Cochrane Library

Cochrane Database of Systematic Reviews

Mini-Mental State Examination (MMSE) for the early detection of dementia in people with mild cognitive impairment (MCI) (Review)

Arevalo-Rodriguez I, Smailagic N, Roqué-Figuls M, Ciapponi A, Sanchez-Perez E, Giannakou A, Pedraza OL, Bonfill Cosp X, Cullum S

Arevalo-Rodriguez I, Smailagic N, Roqué-Figuls M, Ciapponi A, Sanchez-Perez E, Giannakou A, Pedraza OL, Bonfill Cosp X, Cullum S.
Mini-Mental State Examination (MMSE) for the early detection of dementia in people with mild cognitive impairment (MCI).
Cochrane Database of Systematic Reviews 2021, Issue 7. Art. No.: CD010783.
DOI: 10.1002/14651858.CD010783.pub3.

www.cochranelibrary.com

WILEY

TABLE OF CONTENTS

ABSTRACT 1
PLAIN LANGUAGE SUMMARY 2
SUMMARY OF FINDINGS 3
BACKGROUND 8
OBJECTIVES 10
METHODS 10
RESULTS 12
Figure 1. 13
Figure 2. 14
Figure 3. 15
Figure 4. 17
Figure 5. 18
Figure 6. 19
Figure 7. 20
DISCUSSION 21
AUTHORS' CONCLUSIONS 22
ACKNOWLEDGEMENTS 23
REFERENCES 24
CHARACTERISTICS OF STUDIES 30
DATA 54
Test 1. MMSE Conversion to All-cause Dementia 54
Test 2. MMSE Conversion to AD dementia 54
Test 3. MMSE Conversion to Vascular Dementia 54
APPENDICES 55
WHAT'S NEW 66
HISTORY 67
CONTRIBUTIONS OF AUTHORS 67
DECLARATIONS OF INTEREST 67
SOURCES OF SUPPORT 67
DIFFERENCES BETWEEN PROTOCOL AND REVIEW 67
INDEX TERMS 68

[Diagnostic Test Accuracy Review]

Mini-Mental State Examination (MMSE) for the early detection of dementia in people with mild cognitive impairment (MCI)

Ingrid Arevalo-Rodriguez[1], Nadja Smailagic[2], Marta Roqué-Figuls[3], Agustín Ciapponi[4], Erick Sanchez-Perez[5], Antri Giannakou[6], Olga L Pedraza[5], Xavier Bonfill Cosp[7], Sarah Cullum[8]

[1]Clinical Biostatistics Unit, Hospital Universitario Ramón y Cajal (IRYCIS). CIBER Epidemiology and Public Health (CIBERESP), Madrid, Spain. [2]Institute of Public Health, University of Cambridge , Cambridge, UK. [3]Iberoamerican Cochrane Centre, Biomedical Research Institute Sant Pau (IIB Sant Pau), CIBER Epidemiología y Salud Pública (CIBERESP), Barcelona, Spain. [4]Argentine Cochrane Centre, Institute for Clinical Effectiveness and Health Policy (IECS-CONICET), Buenos Aires, Argentina. [5]Neurosciences, Hospital Infantil Universitario de San José-FUCS, Bogotá, Colombia. [6]School of Social and Community Medicine, University of Bristol, Bristol, UK. [7]Iberoamerican Cochrane Centre, Biomedical Research Institute Sant Pau (IIB Sant Pau), Universitat Autònoma de Barcelona, CIBER Epidemiología y Salud Pública (CIBERESP), Barcelona, Spain. [8]Department of Psychological Medicine, University of Auckland, Auckland, New Zealand

Contact: Ingrid Arevalo-Rodriguez, inarev7@yahoo.com.

Editorial group: Cochrane Dementia and Cognitive Improvement Group.
Publication status and date: Edited (no change to conclusions), published in Issue 7, 2021.

Citation: Arevalo-Rodriguez I, Smailagic N, Roqué-Figuls M, Ciapponi A, Sanchez-Perez E, Giannakou A, Pedraza OL, Bonfill Cosp X, Cullum S. Mini-Mental State Examination (MMSE) for the early detection of dementia in people with mild cognitive impairment (MCI). *Cochrane Database of Systematic Reviews* 2021, Issue 7. Art. No.: CD010783. DOI: 10.1002/14651858.CD010783.pub3.

ABSTRACT

Background

Dementia is a progressive global cognitive impairment syndrome. In 2010, more than 35 million people worldwide were estimated to be living with dementia. Some people with mild cognitive impairment (MCI) will progress to dementia but others remain stable or recover full function. There is great interest in finding good predictors of dementia in people with MCI. The Mini-Mental State Examination (MMSE) is the best-known and the most often used short screening tool for providing an overall measure of cognitive impairment in clinical, research and community settings.

Objectives

To determine the accuracy of the Mini Mental State Examination for the early detection of dementia in people with mild cognitive impairment

Search methods

We searched ALOIS (Cochrane Dementia and Cognitive Improvement Specialized Register of diagnostic and intervention studies (inception to May 2014); MEDLINE (OvidSP) (1946 to May 2014); EMBASE (OvidSP) (1980 to May 2014); BIOSIS (Web of Science) (inception to May 2014); Web of Science Core Collection, including the Conference Proceedings Citation Index (ISI Web of Science) (inception to May 2014); PsycINFO (OvidSP) (inception to May 2014), and LILACS (BIREME) (1982 to May 2014). We also searched specialized sources of diagnostic test accuracy studies and reviews, most recently in May 2014: MEDION (Universities of Maastricht and Leuven, www.mediondatabase.nl), DARE (Database of Abstracts of Reviews of Effects, via the Cochrane Library), HTA Database (Health Technology Assessment Database, via the Cochrane Library), and ARIF (University of Birmingham, UK, www.arif.bham.ac.uk). No language or date restrictions were applied to the electronic searches and methodological filters were not used as a method to restrict the search overall so as to maximize sensitivity. We also checked reference lists of relevant studies and reviews, tracked citations in Scopus and Science Citation Index, used searches of known relevant studies in PubMed to track related articles, and contacted research groups conducting work on MMSE for dementia diagnosis to try to locate possibly relevant but unpublished data.

Selection criteria

We considered longitudinal studies in which results of the MMSE administered to MCI participants at baseline were obtained and the reference standard was obtained by follow-up over time. We included participants recruited and clinically classified as individuals with MCI under Petersen and revised Petersen criteria, Matthews criteria, or a Clinical Dementia Rating = 0.5. We used acceptable and commonly used reference standards for dementia in general, Alzheimer's dementia, Lewy body dementia, vascular dementia and frontotemporal dementia.

Data collection and analysis

We screened all titles generated by the electronic database searches. Two review authors independently assessed the abstracts of all potentially relevant studies. We assessed the identified full papers for eligibility and extracted data to create two by two tables for dementia in general and other dementias. Two authors independently performed quality assessment using the QUADAS-2 tool. Due to high heterogeneity and scarcity of data, we derived estimates of sensitivity at fixed values of specificity from the model we fitted to produce the summary receiver operating characteristic curve.

Main results

In this review, we included 11 heterogeneous studies with a total number of 1569 MCI patients followed for conversion to dementia. Four studies assessed the role of baseline scores of the MMSE in conversion from MCI to all-cause dementia and eight studies assessed this test in conversion from MCI to Alzheimer´s disease dementia. Only one study provided information about the MMSE and conversion from MCI to vascular dementia. For conversion from MCI to dementia in general, the accuracy of baseline MMSE scores ranged from sensitivities of 23% to 76% and specificities from 40% to 94%. In relationship to conversion from MCI to Alzheimer's disease dementia, the accuracy of baseline MMSE scores ranged from sensitivities of 27% to 89% and specificities from 32% to 90%. Only one study provided information about conversion from MCI to vascular dementia, presenting a sensitivity of 36% and a specificity of 80% with an incidence of vascular dementia of 6.2%. Although we had planned to explore possible sources of heterogeneity, this was not undertaken due to the scarcity of studies included in our analysis.

Authors' conclusions

Our review did not find evidence supporting a substantial role of MMSE as a stand-alone single-administration test in the identification of MCI patients who could develop dementia. Clinicians could prefer to request additional and extensive tests to be sure about the management of these patients. An important aspect to assess in future updates is if conversion to dementia from MCI stages could be predicted better by MMSE changes over time instead of single measurements. It is also important to assess if a set of tests, rather than an isolated one, may be more successful in predicting conversion from MCI to dementia.

PLAIN LANGUAGE SUMMARY

Baseline scores of Mini-Mental State examination (MMSE) for early prediction of developing dementia in people with mild cognitive impairments (MCI)

Patients with MCI should be evaluated and monitored due to their increased risk of progression to dementia. At present there are no agreements about what the best approach is to register the progression to dementia. Several cognitive function tests have been proposed for this task because most of them are easy to administer, take no longer than 10 minutes to complete, involve major executive functions, and yield an objective score. Our review assessed the current evidence related to one of those brief tests, the Mini-Mental State Examination (MMSE), in the prediction of decline to dementia in people with cognitive impairments. After an extensive search and analysis of available information, we did not find evidence supporting a substantial role of MMSE as a stand-alone single-administration test in the identification of patients who will convert to dementia in the future.

