093/geroni/igaf122.1570