Alzheimer's Disease Dementia Guidelines for Diagnostic Testing: A Systematic Review

Ingrid Arevalo-Rodriguez, MSc[1,6], Olga L. Pedraza, MD, MSc[2], Andrea Rodríguez, MD, MSc[1], Erick Sánchez, MD[2], Ignasi Gich, MSc[3], Ivan Solà, MSc[3], Xavier Bonfill, MD, PhD[4,5], and Pablo Alonso-Coello, MD, PhD[4,5]

Abstract

Alzheimer's disease dementia (AD dementia) is one of the most common neurodegenerative diseases worldwide, with a growing incidence during the last decades. Clinical diagnosis of cognitive impairment and presence of AD biomarkers have become important issues for early and adequate treatment. We performed a systematic literature search and quality appraisal of AD dementia guidelines, published between 2005 and 2011, which contained diagnostic recommendations on AD dementia. We also analyzed diagnostic recommendations related to the use of brief cognitive tests, neuropsychological evaluation, and AD biomarkers. Of the 537 retrieved references, 15 met the selection criteria. We found that Appraisal of Guidelines Research and Evaluation (AGREE)-II domains such as applicability and editorial independence had the lowest scores. The wide variability on assessment of quality of evidence and strength of recommendations were the main concerns identified regarding diagnostic testing. Although the appropriate methodology for clinical practice guideline development is well known, the quality of diagnostic AD dementia guidelines can be significantly improved.

Keywords

diagnosis, Alzheimer's disease dementia, systematic review, clinical practice guidelines

Background

The world population is rapidly aging, which causes a significant increase in chronic and neurodegenerative diseases, with serious consequences in terms of global public health. The prevalence of diseases such as dementia is nearly 42 million patients in 2012, with approximately 4.6 million new cases a year.[1] These figures would increase up to 300% by 2040, with an impact on expenditure of over US$ 422 billion.[1,2]

Alzheimer's disease dementia (AD dementia), the most common dementia in the Western world, is defined as a cognitive decline documented by standardized testing on 2 or more domains, which interferes with daily function and represents a decline in previous levels.[3] Recently, the National Institute of Neurological and Communicative Disorders and Stroke–Alzheimer's Disease and Related Disorders Association criteria have been updated and presented 3 different types of AD dementia according to presence of core clinical criteria only (probable), an atypical course without differential diagnosis (possible), or neuropathological evidence (definitive).[4]

The presence of cognitive impairment, especially in the memory domain, is a fundamental part of the clinical criteria used to diagnose AD dementia. This impairment must be detected and diagnosed through a combination of clinical history analysis and objective cognitive assessment by means of a brief mental evaluation or comprehensive neuropsychological testing.[4] The latter reflects the clinical dimension of the disease and characterizes the type of presentation. However, new trends in AD dementia

[1] Grupo de Evaluación de Tecnologías y Políticas en Salud - GETS, Clinical Research Institute - School of Medicine. National University of Colombia, Bogotá DC, Colombia
[2] Department of Neurosciences and Memory Clinic Unit, Fundacion Universitaria de Ciencias de la Salud, Bogota, Colombia
[3] Iberoamerican Cochrane Centre, Clinical Epidemiology and Public Health Department. Institute of Biomedical Research (IIB Sant Pau), Spain
[4] Iberoamerican Cochrane Centre, Institute of Biomedical Research (IIB Sant Pau), CIBER Epidemiología y Salud Pública (CIBERESP), Spain
[5] Universitat Autònoma de Barcelona, Barcelona, Spain
[6] Doctoral candidate at the Pediatrics, Obstetrics & Gynecology and Preventive Medicine Department, Universitat Autònoma de Barcelona, Barcelona, Spain

Corresponding Author:
Ingrid Arevalo-Rodriguez, MSc, Grupo de Evaluación de Tecnologías y Políticas en Salud - GETS, Clinical Research Institute, School of Medicine, National University of Colombia, Bogotá DC, Colombia.
Email: ingrid.arevalo@e-campus.uab.cat

diagnosis suggest the need to reflect upon the biological dimension of the disease, including biomarkers related to brain amyloid-β protein deposition and downstream neuronal degeneration.[5]

Several clinical practice guidelines (CPGs) for the management of AD dementia have been published in different countries. These offer users (professionals and health administrators, patients and caregivers, among others) clinical recommendations for daily practice based on the best available evidence.[6] Although CPGs have become a useful clinical tool, their quality can vary and could be improved.[7] The CPG assessment must be conducted in an objective manner using standardized instruments that address issues such as rigor of development, clarity of recommendations, and applicability, among others.[8] It is also important to analyze the differences in recommendations among guidelines focused on the same topic, which can vary and even be contradictory.[9,10]

There are previous studies related to the assessment of the quality and content of AD dementia CPGs, but these were not conducted systematically and did not assess recommendations on AD dementia diagnosis.[11-14] The purpose of this review was to assess the quality of CPGs that focus on the diagnosis of AD dementia as well as to examine the recommendations on 3 critical aspects of AD diagnosis: initial testing with brief cognitive screening tools, use of neuropsychological assessment, and use of biomarkers of amyloid-β accumulation and neuronal injury.