SUMMARY OF FINDINGS

Summary of findings 1. Summary of findings table

What is the diagnostic accuracy of MMSE scores at baseline for identifying those MCI participants who would convert to dementia, Alzheimer´s disease dementia or other forms of dementia over time?							
Patient population	Participants diagnosed with MCI at baseline using Petersen and revised Petersen criteria (Petersen 1999; Petersen 2004), Matthews criteria (Matthews 2008) or Clinical Dementia Rating (CDR) = 0.5 (Morris 1993)						
Prior testing	Most tests are usually performed after a cognitive deficit has been identified						
Settings	Participants were recruited from: i) secondary care - outpatient clinic (n = 3); ii) secondary care – memory clinics (n = 6) and iii) populational sources (n = 2)						
Index tests	MMSE scores at baseline						
Reference standard	NINCDS-ADRDA or DSM or ICD criteria for Alzheimer's disease dementia; McKeith criteria for Lewy body dementia; Lund criteria for frontotemporal dementia; and NINDS-AIREN criteria for vascular dementia						
Target condition	Dementia in general (defined by studies), Alzheimer's disease dementia or other forms of dementia						
Included studies	11 studies (1569 participants) of prospectively cohorts with any accepted definition of MCI were included						
Quality concerns	Patient selection and index test domains were insufficiently reported. Seven studies have not pre-specified thresholds Regarding the reference standard domain, a significant number of studies did not have enough information about the independent interpretation between the MMSE scores and the final diagnosis There were not important concerns about applicability domains in general						
Limitations	Limited investigation of heterogeneity due to insufficient number of studies						
Test	**Studies**	**Cases/participants**	**Median specificity from included studies**	**Sensitivityat median specificity (1)**	**Consequences in a cohort of 100**		
					Median percentage converting (range) (2)	**Missed cases (3)**	**Overdiagnosed**
All-cause dementia							
MMSE scores at baseline	4	255/792	88%	40%	36.5 (23.9 to 40.2)	22	8

Alzheimer's disease dementia							
MMSE scores at baseline	8	374/1128	80%	54%	39.1 (13.3 to 47.6)	18	12
Non-Alzheimer's disease dementia (vascular dementia)							
MMSE scores at baseline	1	22/351	80%	36%	6.26	5	19
Investigation of heterogeneity	The planned investigations were not possible due to the limited number of studies available for each analysis						
Conclusions	Our review did not find the evidence for supporting a substantial role of MMSE test in identifying those MCI patients who would develop dementia. The information included in this review is heterogeneous and does not present a definitive answer about critical issues such as an optimal cut-off, the influence of educational background or even the effects of literacy in the accuracy of MMSE						

(1) Meta-analytic estimate of sensitivity derived from the HSROC model at the median value of specificity computed from the included studies. Summary estimates of sensitivity and specificity were not computed due to high heterogeneity derived from included studies.

(2) The median percentage of conversion and range were computed using all the studies included in the analysis for each target condition.

(3) Missed (false negative) and overdiagnosed (false positive) numbers were computed using the median percentage of conversion for each target condition.

Summary of findings 2. Table of descriptive data

ID	Country of Study	Setting	Years of education	Language(s) of tests	MMSE Diagnostic Threshold	Target condition	Definition of dementia	Follow-up	Definition MCI
Buchhave 2008	Sweden	Memory clinic	Not provided	Unclear	28/29	ADD	DSM III-R + NINCDS-ADRDA	5,2 years (4 to 6.8)	Petersen criteria
Conde-Sala 2012	Spain	Memory clinic	50% had attended school for six to eight years	Spanish	21/22	ADD	DSM-IV-TR or NINCDS-ADRDA	5 years	European AD consortium (EADC) criteria for MCI
Chopard 2009	France	Memory clinic	≥ 12 years of education = 8.5%	Unclear	26/27	All-cause dementia	IADLs scores affected + CDR = 0.5 to 1 and DSM-IV + CDR = 1 + plus NINCDS-	14.9 months (± 4.5)	Presence of cognitive complaints +

							ADRDA or NINDS-AIREN or criteria for DLB or Lund and Manchester criteria		no daily activities affected + CDR = 0.5 + no DSM-IV criteria for dementia
Devanand 2008	USA	Memory clinic	Converters = 14 (4.7) Non-converters= 15.4 (4.1)	English	Not provided	ADD	DSM-IV + NINCDS-ADRDA	3 years	Petersen criteria
Meguro 2007a	Japan	Population sample	8.3 years	Japanese	17/18 for 6 or less years of education 20/21 for 7 to 8 years 23/24 for 10 or more years	All-cause dementia	DSM-IV + CDR = 1 + plus NINCDS-ADRDA or NINDS-AIREN or criteria for DLB or Lund and Manchester criteria	5 years	CDR = 0.5
Meguro 2007b	Japan	Population sample	No provided	Japanese	17/18 for 6 or less years of education 20/21 for 7 to 8 years 23/24 for 10 or more years	All-cause dementia	DSM-IV + CDR = 1	7 years	CDR = 0.5
Modrego 2005	Spain	Outpatient clinic	No provided	Spanish	26/27 from 35; 29/30 from 35	ADD	NINCDS-ADRDA	1 to 5 years	Petersen criteria for amnestic MCI
Modrego 2013	Spain	Outpatient clinic	Converters = elementary: 78.9%; high school: 17.5%; university: 3.5% Non-converters = elementary: 79.2%; high school: 12.5%; university: 8.3%	Spanish	30/31 from 35	ADD	NINCDS-ADRDA	2 years (1 to 4)	CDR = 0.5 + BDRS < 4
Palmqvist 2012	Sweden	Memory clinic	No provided	Swedish	26/27; 9/10 for O&R subscale	ADD	NINCDS-ADRDA	5.9 years (3.2 to 8.8)	Petersen criteria
Pozueta 2011	Spain	Memory clinic	% primary school: Pr-AD = 66.6	Spanish	26/27	ADD	Unclear	2 years	Petersen criteria

			S-MCI = 71.42						
Xu 2002	USA	Outpatient research clinic	Non-dementia = 12.17 (5.23) All dementias = 10.06 (6.11)	English	26/27 for dementia or AD; 25/26 for VaD	All-cause dementia, ADD VaD	NINCDS-ADRDA or NINDS-AIREN or criteria for dementia by Lewy bodies or Lund and Manchester criteria or NINDS-SPSP	3.89 (2 to 17)	Unclear

See 'Characteristics of included studies' for more detailed study descriptors.
Abbreviations: ADD = Alzheimer´s Disease Dementia; CDR = Clinical Dementia Rating; DSM = Diagnostic and Statistical Manual of Mental Disorders; ICD-10 = International Statistical Classification of Diseases and Related Health Problems 10th Revision; MCI = Mild cognitive Impairment; MMSE = Mini-Mental State Examination; NINCDS-ADRDA = National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer´s Disease and Related Disorders Association; NINDS-AIREN = National Institute of Neurological Disorders and Stroke and Association Internationales pour la Recherche et l´Ensignement en Neurosciences; O&R = MMSE´s Orientation and Recall items; Pr-AD = Prodormal AD; USA = United States of America; VaD = Vascular Dementia.

Summary of findings 3. Sensitivity and specificity distribution by cut-off

Cut-off	Buchhave 2008	Conde-Sala 2012	Chopard 2009	Devanand 2008	Meguro 2007a	Meguro 2007b	Modrego 2005	Modrego 2013	Palmqvist 2012	Pozueta 2011	Xu 2002
All-cause dementia											
26/27			0.76/0.40								0.57/0.86
By education level					0.35/0.94	0.23/0.92					
ADD											
21/22		0.44/0.45									
26/27									0.62/0.84	0.64/0.80	0.62/0.83
28/29	0.89/0.33										
29/30 (max 35 points)							0.76/0.67				
30/31 (max 35 points)								0.45/0.87			
No provided				0.27/0.90							
VaD											

	26/27	0.36/0.80

Abbreviations: ADD = Alzheimer´s Disease Dementia; VaD = Vascular Dementia.

BACKGROUND

Dementia is a progressive global cognitive impairment syndrome. In 2010, more than 35 million people worldwide were estimated to be living with dementia, a number that will increase to more than 115 million by 2050 (Ferri 2005; Prince 2013; Wimo 2010). Dementia encompasses a group of neurodegenerative disorders that are characterised by progressive loss of both cognitive function and the ability to perform daily living activities. It can be accompanied by neuropsychiatric symptoms and challenging behaviours of varying type and severity. Its underlying pathology is usually degenerative, and subtypes of dementia include Alzheimer's disease dementia (ADD), vascular dementia, dementia with Lewy bodies and frontotemporal dementia, among others. Considerable overlap may be noted in the clinical and pathological presentations of dementia (MRC CFAS 2001), and ADD and vascular dementia often coexist (Matthews 2009; Savva 2009).

Recently a new type of cognitive function stage called mild cognitive impairment (MCI) has been proposed. MCI refers to a heterogeneous condition and currently 16 different classifications are used to define it (Matthews 2008; Petersen 1999; Petersen 2004; Winblad 2004). Prevalence of MCI varies widely (between 0.1% and 42%) according to the criteria applied, with most systems including memory impairment and absence of cognitive decline as basic conditions for diagnosis (Stephan 2007). As part of the Aging, Demographics, and Memory Study (ADAMS) assessment, Plassman et al estimated the prevalence of cognitive impairment without dementia as 22% in people aged 71 years or older (Plassman 2008). MCI may be classified as amnestic or non-amnestic, according to the presence of clinically significant memory impairment that does not meet the criteria for dementia, or a subtle decline in other functions not related to memory (Petersen 2011).

Over time, people with MCI may experience a gradually progressive cognitive decline and changes in personality and behaviour. When the cognitive impairment in memory, reasoning, language and visuospatial abilities interferes with daily function, individuals are diagnosed with dementia. Research studies indicate that an annual average of 10% to 15% of individuals with MCI may progress to dementia, in particular ADD, but with wide variation depending upon the source of study participants, with self-selected clinic attendees having the highest conversion rates (Bruscoli 2004; Mitchell 2008). Information on long-term cohorts suggests that annual conversion rates range from 4.2% (95% confidence interval (CI) 3.9% to 4.6%) for all-cause dementia to 5.8% (95% CI 5.5% to 6.5%) for ADD (Mitchell 2008).

