Alzheimer's Association's Evidence-based Clinical Practice Guideline on the Use of Cognitive Tests for the Early Detection of Cognitive Impairment in Older Adults in Primary Care

What is the ask:

- **Panel recommendations and remarks (Table 1)**: Please review the information starting on Page 2. Use the **online form** to provide feedback on the content or presentation of what are to be the recommendations and associated remarks contained in the green sections in **Table 1**. **Overall, we wish to understand if you believe the recommendations are 1) Clear and 2) Actionable and 3) If not, please provide suggestions for how to improve their usefulness for clinical decision-making.** Your diverse perspectives are essential to ensuring the recommendations are practical, patient-centered, and reflective of real-world experiences. We have also provided a legend (Table 2) informing the interpretation and implementation of these draft recommendations by various users.

- Additional contextual information (Pages 6-10): We briefly describe the overview of the guideline development process, including systematic review methodology. In addition to finalized recommendations and remarks, a full reporting of panel disclosures, summary of findings tables, and methods will be submitted to a scientific journal and peer-reviewed by external reviewers before approval for publication.

Who should comment:

- Clinicians across all disciplines and specialities, researchers**,** patients, caregivers, and family members, patient advocates, health system representatives, healthcare administrators, policy-makers, and any individual or organization with an interest or expertise in this topic can comment.

- If multiple individuals within the same organization/agency wish to provide feedback, we strongly encourage submitting a *single, comprehensive, coordinated response* that integrates all perspectives. This helps ensure clarity and coherence for panel review.

How your comments will be used:

- The methods team and guideline panel will review all feedback received during the public comment period **(June 30 - July 11, 5 p.m. CDT)**. Comments that are within the scope of the guideline question *and supported* by the available evidence will be considered for incorporation into the final guidance. Revisions may be made to improve accuracy, clarity, or applicability.

- Following the publication of the final manuscript, all comments—de-identified where possible—will be made publicly available to promote transparency and acknowledge the contributions of collaborative parties.

Please scroll down to review an overview of the project and recommendations and remarks.

Overview of project:

Background: In 2013, the Alzheimer's Association published recommendations to support cognitive assessment during the Medicare Annual Wellness Visit, offering guidance on workflow, tool selection, and follow-up. Since then, advances in cognitive science, new disease-modifying therapies for AD, and increased emphasis on early detection prompted the Association to update these recommendations using the GRADE methodology. In Spring 2024, the Alzheimer's Association convened a guideline panel of clinical and subject-matter experts to develop an evidence-based clinical practice guideline on the use of cognitive tests in older adults in primary care. In collaboration with systematic review and guideline methodologists, the guideline panel developed the scope, purpose, target audience, and clinical questions for the guideline. In Summer 2025, the panel formulated draft evidence-based recommendations, now available for public comment, and are preparing manuscripts for submission to peer-reviewed journals.

Scope: The scope of this guideline focuses on patients 55 years and older presenting to primary care and appropriate ambulatory settings (i.e. geriatric medicine, psychiatry) in the United States. The recommendations do not apply to specialty care settings such as memory clinics or to patients who already have an established diagnosis of cognitive impairment.

Only index tests that are freely available and take 15 minutes or less to administer were included, and only English and Spanish versions were evaluated. Digital tests were excluded. Importantly, the target condition was defined as any cognitive impairment (vs. no cognitive impairment), including mild cognitive impairment and dementia.

Included tests:

- 5-Cog
- Eight-item Informant Interview to Differentiate Aging and Dementia (AD8)
- General Practitioner assessment of Cognition (GPCOG)
- Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE) (including short version)
- Mini-Cog
- Memory Impairment Screen (MIS)
- Montreal Cognitive Assessment (MoCA) (including short version)
- Quick Dementia Rating System (QDRS)
- Rowland Universal Dementia Assessment Scale (RUDAS)
- Saint Louis University Mental Status (SLUMS)

Methodology: The Alzheimer's Association's methodological team followed the GRADE approach and the Cochrane Handbook for Diagnostic Test Accuracy to synthesize evidence (search conducted between January 1, 1999, to August 13, 2024), assess the certainty of the evidence, move from evidence to decisions, draft recommendations, and assign the strength of recommendations. A priori panel decisions included: development of clinical questions in PICO format, included index tests and reference standards, statistical plan for meta-analysis, and clinical thresholds for decision-making.