Methods

We performed a systematic review of AD dementia CPGs retrieved from the Web sites of guideline developers (Scottish Intercollegiate Guidelines, National Institute for Clinical Excellence, among others), CPG compilers (National Guidelines Clearinghouse, National Library of Guidelines, GIN Web site, among others), and institutions that conduct research related to dementia.

We complemented this search in MEDLINE, using the terms "dementia," "Alzheimer," and "elderly," with a methodological filter to search for CPGs (the full search strategy is available upon request). The CPGs with diagnostic recommendations in English, Spanish, and Portuguese published between January 2005 and December 2011 were included. Previous versions of the same guideline, letters to the editor, or guidelines with a focus on population screening or general mental health were excluded.

Two authors independently selected all potentially eligible studies by reading the titles and abstracts of the retrieved references. Subsequently, 4 trained authors independently assessed the quality of each CPG using the Appraisal of Guidelines Research and Evaluation (AGREE)-II tool.[8] This instrument includes the evaluation of 6 domains: Scope and Purpose, Stakeholder Involvement, Rigor of Development, Clarity of Presentation, Applicability, and Editorial Independence. The intraclass correlation coefficient (ICC) and its 95% confidence interval (95% CI) were calculated to assess interrater agreement. Statistical analyses were performed using SPSS® 15.0 (SPSS Inc, Chicago, IL).

For the analyses of recommendations regarding diagnostic testing, we only included guidelines of high quality, defined as those with scores equal to or greater than 60% in the "Rigor of development" domain. We used this cutoff point as an adequate figure to reflect that a valid and transparent process had been adopted in the development of recommendations.[15] For each guideline, 3 authors extracted recommendations concerning: (1) the brief cognitive instruments recommended for initial evaluation of patients with suspected AD dementia, (2) the use of neuropsychological assessment, and (3) the use of biomarkers in patients with suspected AD dementia. The recommendations (sentences) of each CPG were extracted along with the quality of evidence, strength of the recommendation, and type of studies included. This information was synthesized in a recommendation matrix for further analysis. All disagreements were resolved by consensus.

All paper and pencil tests used for the initial assessment of patients with suspected AD dementia were classified as brief cognitive instruments. Neuropsychological assessment was defined as all comprehensive and standardized cognitive evaluations, including appraisal of memory, language, executive functions, and visual–spatial abilities conducted by trained clinicians. Recommendations on the following AD biomarkers were included: low cerebrospinal fluid (CSF) amyloid β_{42} levels, positive retention of tracer in positron emission tomography (PET), elevated Tau protein levels in CSF (total or phosphorylated), decreased fluorodeoxyglucose (FDG) uptake on PET reflected in the temporoparietal cortex, and presence of patterns of atrophy on magnetic resonance imaging (MRI), involving medial, basal and lateral temporal lobes and medial parietal cortex.[4]

Results

We retrieved a total of 537 citations and excluded 518 that were not related to AD dementia. We also excluded letters to the editor as well as CPGs in languages other than English, Spanish, and Portuguese, those with recommendations concerning only a subset of patients with dementia and those with general recommendations for mental health. In all, 4 CPGs containing therapeutic recommendations only were also excluded (Figure 1). We assessed 15 CPGs distributed among 22 individual references.[16-37]

We found that the domains with the highest scores were "Scope and Purpose" and "Clarity of Presentation" (75.4% and 80%, respectively), whereas "Applicability" and "Editorial Independence" had the lowest figures (32.4% and 47%, respectively). Only 1 (6.7%) CPG had scores above 60% in all domains.[18] All the 6 (40%) guidelines had ratings equal to or above 60% for the "Rigor of Development" domain[18,19,25,30,34,36] (Table 1). We classified these as guidelines of high quality. Agreement among reviewers was considerable (ICC = 0.78, 95% CI: 0.60-0.91).