Establishing a definitive diagnosis of MCI in the presence of subtle symptoms can be challenging. In these cases, it is necessary to document the cognitive decline from the patient's medical history and corroborate it by means of neuropsychological testing, among other suggested tools (Petersen 2001). The American Academy of Neurology recommended in 2001 that patients with MCI should be evaluated and monitored in accordance with their risk of progression to dementia by means of general or brief cognitive screening tools (Petersen 2001). Likewise, the National Institute on Aging and the Alzheimer's Association remarked in 2011 that longitudinal evidence of progressive decline in cognition could support the diagnosis of MCI due to ADD and could allow assessment of the potential benefits of early treatment (Albert 2011).

Usually recognition and assessment of people with suspected dementia in any setting (community, primary care or secondary care) requires a brief test of cognitive function or the use of informant questionnaires, or both (Arevalo-Rodriguez 2013; Moyer 2014). The brief cognitive evaluations needed are usually paper-and-pencil tests that are easy to administer, take no longer than 10 minutes to complete, involve major executive functions and yield an objective score. This final score is useful in determining which individuals need a more comprehensive evaluation (usually identified by low scores) (Boustani 2003). One of these brief cognitive tests is the Mini-Mental State Examination (MMSE) (Folstein 1975), which has become the best-known and the most often used short screening tool for providing an overall measure of cognitive impairment in clinical, research and community settings, although it is now the subject of copyright issues (Nieuwenhuis-Mark 2010).

Systematic assessments of the diagnostic accuracy of brief cognitive tests such as MMSE are scarce (Arevalo-Rodriguez 2014). In 1992, Tombaugh et al presented a narrative review of MMSE studies that emphasised psychometric properties such as reliability and construct validity without evaluating the quality of the included evidence (Tombaugh 1992). Later, Mitchell published a meta-analysis of cross-sectional studies of MMSE and reported different estimations of sensitivity and specificity according to the setting and population (Mitchell 2009). Until now, the relationship between MMSE scores and conversion from MCI to ADD or other dementias has not been evaluated in a systematic fashion.

It is thus the aim of this DTA review for diagnostic test accuracy in dementia to evaluate the ability of the MMSE in such settings as community residences, primary care facilities and memory clinics to identify those people with MCI who will progress to the full clinical syndrome of dementia.

Target condition being diagnosed

In general, dementia as diagnosed is defined by a deficit in more than two cognitive domains that is of sufficient degree to impair functional activities. Symptoms are usually progressive over a period of at least several months and should not be attributable to any other brain disease (American Psychiatric Association 1994). Dementia develops over a trajectory of several years, and it is presumed that during some portion of this time people are asymptomatic and pathology is accumulating (Jack 2011). Individuals or their relatives may notice subtle impairments of recent memory during this time. Gradually, more cognitive domains become involved and difficulty planning complex tasks becomes increasingly apparent. Subtypes of dementia include Alzheimer's disease dementia (ADD) (McKhann 1984; McKhann 2011), vascular dementia (Roman 1993), frontotemporal dementia (Lund and Manchester Groups 1994) and Lewy body dementia (McKeith 1996), among others. Some dementia subtypes are related to other neurological diseases such as Parkinson's disease (Goetz 2008).

This review focused on conversion from MCI to all-cause dementia, ADD, as well as conversion from MCI to other forms of dementia, which were assessed at follow-up. As was previously noted, several studies have shown that most patients with MCI are at increased risk of developing dementia (Petersen 2011). Several medications have been evaluated for use in reducing or delaying the risk of

progression, but none have been adopted for extended clinical use (Farina 2012; Russ 2012; Yue 2012).

Index test(s)

The Folstein Mini-Mental State Examination (MMSE) is a 30-question assessment of cognitive function that evaluates attention and orientation, memory, registration, recall, calculation, language and ability to draw a complex polygon (Folstein 1975). The MMSE has recently been subject to copyright restrictions (de Silva 2010). In its inception, the MMSE was not conceived to identify early stages of dementia, distinguish between different types of dementia or to predict the development of dementia in the long term.

Advantages of the MMSE include rapid administration, availability of multiple language translations and high levels of acceptance as a diagnostic instrument amongst health professionals and researchers (Nieuwenhuis-Mark 2010). The presence of cognitive decline is determined by the total score. Traditionally, a 23/24 cut-off has been used to select patients with suspected cognitive impairment or dementia (Tombaugh 1992). However, several studies have shown that sociocultural variables, age and education, among other factors, could affect individual scores (Bleecker 1988; Brayne 1990; Crum 1993); therefore local standards must be developed for each population and setting evaluated (Diniz 2007; Kulisevsky 2009; Shiroky 2007; Trenkle 2007).

Clinical pathway

Dementia develops over a trajectory of several years. It is presumed that during some portion of this time people are asymptomatic and pathology is accumulating. Individuals or their relatives may notice subtle impairments of recent memory during this time. Gradually, more cognitive domains become involved and difficulty planning complex tasks becomes increasingly apparent. People with memory complaints usually present to their general practitioner (primary care), who may administer one or more brief cognitive tests and potentially refer the individual to a memory clinic (secondary care). However, many people with dementia do not present until much later in the course of the disease and follow a different pathway to diagnosis. In community settings, screening tests are usually administered to estimate the epidemiological figures of dementia, identify cases to be included in clinical trials or even establish a follow-up to detect incident cases or changes in cognitive performance (Brayne 2011). In all cases, a follow-up period is mandatory to detect cognitive changes in populations and conversion of mild cases to dementia (delayed verification).

Standard assessment of dementia includes a history and clinical examination (including neurological, mental state and cognitive examinations); laboratory tests such as thyroid-stimulating hormone, serum folic acid, serum vitamin B_{12} and blood count; an interview with a relative or other informant; and neuroradiological evaluation (Feldman 2008; Hort 2010). Before dementia is diagnosed, other physical and mental disorders (for example hypothyroidism, depression) that might be contributing to cognitive impairment should be excluded or treated. Neuropsychological examination includes full assessment of major cognitive domains, including memory, executive functions, language, attention and visuospatial skills. A neuroradiological examination (computed tomography (CT) or magnetic resonance imaging (MRI) scan of the brain) is also recommended in most recent consensus guidelines (McKhann 2011), although the use of cerebrospinal fluid (CSF) biomarkers is controversial (Dubois 2010). Sometimes the diagnosis is made on the basis of history and presentation alone.

Prior test(s)

Most tests (for example neuroimaging, CSF analysis) are usually performed after a cognitive deficit has been identified. However, it is conceivable that patients with abnormalities on brain imaging, performed for any number of reasons, are likely to be tested subsequently for cognitive deficits.

Role of index test(s)

Accurate diagnosis leads to opportunities for treatment. At the present time, no 'cure' for dementia is known but some treatments can slow cognitive and functional decline or reduce associated behavioural and psychiatric symptoms of dementia (Birks 2006; Clare 2003; McShane 2006). Furthermore, diagnosis of ADD (and other dementias) at an early stage will help people with dementia, their families and potential carers in making timely plans for the future. Coupled with appropriate contingency planning, proper recognition of the disease may help to prevent inappropriate and potentially harmful admissions to hospital or institutional care. In addition, accurate early identification of dementia may increase opportunities for the use of newly evolving interventions designed to delay or prevent progression to more debilitating stages of dementia.

Alternative test(s)

The Cochrane Dementia and Cognitive Improvement Group is undertaking a series of DTA systematic reviews, including a full investigation of other short cognitive tests like the Montreal Cognitive Assessment (Davis 2013a) and the Mini-Cog test (Chan 2014; Fage 2013; Seitz 2014).

Rationale

The public health burden of cognitive and functional impairment due to dementia is of growing concern. With the changing age structure of populations in both high- and low-income countries, the prevalence of dementia is increasing (Ferri 2005; Prince 2013). At the population level, this has major implications for service provision and planning given that the condition leads to progressive functional dependence over several years. Accurate diagnosis leads to opportunities for treatment and appropriate care, but it is also crucial to identify participants for clinical trials of sufficient power to demonstrate the effectiveness of potential treatments.

At the present time, no 'cure' for dementia is known, but some treatments can slow cognitive and functional decline or reduce associated behavioural and psychiatric symptoms of dementia (Birks 2006; Clare 2003; McShane 2006). Furthermore, diagnosis of ADD (and other dementias) at an early stage (that is MCI) will help people with dementia, their families and potential carers in making timely plans for the future. Coupled with appropriate contingency planning, proper recognition of the disease may help prevent inappropriate and potentially harmful admissions to hospital or institutional care. In addition, accurate early identification of dementia may increase opportunities for the use of newly evolving interventions designed to delay or prevent progression to more debilitating stages of disease.

The Cochrane Dementia and Cognitive Improvement Group is undertaking a series of DTA systematic reviews, including three on the accuracy of the MMSE for diagnosing dementia. This review will be focused on evaluation of the MMSE and delayed-verification studies for assessment of conversion from MCI to dementia.

OBJECTIVES

To determine the accuracy of the Mini Mental State Examination for the early detection of dementia in people with mild cognitive impairment

Secondary objectives

To assess the heterogeneity of test accuracy by population (for example memory clinics, community settings) and MMSE thresholds, amongst other factors.

METHODS

Criteria for considering studies for this review

Types of studies

We considered longitudinal studies in which results of the MMSE administered to MCI participants were obtained at baseline and the reference standard was obtained by follow-up over time (at least 12 months). We excluded cross-sectional studies, before-after studies and case reports.