When discussing the body of evidence and drafting recommendations, the panel was blinded to all cognitive assessment names by using placeholders (e.g., Test 1, Test 2, etc.). The panel considered the benefits of true positive and true negative and the harms of false positive and false negative test results. A greater emphasis was placed on maximizing true positives (as opposed to true negatives) and minimizing false negatives (as opposed to false positives), as the intent is early identification and

differentiation of any cognitive impairment or related conditions. Methodologists managed conflicts of interest using predetermined rules set by the Alzheimer's Association to minimize bias.

Recommendations: The panel drafted recommendations that address which cognitive test(s) to use in English and Spanish-speaking adults aged 55-years and older presenting to primary and appropriate ambulatory care settings (i.e. geriatric medicine, psychiatry), in the United States including those aged 65+ as part of an Medicare Annual Wellness Visit (Table 1). A guide for interpreting the certainty of the evidence and strength of recommendations can be found in Table 2.

Table 1. *Clinical question and recommendations and remarks for decision-making.*

Clinical question (closed for comment)
In English- and Spanish-speaking adults aged 55 years and older presenting to primary and appropriate ambulatory care settings, which test (i.e., 5-Cog, AD8, GPCOG, IQCODE, s-IQCODE Mini-Cog, MIS, MoCA, s-MoCA, QDRS, RUDAS, and SLUMS)[1] should be used for the early detection of cognitive impairment (including MCI and dementia)?
Recommendations (open for public comment)
Recommendation 1: In English- and Spanish-speaking adults aged 55 years and older presenting to primary and appropriate ambulatory care settings, the panel suggests **using either AD8 or SLUMS** over no screening for the early detection of cognitive impairment (including mild cognitive impairment and dementia) *(conditional recommendation*, low certainty evidence for AD8 test accuracy, moderate certainty evidence for SLUMS test accuracy; very low certainty evidence for overall test effects)*.
Recommendation 2: In English- and Spanish-speaking adults aged 55 years and older presenting to primary and appropriate ambulatory care settings, the panel suggests **against using MIS, Mini-Cog, and MoCA** for the early detection of cognitive impairment (including mild cognitive impairment and dementia) *(conditional recommendation, low certainty evidence for MIS test accuracy, low certainty evidence for Mini-Cog test accuracy, very low certainty evidence for MoCA test accuracy; very low certainty evidence for overall test effects)*.
Recommendation 3: In English- and Spanish-speaking adults aged 55 years and older presenting to primary and appropriate ambulatory care settings, the panel suggests **neither for or against using IQCODE and s-IQCODE** for the early detection of cognitive impairment (including mild cognitive impairment and dementia) (conditional recommendation, low certainty evidence for IQCODE test accuracy, low certainty evidence for s-IQCODE test accuracy; very low certainty evidence for overall test effects).
Recommendation 4: The panel makes **no recommendation regarding the use of 5-Cog, GPCOG, QDRS, RUDAS, and short MoCA** for English- and Spanish-speaking adults aged 55 years and older presenting to primary and appropriate ambulatory care settings (knowledge gap).

Remarks
● Cognitive testing is not recommended in primary care or ambulatory settings for the following patients, including but not limited to: patients with delirium, agitation, acute or sleep deprivation, or those with unstable medical issues; patients receiving new or recently increased sedative medications or those actively using or withdrawing from drugs of abuse, in the middle of a major depressive episode; patients with an established diagnosis of cognitive impairment or already receiving treatments for cognitive decline; patients with any condition that interferes with having a clear sensorium would contraindicate a screening test. ● These recommendations are not intended to cover acutely ill populations (e.g., urgent care, emergency department) or referred populations seen in specialty clinics (e.g., memory disorder clinics, neurology clinics). The index tests may perform differently in referred populations.

Footnotes:

1. Comparison used for evidence synthesis: Any included cognitive assessment (index tests) vs clinical diagnosis of MCI or dementia using any recognized classification system or consensus diagnosis (reference standards).