Of the 15 identified guidelines, 4 did not report a system to grade evidence or recommendations. Conversely, the 6 high-quality CPGs used 4 different grading systems. The most commonly used system, out of these 4, was that developed by the Scottish Intercollegiate Guidelines Network (66%).[38] Only
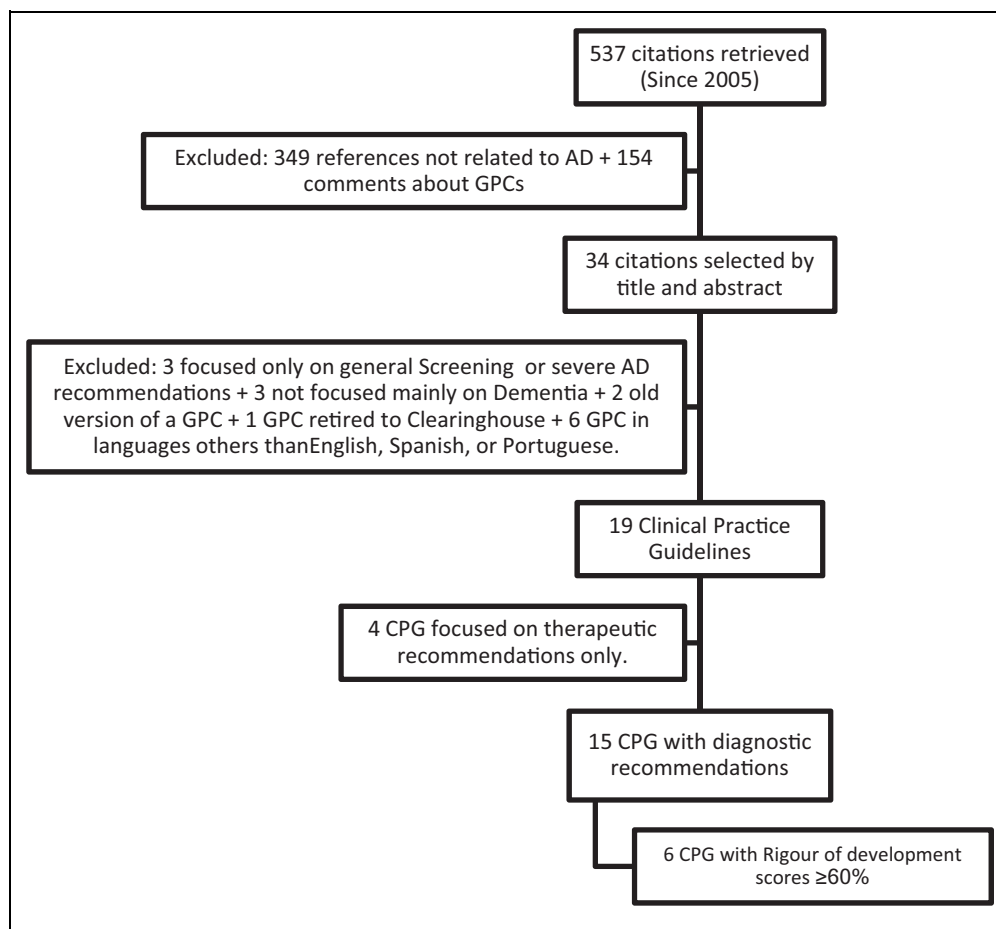
Figure 1. Flowchart of the systematic search.

1 CPG used a tool specifically designed for diagnostic questions[30] (Table 2). One CPG presented a diagnostic classification system in the Methods section, but after concluding that there was no enough information to address the diagnostic topics, they based their recommendations on an expert consensus and presented them without graduation of strength.[18]

All the 15 included guidelines had information on at least 1 of the 3 critical aspects of AD dementia diagnosis included in this review. In all, 14 (93%) provided recommendations related to the use of brief cognitive tests in the diagnosis of patients with suspected AD, 9 (60%) presented recommendations on the use of neuropsychological tests, and 4 (26%) presented explicit recommendations on use of β-amyloid accumulation/neuronal injury biomarkers.

Brief Cognitive Test Recommendations

All 6 (100%) high-quality guidelines provided recommendations on the use of brief cognitive tests in the diagnosis of patients with suspected AD dementia. They emphasized the importance of conducting a formal cognitive assessment using objective and valid instruments, although 3 (50%) of them did not recommend any test in particular (Table 3). A total of 19 brief instruments were recommended in the guidelines

retrieved, the Mini-Mental State Examination (MMSE) being the most frequently recommended. Of the 6 high-quality guidelines, 5 (83%) considered information on national validation and standardization of recommended tests, and 2 (33%) of them, recommended the use of MMSE versions validated in the country where the CPG had been developed.[34,36]

Observational designs, especially cohorts and cross-sectional studies, were the most frequently included study designs. We found disagreements between the quality assessment of evidence and the confidence of recommendation among the high-quality guidelines. The former varied from very low to high depending on the system used. The strength of recommendation ranged from A to C (Table 3).

Neuropsychological Assessment Recommendations

All the 6 (100%) high-quality guidelines presented recommendations related to the use of neuropsychological tests. They recommended neuropsychological testing only when the diagnosis is unclear, when there is uncertainty about the stage of the disease, or when a differential diagnosis needs to be ruled out. Again, observational studies were the most frequently included type of design included. The quality of the evidence varied from very low to high and the strength of recommendations

Table 1. The AGREE II Domain-Standardized Scores for CPG on Alzheimer's Dementia Diagnostic Recommendations.