Participants

We included participants recruited from community, primary care and secondary care settings and clinically classified as individuals with MCI at baseline. We established the diagnosis of MCI using Petersen and revised Petersen criteria (Petersen 1999; Petersen 2004), Matthews criteria (Matthews 2008) or Clinical Dementia Rating (CDR) = 0.5 (Morris 1993). These criteria include subjective complaints, decline in memory objectively verified by neuropsychological testing in combination with patient history, decline in other cognitive domains, minimal or no impairment in activities of daily living and not meeting the criteria for dementia. We included all subtypes of MCI participants (amnestic single domain, amnestic multiple domain, non-amnestic single domain and non-amnestic multiple domain). We excluded studies of participants with a secondary cause of cognitive impairment, namely current or past alcohol or drug abuse, central nervous system (CNS) trauma (for example subdural haematoma), tumour and infection, amongst others.

Index tests

The Mini-Mental State Examination (Folstein 1975), or MMSE, is a simple pen-and-paper test of cognitive function based on a total possible score of 30 points; it includes tests of orientation, concentration, attention, verbal memory, naming and visuospatial skills. In follow-up studies, participants with MCI are evaluated by the MMSE to obtain a baseline score and then are followed for several months to allow identification of new cases of dementia. Its utility as a predictive factor could be evaluated for several thresholds, some of them previously specified or otherwise obtained from statistical methods (for example logistic regression); optimal cut-offs are established according to sensitivity and specificity figures, amongst others.

Target conditions

The target condition was conversion at follow-up from MCI to all-cause dementia, Alzheimer's disease dementia (ADD) or other forms of dementia. We expected to find most studies focused on ADD, vascular dementia, Lewy body dementia and frontotemporal dementia.

Reference standards

Currently, no in vivo gold standard is used for the diagnosis of dementia, and even the value of diagnoses based on neuropathological criteria has been questioned (Scheltens 2011). However, we used acceptable and commonly used reference standards. Clinical diagnosis after follow-up includes all-cause (unspecified) dementia, according to recognised diagnostic criteria for example the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV) and the International Classification of Diseases, 10th Revision (ICD-10). National Institute of Neurological and Communicative Disorders and Stroke (NINCDS)-Alzheimer Disease and Related Disorders Association (ADRDA) criteria (McKhann 1984; McKhann 2011) are the best antemortem clinical consensus gold standard for ADD, defining three antemortem groups: probable, possible and unlikely ADD. DSM and ICD definitions are also acceptable classifications for diagnosis of eventual ADD. The reference standard for Lewy body dementia was the McKeith criteria (McKeith 1996; McKeith 2005), for frontotemporal dementia the Lund-Manchester criteria (Lund and Manchester Groups 1994) and for vascular dementia the National Institute of Neurological Disorders and Stroke (NINDS)-Association Internationale pour la Recherche et l'Enseignement en Neurosciences (AIREN) criteria (Roman 1993).

Search methods for identification of studies

Electronic searches

We searched ALOIS (Cochrane Dementia and Cognitive Improvement Specialized Register of diagnostic and intervention studies (inception to May 2014), MEDLINE (OvidSP) (1946 to May 2014), EMBASE (OvidSP) (1980 to May 2014), BIOSIS (Web of Science) (inception to May 2014), Web of Science Core Collection including the Conference Proceedings Citation Index (ISI Web of Science) (inception to May 2014), PsycINFO (OvidSP) (inception to May 2014) and LILACS (BIREME) (1982 to May 2014). We identified grey literature in the form of conference abstracts in a number of our database searches, especially in EMBASE and the Web of Science Core Collection, which includes the Conference Proceedings Citation Index. We designed similarly structured search strategies using search terms appropriate for each database (see Appendix 1 for all the search strategies). We used standardized database subject headings such as MeSH terms (in MEDLINE) and Emtree (in EMBASE) and other standardized headings (controlled vocabulary) in other databases, as appropriate. We did not use search filters designed to retrieve diagnostic test accuracy studies (collections of terms aimed at reducing the number needed to screen by filtering out irrelevant records and retaining only those that are relevant) as a method to restrict the search overall because available filters have not yet proved sensitive enough for systematic review searches (Whiting 2011). We did not apply any language restriction to the electronic searches. We requested a search of the Cochrane Register of Diagnostic Test Accuracy Studies (hosted and maintained by the Cochrane Renal Group) and the specialised register of the Cochrane Dementia and Cognitive Improvement

Group, ALOIS, which includes both intervention and diagnostic test accuracy studies in dementia. A single researcher with extensive experience of systematic reviewing performed the initial searches.

Searching other resources

We checked the reference lists of all relevant papers for additional studies. We also searched:

- MEDION database (Meta-analyses van Diagnostisch Onderzoek), www.mediondatabase.nl;
- DARE (Database of Abstracts of Reviews of Effects), http://www.crd.york.ac.uk/CRDWeb/;
- HTA Database (Health Technology Assessment Database, the Cochrane Library),
- ARIF database (Aggressive Research Intelligence Facility), www.arif.bham.ac.uk.

Through PubMed, relevant studies were used to search for additional studies using the 'Related Articles' feature. We tracked key studies in citation databases such as the Science Citation Index and Scopus to ascertain further relevant studies. We identified grey literature in the form of conference abstracts in a number of our database searches, especially in EMBASE and the Web of Science Core Collection, which includes the Conference Proceedings Citation Index. We also attempted to contact researchers involved in studies with possibly relevant but unpublished data. We did not perform handsearching as the evidence for the benefits of handsearching is not certain. The findings of a recent study investigating handsearching as a method for identifying diagnostic test accuracy studies suggested little additional benefit for handsearching above a robust initial search strategy in a well-indexed and clearly defined subject area (Glanville 2012).

Data collection and analysis

Selection of studies

We selected studies on the basis of title and abstract screening undertaken by the review authors or by teams of experienced assessors. We then located the full paper for each potentially eligible study identified by the search, and two review authors independently evaluated each study for inclusion or exclusion. We resolved disagreements by discussion. If this did not prove conclusive, the default position was to include the study. We presented the study selection process in a PRISMA flow diagram.

Data extraction and management

We extracted data on study characteristics into a study-specific proforma and included data on assessment of quality and investigation of heterogeneity, as described in Appendix 2. The proforma was piloted against five primary diagnostic studies. Two review authors extracted data. We cross-tabulated in 2 × 2 tables the index test results (positive or negative) against the target disorder (positive or negative) and showed results in RevMan tables.

Assessment of methodological quality

We assessed the methodological quality of each study by using the Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) (Whiting 2011a), as recommended by The Cochrane Collaboration. This tool is made up of four domains: patient selection, index test, reference standard and patient flow (see Appendix 3). Each domain was assessed in terms of risk of bias, and the first three domains were also considered in terms of applicability. We reported the QUADAS-2 methodological assessment of studies using bespoke tables. Operational definitions describing the use of QUADAS-2 are detailed in Appendix 4.

Statistical analysis and data synthesis

The target condition comprised two categories: (1) all-cause dementia (not otherwise specified) and (2) dementia subtypes (Alzheimer's, vascular, Lewy body, etc.). Studies may detail one or both outcomes. Each of these target conditions deserved a separate meta-analysis.

For all included studies, the data in the 2 × 2 tables (showing binary test results cross-classified with the binary reference standard) were used to calculate sensitivities and specificities, with their 95% confidence intervals. We presented individual study results graphically by plotting estimates of sensitivities and specificities both in a forest plot and in a receiver operating characteristic (ROC) space. We considered these findings in the light of the previous systematic assessment (using QUADAS-2) of the methodological quality of individual studies. We used RevMan software to document these descriptive analyses. When more than one threshold were reported in an individual study, we presented the graphical findings for all thresholds reported.

Meta-analyses was performed with the metandi command in Stata, version 13 (StataCorp, College Station, Texas) that estimates the parameters for bivariate random-effects models (Leeflang 2008; Reitsma 2005). If no covariates are included in the model, a function of the parameter estimates for the bivariate model allows one to obtain the parameters of the equivalent hierarchical summary ROC model (HSROC) (Harbord 2007). The HSROC model's parameters were used to plot the ROC curve in RevMan. For descriptive purposes, and due to the considerable uncertainty regarding the pooled results, we expressed our results in absolute terms by estimating the diagnostic accuracy of the index test in an hypothetical cohort of 100 MCI patients, using the median specificity and estimating the corresponding sensitivity according to the HSROC model's parameters. This approach has been used in previous diagnostic test accuracy (DTA) Cochrane reviews about dementia (Ritchie 2014; Zhang 2014).

Investigations of heterogeneity

In preliminary analyses, we visually examined forest plots of sensitivity and specificity and summary ROC (SROC) plots to explore the effect of the sources of heterogeneity. We had planned to explore potential sources of heterogeneity across studies through meta-regression, by fitting HSROC models with several pre-specified covariates in SAS. However, there were insufficient studies to conduct a formal investigation of heterogeneity as planned.

Sensitivity analyses

Due to the limited number of studies evaluating MMSE for all-cause dementia, we performed sensitivity analyses only for studies of ADD regarding the version of MMSE test used. This was different to the protocol (Arevalo-Rodriguez 2013a) and is explained in the Differences between protocol and review section.

Assessment of reporting bias

Quantitative methods for exploring reporting bias are not well established for studies of DTA. Specifically, we did not consider funnel plots of the diagnostic odds ratio (DOR) versus the standard error of this estimate.