Table 2. *Definitions for interpreting the certainty of the evidence and implementing strong vs. conditional recommendations.*

DEFINITION OF CERTAINTY OF THE EVIDENCE	
Category	**Definition**
High	Very confident that the true effect lies close to that of the estimate of the effect.
Moderate	Moderately confident in the effect estimate; the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.
Low	Confidence in the effect estimate is limited; the true effect may be substantially different from the estimate of the effect.
Very Low	Very little confidence in the effect estimate; the true effect is likely to be substantially different from the estimate of the effect.

DEFINITION OF STRONG VS. CONDITIONAL RECOMMENDATIONS AND IMPLICATIONS FOR COLLABORATIVE PARTIES		
Implications	**Strong Recommendations**	**Conditional Recommendations**

For Patients	Most patients in this situation would want the recommended course of action, and only a small proportion would not. Formal decision aids are not likely to be needed to help individuals make decisions consistent with their values and preferences.	Most patients in this situation would want the suggested course of action, but many would not.
For Clinicians	Most patients should receive this course of action. Adherence to this recommendation, according to the guideline, could be used as a quality criterion or performance indicator.	Recognize that different choices will be appropriate for individual patients and that you must help each patient arrive at a management decision consistent with his or her values and preferences. Decision aids may be useful in helping patients make decisions consistent with their values and preferences.
For Policy Makers	The recommendation can be adapted as policy in most situations.	Policy making will require substantial debate and the involvement of various collaborative parties.
Researchers	The recommendation is supported by credible research or other convincing judgments that make additional research unlikely to alter the recommendation. On occasion, a strong recommendation is based on low or very low certainty in the evidence. In such instances, further research may provide important information that alters the recommendations.	The recommendation is likely to be strengthened (for future updates or adaptation) by additional research. An evaluation of the conditions and criteria (and the related judgments, research evidence, and additional considerations) that determined the conditional (rather than strong) recommendation will help to identify possible research gaps.

Conclusions: The panel judged the benefits of using AD8 or SLUMS in patients aged 55 and older presenting to primary care to outweigh the harms, and therefore made conditional recommendations for the use of either test. The panel judged the harms of using MIS, Mini-Cog, and MoCA to outweigh the benefits and made conditional recommendations against their use. The panel judged the benefits and harms of using IQCODE and s-IQCODE to be balanced and therefore does not recommend for or against either test. Lastly, the panel identified no studies meeting eligibility criteria in primary care or community-based settings on the use of 5-Cog, GPCOG, QDRS, RUDAS, and short MoCA and chose not to formulate recommendations for those tests.

Next Steps: This clinical practice guideline (and associated systematic review) will be published later this year and will provide finalized recommendations based on the best available evidence.

------BELOW IS CONTEXTUAL INFORMATION FOR REFERENCE ONLY--------

Additional information on systematic review and guideline methodology:

Index tests of interest for which evidence was sought (including short forms)**:**

- 5-Cog
- AD8
- GPCOG
- IQCODE
- Mini-Cog
- MIS
- MoCA
- QDRS
- RUDAS
- SLUMS

Excluded index tests:

- Tests that only assess function and behavior
- Tests that rely solely on subjective assessments
- Tests only available in digital format
- Tests requiring a license for use
- Combinations, composites, or components of index tests
- Specific excluded tools include:
    - Mini-Mental State Examination (MMSE)
    - Clock Drawing Test
    - Cognitive Function Index

Excluded settings:

- Specialty care settings/referred populations

Reference Standards:

- Clinical diagnosis using recognized classification systems or consensus diagnosis. For example:
    - Petersen criteria (MCI)
    - DSM criteria (dementia)
    - NINCDS-ADRDA criteria (Alzheimer's)
    - Clinical Dementia Rating (CDR)

Outcomes:

- Sensitivity and specificity
- Positive/negative predictive value
- Area under the curve (AUC)
- Likelihood ratios
- Frequencies of TP, FP, TN, FN

A priori decisions made by the panel for decision making:

The panel set thresholds for test accuracy based on the number of false negative and false positive results they would be willing to accept:

- False negative threshold: 5 per 100
- False positive threshold: 10 per 100

The panel agreed on a prevalence of 34.2% for cognitive impairment, which was used to estimate the number of true positives, true negatives, false positives, and false negatives each test would result in based on data analyzed in the systematic review.

For decision-making, the panel placed a greater emphasis on maximizing true positives (as opposed to true negatives) and minimizing false negatives (as opposed to false positives), as the intent is early identification and differentiation of any cognitive impairment or related conditions.