Guideline	Organization	Country and Language	Publication Year	Scope and Purpose	Stakeholder Involvement	Rigor of Development	Clarity of Presentation	Applicability	Editorial Independence
Diagnosis of Alzheimer's disease in Brazil[16,17]	Scientific Department of Cognitive Neurology and Aging of the Brazilian Academy of Neurology	Brazil, Portuguese	2005	64	26	57	88	18	15
Dementia: supporting people with dementia and their carers in health and social care[18]	National Institute for Health and Clinical Excellence (NICE) & Social Care Institute for Excellence	United Kingdom, English	2006	92	88	83	86	65	77
Management of patients with dementia: A national clinical guideline[19]	Scottish Intercollegiate Guidelines Network	United Kingdom, English	2006	69	64	92	93	93	58
Recommendations for best practices in the treatment of Alzheimer's disease in managed care[31]	The Alzheimer's Drug Discovery Foundation	United States, English	2006	74	64	49	86	29	40
Dementia[21]	Ministry of Health, Singapore	Singapore, English	2007	86	69	49	90	52	4
Cognitive impairment in the elderly recognition, diagnosis, and management[20]	British Columbia Medical Association	Canada, English	2007	68	17	16	72	29	10
Diagnosis and treatment of dementia[23-29]	Third Canadian Consensus Conference on the Diagnosis and Treatment of Dementia	Canada, English	2008	68	57	74	93	24	92
Guideline for Alzheimer's disease management[22]	California Workgroup on Guidelines for Alzheimer's Disease Management	United States, English	2008	76	53	38	75	22	19
Management of dementia[36]	Ministry of Health, Malaysia	Malaysia, English	2009	99	71	81	93	34	94
Dementia: diagnosis and treatment guideline[33]	Group Health Cooperative	United States, English	2009	26	22	18	49	11	8
Clinical practice guideline. Diagnosis and treatment of dementia in elderly in primary care[35]	Instituto Mexicano del Seguro Social	México, Spanish	2009	92	53	51	54	21	48
Clinical practice guideline about comprehensive care for people with Alzheimer's disease and other dementias[34]	Agència d'Informació, Avaluació i Qualitat en Salut de Cataluña	Spain, Spanish	2010	100	88	80	92	38	92
EFNS guidelines for the diagnosis and management of Alzheimer's disease.[30]	European Federation of Neurological Societies	International (Europe), English	2010	65	47	70	88	28	83
Dementia: Diagnosis and Treatment[37]	Regional Health Council—Tuscany	Italy, English	2011	75	63	27	79	3	15
Alzheimer's disease. Clinical practice guideline[32]	Argentine Neurological Society Research Group on Behavioral Neurology and Cognitive Neurosciences	Argentina, Spanish	2011	76	49	44	63	20	58
Mean scores				75.4	55.3	55.3	80.0	32.4	47.5
Range				26-100	17-88	16-92	49-93	3-93	4-94

Abbreviations: CPG, clinical practice guideline; EFNS, European Federation of Neurological Societies.

Table 2. Quality of Evidence and Strength of Recommendation Systems Used.

	Quality of the Evidence	Strength of Recommendation	CPG
SIGN[38]	1++ = High-quality meta-analyses, systematic reviews of randomized controlled trials (RCTs), or RCTs with a very low risk of bias. 1+ = Well-conducted meta-analyses, systematic reviews of RCTs, or RCTs with a low risk of bias. 2++ = High-quality systematic reviews of case control or cohort studies; high-quality case–control or cohort studies with a very low risk of confounding or bias and a high probability that the relationship is causal. 3 = Nonanalytic studies, eg, case reports, case series. 4 = Expert opinion.	Level A = At least 1 meta-analysis, systematic review of RCTs, or RCT rated as 1++ and directly applicable to the target population; or a body of evidence consisting principally of studies rated as 1+, directly applicable to the target population, and demonstrating overall consistency of results. Level B = A body of evidence including studies rated as 2++, directly applicable to the target population, and demonstrating overall consistency of results; or Extrapolated evidence from studies rated as 1++ or 1+ Level C = A body of evidence including studies rated as 2+, directly applicable to the target population and demonstrating overall consistency of results; or Extrapolated evidence from studies rated as 2++ Level D = Evidence level 3 or 4; or extrapolated evidence from studies rated as 2+.	19,34
Canadian Task Force on Preventive Health Care	1 = Evidence obtained from at least 1 proper RCT. 2 = Evidence obtained from well-designed controlled trials without randomization, or evidence obtained from well-designed cohort or case–control analytic studies preferably from more than 1 center or research group, or evidence obtained from comparisons between times or places with or without the intervention. Dramatic results in uncontrolled experiments are included in this category. 3 = Opinions of respected authorities based on clinical experience, descriptive studies, or reports of expert committees.	Level A. There is good evidence to support this maneuver. Level B. There is fair evidence to support this maneuver. Level C. There is insufficient evidence to recommend for or against this maneuver, but recommendations may be made on other grounds. Level D. There is a fair evidence to recommend against this procedure. Level E. There is good evidence to recommend against this procedure.	25
EFNS Classification	I = A prospective study in a broad spectrum of persons with the suspected condition, using a "Gold Standard" for case definition, where the test is applied in a blinded evaluation, and enabling the assessment of appropriate tests of diagnostic accuracy. II = A prospective study of a narrow spectrum of persons with the suspected condition, or a well-designed retrospective study of a broad spectrum of persons with an established condition (by Gold Standard) compared to a broad spectrum of controls, where test is applied in a blinded evaluation, and enabling the assessment of appropriate tests of diagnostic accuracy. III = Evidence provided by a retrospective study where either persons with the established condition or controls are of a narrow spectrum, and where test is applied in a blinded evaluation. Class IV = Any design where test is not applied in blinded evaluation or evidence provided by expert opinion alone or in descriptive case series (without controls).	Level A rating (established as useful/predictive or not useful/predictive). Level B rating (established as probably useful/predictive or not useful/predictive). Level C rating (established as possibly useful/predictive or not useful/predictive) requires at least 2 convincing class III studies.	30
Canadian Task Force on Preventive Health Care and SIGN Modified	I = Evidence obtained from at least 1 properly designed RCT. II-1 = Evidence obtained from well-designed controlled trials without randomization. II-2 = Evidence obtained from well-designed cohort or case–control analytic studies, preferably from more than 1 center or research group. II-3 = Evidence obtained from multiple time series with or without the intervention. Dramatic results in uncontrolled experiments could also be regarded as this type of evidence. III = Opinions of respected authorities, based on clinical experience; descriptive studies and case reports; or reports of expert committees.	Level A = At least 1 meta-analysis, systematic review, or RCT, or evidence rated as good and directly applicable to the target population Level B = Evidence from well-conducted clinical trials, directly applicable to the target population, and demonstrating overall consistency of results; or evidence extrapolated from meta-analysis, systematic review, or RCT Level C = Evidence from expert committee reports, or opinions and/or clinical experiences of respected authorities; indicates absence of directly applicable clinical studies of good quality.	36