RESULTS

Results of the search

Our search resulted in 24,357 citations (47 identified in congress reports), of which 17,513 references that were not related to MMSE information were excluded (Figure 1). After that, 6844 records were screened by the authors; excluding 6611 citations. Then 233 records were assessed in more depth in order to apply the inclusion criteria; 186 studies were excluded as they did not include participants with MCI at baseline, and did not assess the conversion from MCI to dementia. Finally, 47 references were retrieved in full text. The review team excluded 35 of them, mainly because they did not provide data about the accuracy of MMSE for conversion to dementia from MCI (see Characteristics of excluded studies). We contacted nine authors to request useable data, of which six responded. Two studies with insufficient data were therefore excluded (Li 2011; Mauri 2012). One study retrieved in abstract form was classified as an 'ongoing study' because the authors presented a protocol in progress but without information about the accuracy of MMSE scores (Hall 2012). The review included 12 references representing 11 datasets with a total of 1569 participants (Summary of findings 2).

Figure 1. Study flow diagram.

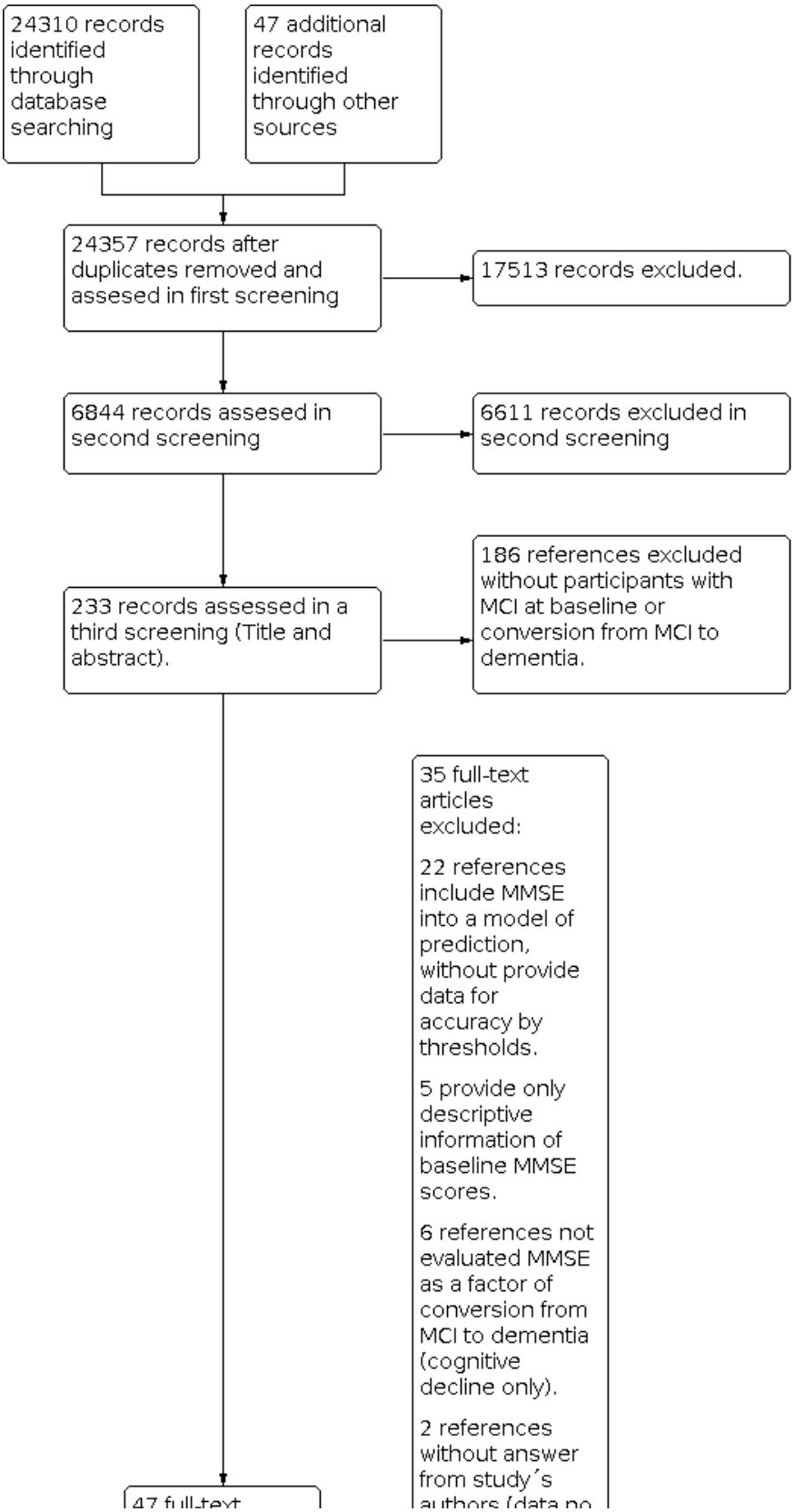
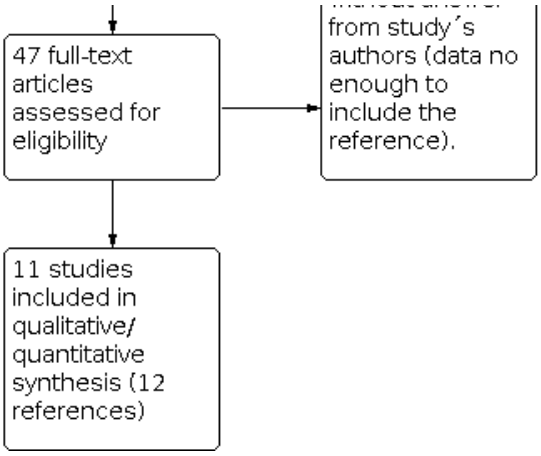


Figure 1. (Continued)



Methodological quality of included studies

We assessed the risk of bias using the QUADAS-2 tool (Appendix 3; Appendix 4). The main results are summarized below (Figure 2; Figure 3).

Figure 2. Risk of bias and applicability concerns graph: review authors' judgements about each domain presented as percentages across included studies.

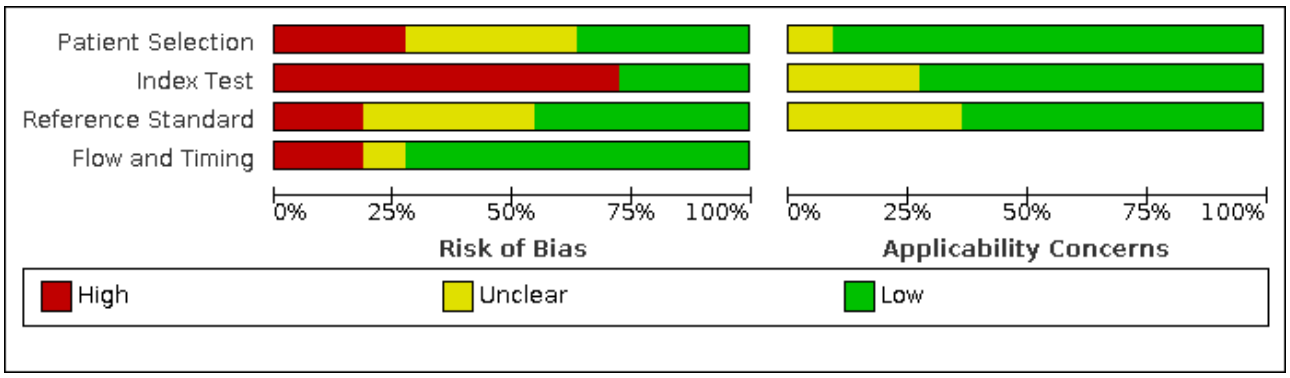


Figure 3. Risk of bias and applicability concerns summary: review authors' judgements about each domain for each included study.

In the patient selection domain we judged three studies (Conde-Sala 2012; Meguro 2007a; Meguro 2007b) to be at high risk of bias due to poor reporting of both the sampling procedure and exclusion criteria. We considered four studies (Buchhave 2008; Chopard 2009; Modrego 2005; Palmqvist 2012) to be at unclear risk of bias because they did not report whether the participants were systematically enrolled. We considered the remaining four studies (Devanand 2008; Modrego 2013; Pozueta 2011; Xu 2002) to be at low risk of bias. We stated that all included studies avoided a case-control design because we only considered data on the performance of the index test to discriminate between patients with MCI who converted to dementia and those who remained stable (that is delayed verification cohort studies).

In the index test domain we considered eight studies (Buchhave 2008; Chopard 2009; Devanand 2008; Modrego 2005; Modrego 2013; Palmqvist 2012; Pozueta 2011; Xu 2002) to be at high risk of bias because the threshold used was not pre-specified and the optimal cut-off level was determined from ROC analyses; therefore, the accuracy of the MMSE reported in these studies appeared to be overestimated. Some studies reported poorly which MMSE version was used and who administered and interpreted the test. We judged the remaining three studies (Conde-Sala 2012; Meguro 2007a; Meguro 2007b) to be at low risk of bias.

In the references´ standard domain we considered four studies (Buchhave 2008; Conde-Sala 2012; Modrego 2013; Xu 2002) to be at 'unclear' risk of bias and two more at high risk (Modrego 2005; Pozueta 2011) because they did not provide enough information about independence and blinding between baseline MMSE scores and the final diagnosis of dementia. These last studies also did not provide enough information about the criteria to establish the conversion from MCI to dementia at follow-up.

In the flow and timing domain we considered the majority of studies (eight) to be at low risk of bias. We judged two (Devanand 2008; Pozueta 2011) to be at high risk of bias and one (Modrego 2005) to be at unclear risk of bias due to loss to follow-up (5% to 15% of losses at follow-up) or poor reporting, or both.