The panel decided to only use community and primary care data to inform decision-making for which tests to recommend for primary care. Although the systematic review identified studies evaluating the cognitive assessments in specialty care or referred populations, the panel agreed that these settings are too indirect for the clinical question, and the pre-test probability will differ from that of primary care.

Results of main analysis:
15 observational studies were identified that assessed the accuracy of the cognitive tests in the population and setting of interest (i.e., patients 55 years and older in primary care or community-based settings). Where possible, data were pooled across studies to estimate a single sensitivity and specificity. When pooling was not possible, ranges were reported **(Table 3)**.

Table 3. *Summary of findings and accuracy judgments for the cognitive assessments.*

Test	N studies (n participants) Pooled or range	Sensitivity	Specificity	TP/TN/FP/FN frequencies[1] (95% CI)	Accuracy judgement	Certainty of the evidence
AD8	3[2,3,4] (881) Pooled	0.77 (95% CI: 0.67-0.85)	0.87 (95% CI: 0.82-0.91)	TP: 26 (23 to 29) TN: 57 (54 to 60) FN: 8 (5 to 11) FP: 9 (6 to 12)	Inaccurate	Low
IQCODE	1[5] (160)	0.71 (95% CI: 0.61-0.80)	0.74 (95%CI: 0.62-0.84)	TP: 24 (21 to 27) TN: 49 (41 to 55) FN : 10 (7 to 13) FP: 17 (11 to 25)	Very inaccurate	Low
Mini-Cog	2[6,7] (721) Range	0.39 to 0.50	0.73 to 0.78	TP: 13 to 17 TN: 48 to 51 FN: 17 to 21 FP: 15 to 18	Very inaccurate	Low
MIS	2[6,8] (747) Range	0.17 to 0.73	0.87 to 0.98	TP: 6 to 25 TN: 57 to 64 FN): 9 to 28 FP: 2 to 9	Very inaccurate	Low
MoCA	5[9,10,11,12,13] (1,040) Range	0.57 to 0.98	0.10 to 0.72	TP: 19 to 34 TN: 7 to 47 FN: 0 to 15 FP: 19 to 59	Very inaccurate	Very low
s-IQCODE	2[14,15] (1,004)	0.55 to 0.80	0.82 to 0.93	TP: 19 to 27 TN: 54 to 61 FN: 7 to 15	Inaccurate	Low

	Range			FP: 5 to 12		
SLUMS	1[17] (433)	0.98 (95% CI: 0.93-1.00)	0.61 (95% CI: 0.55-0.67)	TP:34 (32 to 34) TN: 40 (36 to 44) FN:0 (0 to 2) FP: 26 (22 to 30)	Accurate	Moderate
5-Cog	No included studies	-	-	-	-	-
GPCOG	No included studies	-	-	-	-	-
QDRS	No included studies	-	-	-	-	-
RUDAS	No included studies	-	-	-	-	-
s-MoCA	No included studies	-	-	-	-	-

1. True positive, true negative, false negative, false positive; based on a prevalence of 34.2%
2. Galvin 2005
3. Malmstrom 2009
4. Tainta 2022
5. Cruz-Orduna 2012
6. Holsinger 2012
7. Kaufer 2008
8. Carnero-Pardo 2011
9. Alfano 2022
10. Katz 2021
11. McLennan 2011
12. Rossetti 2019
13. Stimmel 2024
14. Ayalon 2011
15. Grober 2017
16. Tariq 2006

Additional contextual factors considered as part of GRADE evidence-to-decision framework:

Additional contextual factors, using the GRADE approach, regarding the use of each cognitive assessment (vs. no testing) were considered. *We acknowledge this section is methodologically jargon-heavy, and will fully explain our methodology, the evidence, and our judgments on the evidence in our final manuscripts.*

Desirable and undesirable effects (i.e., benefits and harms) of each test with available data were judged by the panel based on downstream consequences of true positives, true negatives, false negatives, and false positives. Desirable effects, undesirable effects, and accuracy were all considered when judging the balance of effects **(Table 4).**