Abbreviations: EFNS, European Federation of Neurological Societies; SIGN, Scottish Intercollegiate Guidelines Network; RCT, randomized controlled trials.

Table 3. Recommendations for Brief Cognitive Tests in AD Dementia.

	N	Level of Evidence	Grade of Recommendation
MMSE and/or other tests	3	2[++19,34] and III[36]	B,[19] C,[36] and A[34]
Generic recommendation about brief cognitive test	3	2,[25] I,[30] and expert consensus[18]	B,[25] A,[30] and no strength[18]

Abbreviations: AD, Alzheimer's disease; MMSE, Mini-Mental State Examination.

Table 4. Neuropsychological Assessment Recommendations in AD Dementia Guidelines.

.	Level of Evidence	Grade of Recommendation
"Fomal neuropsychological testing should form part of the assessment in cases of mild or questionable dementia"[18]	Expert consensus	No strength
"Neuropsychological testing should be used in the diagnosis of dementia, especially in patients where dementia is not clinically obvious"[19]	I[++], 2[++]	B
"The diagnosis and differential diagnosis of dementia is currently a clinically integrative one. Neuropsychological testing alone cannot be used for this purpose and should be used selectively in clinical settings. Neuropsychological testing may aid in: addressing the distinction between normal aging, mild cognitive impairment or cognitive impairment without dementia, and early dementia; addressing the risk of progression from mild cognitive impairment or cognitive impairment without dementia to dementia or Alzheimer dementia; determining the differential diagnosis of dementia and other syndromes of cognitive impairment."[25]	2	B
"When the diagnosis of dementia is inconclusive, then neuropsychological tests will be required"[36]	III	C
"We recommend performed a detailed neuropsychological evaluation by specific tests when there are discrepancies between the clinical impression and screening tests, diagnostic concerns or when the complaints are of short duration or limited to a single cognitive domain"[34]	4	D
"Quantitative neuropsychological testing should be made in patients with questionable or very early Alzheimer Dementia"[30] "The assessment of cognitive functions should include a general cognitive measure and more detailed testing of the main cognitive domains, and in particular an assessment of delay recall"[30]	III, I	B, A

Abbreviation: AD, Alzheimer's disease.

from A to D, similar to brief cognitive recommendations. None of the CPGs proposed a specific neuropsychological battery of instruments and only 1 guideline included a recommendation about the assessment of major cognitive domains (memory, executive functions, language, attention, and visuospatial skills; Table 4).[30]

The AD Biomarkers Recommendations

Only 4 (66%) of the 6 high-quality CPGs presented explicit recommendations for the use of β-amyloid accumulation/neuronal injury biomarkers (Table 5). In all, 3 of these emphasized a lack of clinical studies on these biomarkers and presented recommendations against their use.[19,34,36] The fourth high-quality CPG found FDG–PET and Tau levels to be useful,[30] although the direction of recommendation (for or against) remains unclear. Again, observational designs were the most frequently included studies. The quality of evidence varied widely, from low to high, depending on the system used. Similarly, the strength of recommendation ranged from B to C.

Discussion

This review shows that CPGs addressing diagnostic issues related to AD dementia published between 2005 and 2011 are very heterogeneous in terms of quality according to their AGREE-II scores. Domains such as applicability and editorial independence had very low scores and deserve special attention. Previous reviews of AD dementia CPGs highlight similar difficulties on these topics, even though such shortcomings could be easily improved.[7,12] For example, if the applicability domain addresses organizational barriers, cost implications, and monitoring criteria related to guideline implementation,[39-42] it is important that guideline developers identify a priori factors related to transferability, such as values, costs, availability of resources, and influence of economic and intellectual interest on implementation. Regarding editorial independence issues,[41,43,44] a complete and valid documentation of conflicts of interest for all participants in all stages of guideline development is required.[45]

Issues concerning CPG rigor of development deserve closer attention. In our review, we noticed that 26% of the CPGs did

Table 5. Recommendations for the Use of Alzheimer's Dementia Biomarkers.