For assessment of applicability concerns, for the majority of the studies (seven) there was no concern that the included patients and setting, the conduct and interpretation of the index test, and the target condition (as defined by the reference standard) did not match the review question. We judged that there was unclear concern about applicability for Modrego 2005 regarding all three domains and for Buchhave 2008, Modrego 2013 and Pozueta 2011 regarding the reference standard domain. It should be noted that the lack of concern about applicability of the three domains mentioned above was based on the inclusion criteria set in the review.

Findings

Included studies are detailed in Characteristics of included studies and Summary of findings 2 and Summary of findings 3. The total number of participants across all included studies was 1569 (median = 109; inter-quartile range (IQR) = 105 to 140). The maximum percentage of losses to follow-up was 15% (Devanand 2008).

One of the references (Meguro 2007a; Meguro 2007b) contained two independent datasets with different follow-ups and we included these as separate entries. Another reference had a single population followed in two different time frames and thresholds (Modrego 2005). We included the information from the longest follow-up (three years) in general analysis. Finally, one of the studies showed information about the accuracy of the Orientation and Recall MMSE subscales (Palmqvist 2012) but this information was not included in our analysis.

More than half of the studies were developed with patients from memory clinics (Buchhave 2008; Chopard 2009; Conde-Sala 2012; Devanand 2008; Palmqvist 2012; Pozueta 2011) with average ages greater than 60 years. In all studies, between 36.3% and 70% of participants were women. Few studies provided descriptive information about social class, years of education, MMSE version used, comorbidities or APOE-ε4 status. No study provided information about pharmacological or non-pharmacological interventions for MCI during the follow-up.

Four different diagnostic thresholds were used to define a positive MMSE (≤ 21, ≤ 26, ≤ 28, ≤ 29). Two additional datasets (Meguro 2007a and Meguro 2007b) considered cut-offs according to individual years of education (≤ 17 for less or equal to 6 years of education, ≤ 20 for 7 to 8 years of education, and ≤ 23 for 10 or more years of education). One additional study provided accuracy for a predicted risk of 0.5 derived from a univariate logistic regression model, instead of a MMSE threshold (Devanand 2008).

Average follow-up times ranged from 15 months to seven years. Median incidence of all-cause dementia in general was 36.5% (4 datasets; IQR = 32.9 to 37.8), while the median incidence of ADD was 39.4% (8 datasets; IQR = 13.3 to 54.2). Only one study provided data for vascular dementia (VaD) incidence (Xu 2002, 6.26%). The scope of the studies included data from five different countries: four studies from Spain; two from Japan, Switzerland and USA; and one from France. In general, sensitivity and specificity figures ranged between 23% and 88% and 32% and 94%, respectively (Summary of findings 3).

Baseline MMSE scores for conversion from MCI to all-cause dementia

Three studies provided numerical data for conversion from MCI to all-cause dementia, with four datasets (n = 792). Sensitivity ranged from 23% to 76% and specificity ranged from 40% to 94% (Figure 4). Figure 5 shows the accuracy estimations of included studies in ROC space along with the SROC curve fitted by the model. We noticed a large lack of precision represented by the 95% confidence interval around the pooled estimates as well as a wide region of prediction, showing a high degree of influence of heterogeneity in this analysis. Under these conditions, there was considerable uncertainty regarding the combination of results from these studies, and the pooled results derived from this model need to be interpreted with caution. The degree of heterogeneity as well as the quality of evidence lowers the level of confidence in the strength of the results. To translate the meta-analysis results into absolute effects, at the median specificity of 88%, we estimated the sensitivity to be 40%. In a hypothetical cohort of 100 MCI patients with a 36.5% incidence of dementia, the number of missed cases would be 18 patients, while 8 MCI patients would be overdiagnosed.

Figure 4. Forest plot of 1 MMSE conversion to all-cause dementia.

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Chopard 2009	29	41	9	27	0.76 [0.60, 0.89]	0.40 [0.28, 0.52]
Meguro 2007a	7	2	13	32	0.35 [0.15, 0.59]	0.94 [0.80, 0.99]
Meguro 2007b	26	14	87	154	0.23 [0.16, 0.32]	0.92 [0.86, 0.95]
Xu 2002	48	38	36	229	0.57 [0.46, 0.68]	0.86 [0.81, 0.90]

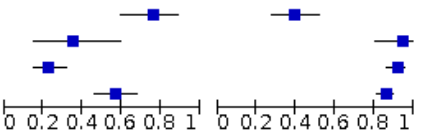
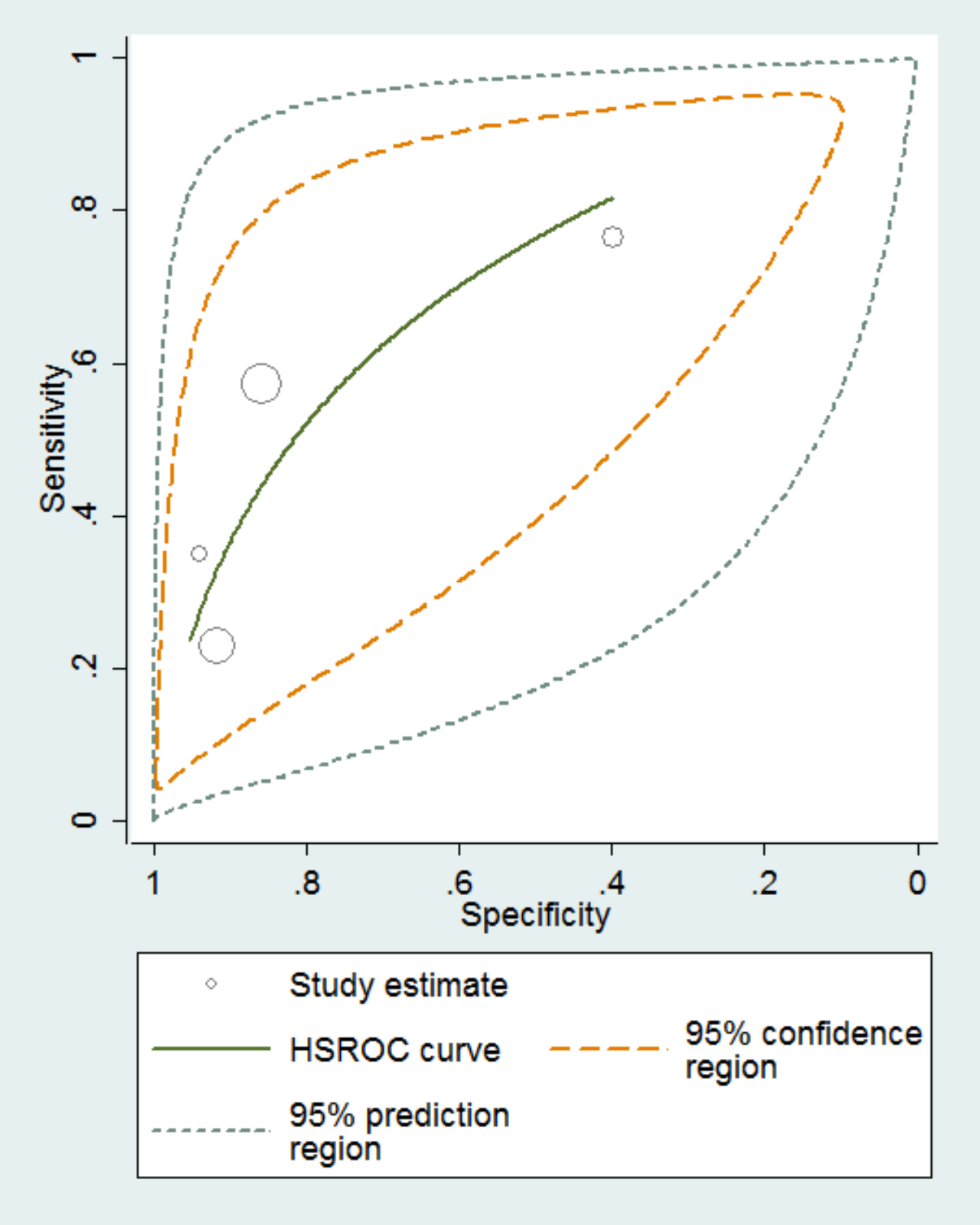


Figure 5. Baseline MMSE scores - conversion from MCI to all-cause dementia.



Baseline MMSE scores for conversion from MCI to Alzheimer's disease dementia (ADD)

Eight studies provided numerical data for conversion from MCI to ADD (n = 1128). Sensitivity ranged from 27% to 89% and specificity ranged from 32% to 90% (Figure 6). Figure 7 shows the accuracy estimates of included studies in ROC space along with the SROC curve fitted by the model. Again, we noticed a high lack of precision represented by the 95% confidence intervals around the pooled estimations as well as a wide region of prediction, showing the large influence of heterogeneity in this analysis. For example, the study of Conde-Sala 2012 (outside the estimated curve) was very different to the rest of studies with its sensitivity and specificity (44% and 45%, respectively) and its cut-off for ADD (21/22 points). Under these conditions, there was considerable uncertainty regarding the combination of results from these studies, and the pooled results derived from this model need to be interpreted with caution. The degree of heterogeneity as well as the quality of evidence lowered our confidence in the strength of the results. To obtain absolute effects, at the median specificity of 80%, the estimated sensitivity was 54%. In a hypothetical cohort of 100 MCI patients with a 39.2% incidence of ADD the number of missed cases would be 18 patients, and 12 MCI patients would be overdiagnosed.

Figure 6. Forest plot of 2 MMSE conversion to ADD.