Table 4. *Balance of effects for cognitive assessments with available data.*

Test	Desirable Effects	Undesirable Effects	Accuracy*	Balance of effects
AD8	Moderate	Small	Inaccurate	Probably favors using AD8
IQCODE	Moderate	Small	Very inaccurate	Does not favor either/balanced effects
Mini-Cog	Trivial	Large	Very inaccurate	Probably favors not using Mini-Cog
MIS	Trivial	Moderate	Very inaccurate	Probably favors not using MIS
MoCA	Small	Moderate	Very inaccurate	Probably favors not using MoCA
s-IQCODE	Small	Small	Inaccurate	Does not favor either/balanced effects
SLUMS	Large	Small	Accurate	Probably favors using SLUMS

*Based on true positives, true negatives, false positives, false negatives, sensitivity, and specificity

Other evidence-to-decision factors were discussed as they relate to cognitive testing more generally (i.e., unlikely to differ depending on which assessment is given). This information will be provided in the manuscript, and final judgments for those factors are reported below:

- Certainty of the evidence of management's effects: How certain is the panel that treatment and management for MCI and dementia is effective? *Low to very low certainty.*
- Certainty of the evidence of the link between test result and management: How certain is the panel that patients who are positive based on a cognitive assessment will proceed to further testing/diagnosis and management? *Low to very low certainty.*
- Overall certainty of test effects: Based on the certainty of the evidence of management's effects and the link between test result and management: *Low to very low certainty.*
- Patients' values and preferences: Is there important uncertainty about or variability in how much people value the main outcomes? *Possibly important uncertainty or variability.*
- Resources required: How large are the resource requirements/costs? *Negligible costs.*
- Cost-effectiveness: Does the cost effectiveness of using cognitive assessments favor their use or non-use? *No included studies.*
- Equity: What would be the impact on health equity? *Increased*
- Acceptability: Is the use of cognitive assessments acceptable to collaborative parties? *Yes.*
- Feasibility: Are cognitive assessments feasible to implement? *Varies.*

Limitations of the evidence synthesis and evidence-to-decision process:
The evidence base was restricted to studies published in English, which may have excluded relevant data published in other languages, such as Spanish. Additionally, only index tests administered in English and Spanish were considered, potentially limiting the generalizability of findings to more diverse populations. The inclusion criteria also restricted the population to individuals aged 55 years and older, which may not capture data relevant to younger individuals. However, dementia occurring before age 55 is typically attributed to early-onset forms, which often have different etiologies, including those not related to AD.

The Mini-Mental State Examination (MMSE) was not included in this review, which limits its completeness. The MMSE was excluded due to licensing requirements that restrict its accessibility in many primary care and community-based settings. Furthermore, this review and guideline did not consider or evaluate the performance of repeat administrations of the index tests over time, did not consider adjustments for potential confounding factors such as age or education, and did not evaluate different cutoffs for index tests as this is outside the scope of the clinical question. Finally, because of significant heterogeneity among the included studies, it was not possible to pool data across studies for all cognitive tests, underscoring the need for further research in these areas.

Contact information and authors list:

Contact: *Please use the online form to provide feedback on this guideline. For any general questions about the Alzheimer's Association's Guideline Development Program, please contact Malavika Tampi, Director, Clinical Practice Guidelines Program and Methodology Lead (mptampi@alz.org).*

This particular document was prepared by the following guideline panel and methodology team members. Additional authors contributed to the systematic review and guideline manuscripts and will be appropriately included in publications along with conflict of interest disclosure forms for all.

Guideline panel (chairs, panel in alphabetical order):

Dorene Rentz (Co-chair), Fred Ko (Co-chair), Karen L. Bell, Sharon A. Brangman, Noelle Fields, Brent Forester, Jennifer Gatchel, Dana Jones, Ambar Kulshreshtha, Martin J. Sliwinski, Peter Snyder, Eric Tangalos

Methodology team:

- Malavika Tampi, Director, Clinical Practice Guidelines Methodology and Program Lead, Alzheimer's Association (mptampi@alz.org)
- Sarah Pahlke, Director, Clinical Practice Guidelines Methodology, Alzheimer's Association
- Jennifer Loveless, Contract Methodologist
- Lauren Pilcher, Contract Methodologist
- Heather Snyder, SVP, Med & Scientific Operations, Alzheimer's Association
- Simin Mahinrad, Director, Medical Writing, Alzheimer's Association
- Mary Beth McAteer, Contract Medical Librarian