Level of evidence and strength of recommendation[a]	Biomarkers of Aβ Accumulation		Biomarkers of Neuronal Injury	
	Abnormal tracer retention on amyloid PET imaging	Low CSF Aβ42	Elevated CSF Tau (total or phosphorylated)	Decreases FDG uptake on PET involving temporoparietal cortex	Atrophy on MRI, involving medial, basal and lateral temporal lobes and medial and lateral cortices.
Management of patients with dementia: A national clinical guideline[19]		2^{++}, B[b]	2^{++}, B[b]		
Management of dementia[36]	III, C[c]	III, C[b]	III, C[b]	III, C[b]	III, C[b]
Clinical practice guidelines about comprehensive care for people with Alzheimer's disease and other dementias[34]		2^{++}, B[b]	2^{++}, B[b]		
EFNS guidelines for the diagnosis and management of Alzheimer's disease[30]	I, B[c]	I, B[c]	I, B[c]	I, B[c]	I, B[c]

Abbreviations: CSF, cerebrospinal fluid; FDG, fluorodeoxyglucose; EFNS, European Federation of Neurological Societies; MRI, magnetic resonance imaging; PET, positron emission tomography.
[a] Inside cells: level of evidence, grade of recommendation.
[b] Recommendation against use of biomarker.
[c] Recommendation unclear for use of biomarker.

not use a system to grade the quality of the evidence and the strength of recommendations. However, among those that did use a system, only 1 included diagnostic-specific elements in their framework (eg, appropriate reference standard or adequate spectrum of patients).[46] This figure is disturbing because CPG users need to know how much they can trust the evaluated information and whether adherence to the recommendations will yield more benefits than harm.[47] In our opinion, the use of generic systems developed by therapeutic studies would lead to inaccurate ratings of the quality of the evidence and would generate uncertainty about the recommendations. An example is the use of neuropsychological testing. Although recommendations were fairly homogeneous and specific in terms of content, their strength and quality of the evidence on which they were based varied from very low (experts opinion, D) to very high (1^{++}, A). This situation is worrying because although we selected only high-quality guidelines for these comparisons, it is not clear whether the evidence related to neuropsychological tests is of good quality or whether developers are confident about implementing the recommendations provided.

We found a similar situation regarding the use of brief cognitive tests, which added to other concerns about the clinical utility of these tools. In our review, we found a total of 19 tests recommended for initial assessment, but 3 of the high-quality guidelines only prescribed a cognitive assessment, although without identifying any instrument to achieve that goal. Once again, we encountered difficulties in assessing the quality of the diagnostic evidence and the strength of recommendations, but this variability also reflects doubts about the advantages and limitations of the tests.[25] We think that valid and global figures about accuracy of these tools, especially those related to subgroups such as patients with mild cognitive impairment or the general population, are imperative. This information could help us determine which test(s) would be more useful for the initial assessment of patients with suspected AD dementia.

We found that only a few (26%) CPGs addressed the use of biomarkers. Three of them did not recommend their use in daily clinical practice and only one suggested that they could be potentially useful.[30] Based on the currently available evidence, the uncertainty with respect to the applicability and benefit of these biomarkers in clinical settings does not support their inclusion in CPGs. The CPG users, who still wish to adopt these diagnostic tools in daily practice, could face difficulties related to local standardization, development of cutoff points for different stages of AD, and availability of trained staff (eg, radiologists in measurement of MRI structures). Once again, we noticed that formal evaluation of clinical accuracy of these biomarkers is necessary.[48]

A previous review by Beck et al also found considerable differences among diagnostic recommendations in CPGs on dementia published until 1999.[11] They attributed these differences to the methods used to formulate recommendations and to the need for an evidence-based approach for future guideline development. We found that disagreement persisted even in rigorously developed evidence-based CPGs, causing significant consequences in terms of interpretation, comparison, and generalizability of the recommendations.[12] In our opinion, adopting a system that took into account the nature of the evidence evaluated would be helpful. The framework outlined by Grading of Recommendations Assessment, Development and Evaluation approach has been developed to provide a common, rigorous, and comprehensive framework with a specific consideration of diagnostic issues that could be helpful in solving these difficulties.[49]

Our review has several strengths and limitations. We conducted an exhaustive and systematic search strategy of the available literature that was not limited to indexed databases, given that it included references that had not been used in other reviews on AD dementia. However, we had to exclude 6 CPGs in languages such as German and Chinese as well as 4 references with only therapeutic recommendations, which could have facilitated a more complete evaluation of AD dementia guidelines. Similarly, we only included guidelines published between 2005 and 2011. Although we selected this time span in order to be more certain about the appropriateness of the guidelines, CPG developers may have had different interpretations of the same evidence.

We used a valid and widely disseminated instrument (AGREE-II) to assess guideline quality as well as a reliable process based on independent assessment by 4 appraisers. Nevertheless, the use of the AGREE instrument has potential drawbacks related to a distinctive focus on the development process of the CPG instead of on the quality of individual recommendations within the guideline.[50] In addition, we only used 1 domain of the AGREE-II tool to select the guidelines for individual analyses, whereas other reviews recommend the assessment of 2 domains, namely rigor of development and editorial independence. In our review, we aimed to assess the methodological issues related to the development of CPGs in order to highlight the issues associated with using generic grading systems, but we also noticed that all selected guidelines obtained adequate scores on the editorial independence domain ($\geq 50\%$).

As a conclusion, it is necessary that clinicians have access to high-quality and up-to-date recommendations from the best available evidence for the adequate management of patients with AD dementia. The quality of CPGs on the diagnosis of AD dementia can be improved not only through the inclusion of ethical aspects, the participation of stakeholders, and a clear strategy for their implementation, but also through a thorough evaluation of the quality of the available evidence using a unified framework to present recommendations.