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Buchhave 2008	56	56	7	28	0.89 [0.78, 0.95]	0.33 [0.23, 0.44]
Conde-Sala 2012	19	36	24	30	0.44 [0.29, 0.60]	0.45 [0.33, 0.58]
Devanand 2008	9	9	24	83	0.27 [0.13, 0.46]	0.90 [0.82, 0.95]
Modrego 2005	22	8	7	16	0.76 [0.56, 0.90]	0.67 [0.45, 0.84]
Modrego 2013	26	6	31	42	0.46 [0.32, 0.59]	0.88 [0.75, 0.95]
Palmqvist 2012	32	13	20	68	0.62 [0.47, 0.75]	0.84 [0.74, 0.91]
Pozueta 2011	32	11	18	44	0.64 [0.49, 0.77]	0.80 [0.67, 0.90]
Xu 2002	29	52	18	252	0.62 [0.46, 0.75]	0.83 [0.78, 0.87]

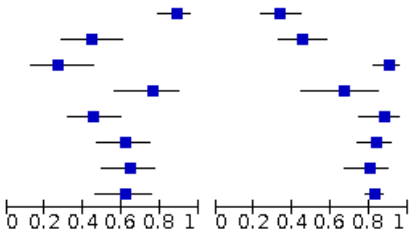
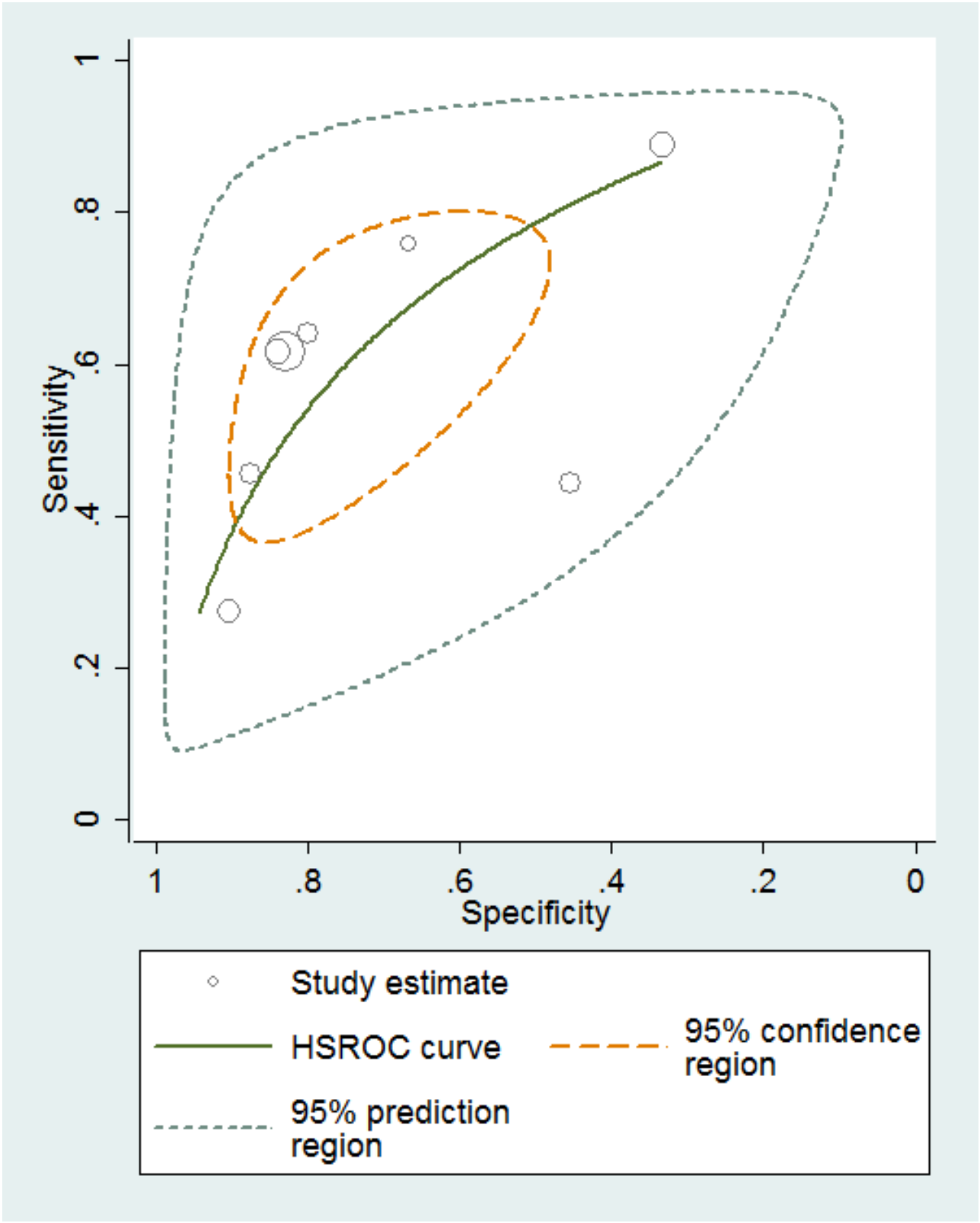


Figure 7. MMSE scores conversion from MCI to ADD.



Baseline MMSE scores for conversion from MCI to vascular dementia (VaD)

Only one study provided information about conversion from MCI to VaD (Xu 2002). This study presented a sensitivity of 36% and a specificity of 80% with an incidence of VaD of 6.2%. In a hypothetical cohort of 100 MCI patients the number of missed cases would be 5 patients, and 19 MCI patients would be overdiagnosed.

Analysis of heterogeneity and sensitivity analysis

Although we had planned to explore possible sources of heterogeneity through meta-regression, including several pre-specified variables, this was not undertaken due to the scarcity of studies included in our meta-analysis. In a narrative description, we noticed that the index test threshold was one of the main sources of variability between included studies. Only five studies (two for conversion to all-cause dementia, three for conversion to ADD and one for conversion to VaD) shared a common threshold (26/27 points), the remaining studies used other cut-offs to classify converters.

We also noticed an important variability in the estimated incidence of dementia. For instance, the incidence for ADD varied between 13% and 54% in MCI samples analysed. The influence of factors such as training of evaluators, education, presence of APOE-4 or onset of medical management was not assessed due to lack of reporting of these variables in the included studies. Related to version of MMSE used, we performed a sensitivity analysis for conversion from MCI to ADD. We removed the data of Modrego 2005 and Modrego 2013 because these studies used a MMSE version with a different scale (35 points). We did not find a significant difference in test accuracy or in perception of heterogeneity when these studies were removed.

Given the modest number of papers and the clinical heterogeneity registered, we did not perform any further sensitivity analysis by risk of bias measured with QUADAS-II items.

DISCUSSION

Summary of main results

In this review we included 11 heterogeneous studies with a total number of 1569 MCI patients followed for conversion to dementia. Four studies assessed the role of baseline scores of MMSE in conversion from MCI to all-cause dementia (defined by the authors of each study) and eight studies assessed this test in conversion from MCI to Alzheimer´s disease dementia (ADD). Only one study provided information about MMSE and conversion from MCI to vascular dementia (VaD). Other dementias, such as frontotemporal dementia or Lewy body dementia, were not assessed by any study. Due to the high heterogeneity and the scarcity of data we could not formally evaluate the influence of factors such as the threshold, the follow-up times, or even the incidence of dementia in the accuracy of this test.

For conversion from MCI to all-cause dementia, we included information from 792 patients (Chopard 2009; Meguro 2007a; Modrego 2005; Xu 2002), 255 of whom developed dementia. Two out of four included studies used CDR scores of 0.5 as the definition of cognitive impairment without dementia. The follow-up time frames were wider in comparison with the rest of studies included in this review (from 14 months to 7 years). Only two studies shared a common cut-off to define dementia (scores ≤ 26) and showed higher specificities and lower sensitivities in comparison with the rest of studies in this group. The accuracy of baseline MMSE scores ranged from sensitivities of 23% (Meguro 2007b) to 76% (Chopard 2009) and specificities from 40% (Chopard 2009) to 94% (Meguro 2007a). We obtained a summary sensitivity of 40% from the SROC curve at the median specificity of 88%. According to this information, MMSE scores appear to have a modest specificity but without the capacity to detect more than half of the MCI converters.

Related to conversion from MCI to ADD, we included information from 1128 patients (Buchhave 2008; Conde-Sala 2012; Devanand 2008; Modrego 2005; Modrego 2013; Palmqvist 2012; Pozueta 2011; Xu 2002) 374 of whom developed ADD. Five out of eight included studies used the Petersen diagnostic criteria (Petersen 1999; Petersen 2004) for defining MCI. The follow-up times in this group of studies were between two and six years. Only three studies (Modrego 2005; Pozueta 2011; Xu 2002) shared a common threshold (scores ≤ 26). The accuracy of baseline MMSE scores ranged from sensitivities of 27% (Devanand 2008) to 89% (Buchhave 2008) and specificities from 32% (Buchhave 2008) to 90% (Devanand 2008). We obtained a summary sensitivity of 54% from the SROC curve at the median specificity of 80%. Again, MMSE scores appear to have a modest specificity but without the capacity to detect near to half of MCI-converters.

Strengths and weaknesses of the review

Our review is part of a series of diagnostic test accuracy (DTA) reviews related to neuropsychological tests on dementia, for which a generic protocol was developed. This protocol identified a priori the best methodology in order to assess the accuracy of cognitive tests in the identification or conversion to any type of dementia (Davis 2013).

This review is the first Cochrane DTA review to assess the role of a well-known paper-and-pencil test (MMSE) in the evaluation of people in the early dementia stage, an entity recently recognized as an important frontier for successful management of this condition. We were challenged in the selection of studies for the research question because we shared a search strategy designed to cover all the DTA reviews related to MMSE and its accuracy in different settings (that is in people over 65 years within a secondary healthcare setting, and asymptomatic and previously clinically unevaluated people aged over 65 years in community and primary care populations). At the same time, with the use of such an exhaustive search strategy we could be sure we included all possible studies, even those with smaller sample sizes, to determine the accuracy of MMSE in predicting conversion from MCI to full dementia.