Declaration of Conflicting Interests

Funding

References

1. Ferri CP, Prince M, Brayne C, et al. Global prevalence of dementia: a Delphi consensus study. *Lancet*. 2005;366(9503):2112-2117.
2. Wimo A, Winblad B, Jonsson L. The worldwide societal costs of dementia: Estimates for 2009. *Alzheimer's Dement*. 2010;6:98-103.
3. McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM. Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology*. 1984;34(7):939-944.
4. McKhann GM, Knopman DS, Chertkow H, et al. The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's Dement*. 2011;7:263-269.
5. Dubois B, Feldman HH, Jacova C, et al. Revising the definition of Alzheimer's disease: a new lexicon. *Lancet Neurol*. 2010;9: 1118-1127.
6. Woolf SH, Grol R, Hutchinson A, Eccles M, Grimshaw J. Clinical guidelines: potential benefits, limitations, and harms of clinical guidelines. *BMJ*. 1999;318(7182):527-530.
7. Alonso-Coello P, Irfan A, Sola I, et al. The quality of clinical practice guidelines over the last two decades: a systematic review of guideline appraisal studies. *Qual Saf Health Care*. 2010;19(6):e58.
8. Brouwers MC, Kho ME, Browman GP, et al. AGREE II: advancing guideline development, reporting and evaluation in health care. *J Clin Epidemiol*. 2010;63(12):1308-1311.
9. Ferket BS, Colkesen EB, Visser JJ, et al. Systematic review of guidelines on cardiovascular risk assessment: which recommendations should clinicians follow for a cardiovascular health check? *Arch Intern Med*. 2010;170(1):27-40.
10. Bennett WL, Odelola OA, Wilson LM, et al. Evaluation of guideline recommendations on oral medications for type 2 diabetes mellitus: a systematic review. *Ann Intern Med*. 2012;156(1 pt 1):27-36.
11. Beck C, Cody M, Souder E, Zhang M, Small GW. Dementia diagnostic guidelines: methodologies, results, and implementation costs. *J Am Geriatr Soc*. 2000;48(10):1195-1203.
12. Azermai M, Petrovic M, Elseviers MM, Bourgeois J, Van Bortel LM, Vander Stichele RH. Systematic appraisal of dementia guidelines for the management of behavioural and psychological symptoms. *Ageing Res Rev*. 2011;11:78-86.
13. Cummings JL, Frank JC, Cherry D, et al. Guidelines for managing Alzheimer's disease: Part II. Treatment. *Am Fam Physician*. 2002;65(12):2525-2534.
14. Cummings JL, Frank JC, Cherry D, et al. Guidelines for managing Alzheimer's disease: part I. Assessment. *Am Fam Physician*. 2002;65:2263-2272.
15. Watine J, Friedberg B, Nagy E, et al. Conflict between guideline methodologic quality and recommendation validity: a potential problem for practitioners. *Clin Chem*. 2006;52(1):65-72.
16. Nitrini R, Caramelli P, Bottino CM, Damasceno BP, Brucki SM, Anghinah R. [Diagnosis of Alzheimer's disease in Brazil: diagnostic criteria and auxiliary tests. Recommendations of the Scientific Department of Cognitive Neurology and Aging of the Brazilian Academy of Neurology]. *Arq Neuropsiquiatr*. 2005; 63(3A):713-719.
17. Nitrini R, Caramelli P, Bottino CM, Damasceno BP, Brucki SM, Anghinah R. [Diagnosis of Alzheimer's disease in Brazil: cognitive and functional evaluation. Recommendations of the Scientific Department of Cognitive Neurology and Aging of the Brazilian Academy of Neurology]. *Arq Neuropsiquiatr*. 2005; 63(3A):720-727.

18. National Collaborating Centre for Mental Health, National Institute for Health and Clinical Excellence (NICE). Dementia: supporting people with dementia and their carers in health and social care. London, England: National Institute for Health and Clinical Excellence (NICE); 2006:417.

19. Scottish Intercollegiate Guidelines Network. *Management of patients with dementia. A national clinical guideline*. Edinburgh, Scotland: Scottish Intercollegiate Guidelines Network (SIGN); 2006:53.

20. British Columbia Medical Association. Cognitive Impairment in the Elderly-Recognition, Diagnosis and Management. Clinical Practice Guidelines and Protocols in British Columbia: 2007 [cited 2011 July 01]. http://www.bcguidelines.ca/pdf/cognitive.pdf.

21. Singapore Ministry of Health. *Dementia*. Singapore: Singapore Ministry of Health; 2007:80.

22. California Workgroup on Guidelines for Alzheimer's Disease Management. Guideline for Alzheimer's disease management. Chicago, IL: Alzheimer's Association; 2008:61.

23. Chertkow H. Diagnosis and treatment of dementia: introduction. Introducing a series based on the Third Canadian Consensus Conference on the Diagnosis and Treatment of Dementia. *CMAJ.* 2008;178(3):316-321.

24. Chertkow H, Massoud F, Nasreddine Z, et al. Diagnosis and treatment of dementia: 3. Mild cognitive impairment and cognitive impairment without dementia. *CMAJ.* 2008;178(10):1273-85.

25. Feldman HH, Jacova C, Robillard A, et al. Diagnosis and treatment of dementia: 2. Diagnosis. *CMAJ.* 2008;178(7):825-836.

26. Herrmann N, Gauthier S. Diagnosis and treatment of dementia: 6. Management of severe Alzheimer disease. *CMAJ.* 2008;179(12):1279-1287.

27. Hogan DB, Bailey P, Black S, et al. Diagnosis and treatment of dementia: 4. Approach to management of mild to moderate dementia. *CMAJ.* 2008;179(8):787-793.

28. Hogan DB, Bailey P, Black S, et al. Diagnosis and treatment of dementia: 5. Nonpharmacologic and pharmacologic therapy for mild to moderate dementia. *CMAJ.* 2008;179:1019-1026.

29. Patterson C, Feightner JW, Garcia A, Hsiung GY, MacKnight C, Sadovnick AD. Diagnosis and treatment of dementia: 1. Risk assessment and primary prevention of Alzheimer disease. *CMAJ.* 2008;178(5):548-556.

30. Hort J, O'Brien JT, Gainotti G, et al. EFNS guidelines for the diagnosis and management of Alzheimer's disease. *Eur J Neurol.* 2010;17(10):1236-1248.

31. Fillit HM, Doody RS, Binaso K, et al. Recommendations for best practices in the treatment of Alzheimer's disease in managed care. *Am J Geriatr Pharmacother.* 2006;4(Suppl A):S9-S24;quiz S5-S8.

32. Allegri RF, Arizaga RlL, Bavec CV, et al. Enfermedad de Alzheimer. Guía de práctica clínica. *Neurol Argent.* 2011;3(2):120-137.

33. Group Health Cooperative. *Dementia*. Diagnosis and treatment Guideline. Seattle, Wash; 2009:13.

34. Grupo de trabajo de la Guía de práctica clínica sobre la atención integral a las personas con enfermedad de Alzheimer y otras demencias. Guía de práctica clínica sobre la atención integral a las personas con enfermedad de Alzheimer y otras demencias. Agència d'Informació, Avaluació i Qualitat en Salut. Servei Català de la Salut. Pla Director Sociosanitario. Departament de Salut. Generalitat de Catalunya; 2011:508.

35. Instituto Mexicano del Seguro Social. *Guía de Práctica Clínica: Diagnóstico y Tratamiento de la Demencia en el Adulto Mayor en el Primer Nivel de Atención*. México: Instituto Mexicano del Seguro Social; 2010:45.

36. Ministry of Health Malaysia. *Management of Dementia*. 2nd ed. Malaysia: Ministry of Health Malaysia; 2009:144.

37. Regional Health Council. *Dementia. Diagnosis and treatment*. Milan, Italy: Regione Toscana, Consiglio Sanitario Regionale; 2011:38.

38. Scottish Intercollegiate Guidelines Network. SIGN 50: A guideline developer's handbook. Edinburgh, England: 2008 [cited 2012 Jan 01]. http://www.sign.ac.uk/guidelines/fulltext/50/index.html.

39. Schunemann HJ, Fretheim A, Oxman AD. Improving the use of research evidence in guideline development: 13. Applicability, transferability and adaptation. *Health Res Policy Syst.* 2006;4:25.

40. Fretheim A, Schunemann HJ, Oxman AD. Improving the use of research evidence in guideline development: 15. Disseminating and implementing guidelines. *Health Res Policy Syst.* 2006;4:27.

41. Lexchin J, Bero LA, Djulbegovic B, Clark O. Pharmaceutical industry sponsorship and research outcome and quality: systematic review. *BMJ.* 2003;326:1167-1170.

42. Francke AL, Smit MC, de Veer AJ, Mistiaen P. Factors influencing the implementation of clinical guidelines for health care professionals: a systematic meta-review. *BMC Med Inform Decis Mak.* 2008;8:38.

43. Boyd EA, Bero LA. Improving the use of research evidence in guideline development: 4. Managing conflicts of interests. *Health Res Policy Syst.* 2006;4:16.

44. Detsky AS. Sources of bias for authors of clinical practice guidelines. *CMAJ.* 2006;175(9):1033,1055.

45. Guyatt G, Akl EA, Hirsh J, et al. The vexing problem of guidelines and conflict of interest: a potential solution. *Ann Intern Med.* 2010;152(11):738-41.

46. Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med.* 2011;155(8):529-536.

47. Schunemann HJ, Fretheim A, Oxman AD. Improving the use of research evidence in guideline development: 9. Grading evidence and recommendations. *Health Res Policy Syst.* 2006;4:21.

48. McShane R, Noel-Storr A, Ritchie C, Flicker L. The quality and extent of evidence for biomarkers: a Cochrane systematic review. *Alzheimer's & dementia: the journal of the Alzheimer's Association.* 2011;7(4):S100-S1.

49. Schunemann HJ, Oxman AD, Brozek J, et al. GRADE: assessing the quality of evidence for diagnostic recommendations. *Evid Based Med.* 2008;13(6):162-163.

50. Brouwers MC, Kho ME, Browman GP, et al. The Global Rating Scale complements the AGREE II in advancing the quality of practice guidelines. *J Clin Epidemiol.* 2012;65(5):526-534.