Although the MMSE is a cognitive test with more than 40 years of use, we only identified studies published since 2002 to answer our research question, possibly due to our specific baseline population (MCI). Also, the review only included studies with a delayed verification of the diagnosis, in contrast to the classic cross-sectional assessment of test accuracy. Such a combination of inclusion criteria had a direct influence on the number of retrieved and included studies in our review. For instance, a considerable number of studies with potential information were excluded because MMSE scores were evaluated as a part of prediction models without providing information about the accuracy of a specified threshold. We hope that in the future we can update our review with

information provided by the contacted authors of excluded studies as well as ongoing studies.

We noticed that most of the included studies did not have the assessment of baseline MMSE scores as a main purpose. In most studies the MMSE was included as part of the usual diagnostic pathway for MCI patients and a common comparator for the principal test assessed. This directly affected the reporting quality and, most importantly, the methodological evaluation of included studies. Our results showed that the index test domain had the greater risk of bias, in most of the cases due to lack of reporting related to administration of MMSE as well as the absence of pre-specified thresholds. We think that the lack of pre-specified thresholds partially explained the high level of heterogeneity among the included studies. The scarcity of information did not allow us to formally assess the influence of this factor in the accuracy of MMSE, but the differences are easily noticed in the analysis of SROC curves. Additional factors that could affect the operative characteristics of MMSE, such as APOE status, the duration of MCI stage at the beginning of the study or even the administration of pharmacological or non-pharmacological interventions were not reported in a consistent way to be considered in our analyses.

Applicability of findings to the review question

Although the MMSE is a test with quick and easy administration, needs few resources and covers multiple cognitive domains at once, it is necessary to remark that this test was not developed to identify the early stages of dementia or even to predict the development of dementia in the long term. In our review, although it was not possible to estimate in a valid way the pooled operating characteristics, the descriptive data provided by the studies showed that neither the sensitivity nor specificity exceeded 80% at the same time. Only one study (Modrego 2005) shows a balance between accuracy figures, estimating a sensitivity of 76% and a specificity of 67% for the conversion of MCI to ADD, but derived from information with a high risk of bias. Our results suggest that MMSE may be of value to decrease the post-test probability of progression to dementia in the presence of normal test scores that confirm the possible conversion. These results may show that the items of the MMSE are insufficient to detect the change from mild to advanced cognitive decline, or even that some factors such as age, education and literacy must be taken into account to determine its true value in MCI patients. Likewise, this brief cognitive test may be more useful to document cognitive changes over time rather than to predict future progression with a single measurement. The verification of this hypothesis requires the assessment of evidence not presented in this review.

AUTHORS' CONCLUSIONS

Implications for practice

At present, there is consensus about the clinical relevance of MCI because this stage represents an opportunity to prevent or delay progression to dementia through modifications of risk factors such as depression and hypertension. The identification of diagnostic tools to predict which patients may progress to more severe stages of the disease has become a priority. The role of cognitive tests in the diagnosis of dementia is not questioned,because they show the clinical decline in areas like speed of information processing, executive functions and reasoning (Sperling 2011), while the role of biomarkers remains under evaluation (Kokkinou 2014; Ritchie 2014; Vacante 2013; Zhang 2014).

The Mini-Mental State Examination (MMSE) is a brief neuropsychological test that provides an overview of cognitive function which, in the setting of patients with MCI, is supplemented with more specialized neuropsychological tests for other domains of language, praxis and executive functions, among others. MMSE advantages reside in the easy way of administration (especially in terms of time and resources) without direct harmful effects, as well as a high acceptability by the health professionals involved in the management of people with dementia. In fact this popular test is frequently administered by clinicians to MCI patients and our review could help them to interpret the results of MMSE of their patients. Ideally, this brief cognitive test could be used for initial classification of MCI patients in order to determine their needs in a further and comprehensive assessment. However, there might be occasions where an extensive and formal neuropsychological evaluation is not available. In both cases it is important that clinicians know the limitations related to the use of this test in the prediction of dementia for MCI patients.

Our review did not find evidence supporting a substantial role of MMSE as a stand-alone single-administration test in the identification of MCI patients who may develop dementia. For example, a MCI patient with a baseline probability of 39% to develop ADD in the next three to four years (median incidence of ADD in our studies) only increases his or her post-test probability to 63% (95% CI 49 to 75) using a MMSE score indicator of progression (LR+ = 2.67), while a negative MMSE score for progression only decreases his or her post-test probability to 27% (LR - = 0.58, 95% CI 20 to 34). In the case of progression to dementia in general we found similar results. In all cases, clinicians would prefer to request additional and extensive tests to be sure about the management of these patients. Also, the review has not been able to address some critical issues such as an optimal cut-off, the influence of educational background, or even the effects of literacy in the accuracy of MMSE.

We think that MMSE items, despite the fact that they cover several cognitive domains, are insufficient for registering subtle cognition changes in MCI patients, especially for detecting those dementias without an important decline in the memory domain (such as frontotemporal dementia and primary progressive apraxia). It is important that clinicians are aware of the limitations related to the use of MMSE as a stand-alone single-administration test and seek to either use MMSE as a follow-up to detect changes in time, or to use it in the context of comprehensive assessments with more specialized neuropsychological tests for other domains of language, praxis and executive functions.

Implications for research

At present, the identification of useful cognitive tests that are able to detect subtle cognitive changes in people at different stages of dementia has become an important challenge. Although the information included in this review does not support the extended use of the MMSE in the stage of progression of MCI to dementia, we should not forget that this kind of test could be useful in settings where formal neuropsychological assessment is not available. In order to determine, with more information, the true operative characteristics of this test future research could focus on the evaluation of unique and pre-specified diagnostic

thresholds, as well as improving the appropriate reporting of variables such as education levels and literacy. Likewise, it is essential to know if there are subsets of MMSE items which could be more strongly associated with progression to dementia. If such a subset of items exists, clinicians will be confident that those results denote conversion to a particular kind of dementia (for instance, conversion to ADD determined by the memory items of the MMSE).

An important aspect to consider in future updates is if conversion to dementia from MCI stages could be predicted better by MMSE changes over time instead of isolated measurements. The role of repeated measurements of MMSE might be more informative than baseline scores, but a formal evaluation of the utility of numerical differences between serial MMSEs has not been performed yet. Finally, although the alternative of a single indicator for progression to dementia is attractive, it would be more interesting to assess if a set of tests, rather than an isolated one, may be more successful in predicting conversion from MCI to dementia. A diagnostic model composed of different cognitive and biomarker tools could be more useful for this condition.

ACKNOWLEDGEMENTS

Ingrid Arévalo-Rodríguez is a PhD student at the Department of Pediatrics, Obstetrics and Gynecology and Preventive Medicine of the Universitat Autònoma de Barcelona.

The authors would like to thank Anna Noel-Storr, Trials Search Co-ordinator of the Cochrane Dementia and Cognitive Improvement Group, for her assistance with writing the search strategy, searching and initial screening of search results.

REFERENCES

References to studies included in this review

Buchhave 2008 *{published data only}*

Buchhave P, Stomrud E, Warkentin S, Blennow K, Minthon L, Hansson O. Cube copying test in combination with rCBF or CSF A beta 42 predicts development of Alzheimer's disease. *Dementia and Geriatric Cognitive Disorders.* 2008;**25**(6):544-52. [PMID: 18535375]

Chopard 2009 *{published data only}*

Chopard G, Vanholsbeeck G, Tio G, Pitard A, Binetruy M, Rumbach L, et al. Rapid screening of cognitive change in patients with questionable dementia using the Memory Impairment Screen and the Isaacs Set Test. *Journal of the American Geriatrics Society* 2009;**57**(4):703-8. [PMID: 19220561]

Conde-Sala 2012 *{published data only}*

* Conde-Sala JL, Garre-Olmo J, Vilalta-Franch J, Llinas-Regla J, Turro-Garriga O, Lozano-Gallego M, et al. Predictors of cognitive decline in Alzheimer's disease and mild cognitive impairment using the CAMCOG: a five-year follow-up. *International Psychogeriatrics* 2012;**24**(6):948-58. [PMID: 22278151]

Devanand 2008 *{published data only}*

* Devanand DP, Liu X, Tabert MH, Pradhaban G, Cuasay K, Bell K, et al. Combining early markers strongly predicts conversion from mild cognitive impairment to Alzheimer's disease. *Biological Psychiatry* 2008;**64**(10):871-9. [PMID: 18723162]

Devanand DP, Pradhaban G, Liu X, Khandji A, De Santi S, Segal S, et al. Hippocampal and entorhinal atrophy in mild cognitive impairment: prediction of Alzheimer disease. *Neurology* 2007;**68**(11):828-36. [PMID: 17353470]

Meguro 2007a *{published data only}*

* Meguro K, Ishii H, Kasuya M, Akanuma K, Meguro M, Kasai M, et al. Incidence of dementia and associated risk factors in Japan: The Osaki-Tajiri Project. *Journal of the Neurological Sciences.* 2007;**260**(1-2):175-82. [PMID: 17553526]

Nakata E, Kasai M, Kasuya M, Akanuma K, Meguro M, Ishii H, et al. Combined memory and executive function tests can screen mild cognitive impairment and converters to dementia in a community: the Osaki-Tajiri project. *Neuroepidemiology* 2009;**33**(2):103-10. [PMID: 19494551]

Meguro 2007b *{published data only}*

Meguro K, Ishii H, Kasuya M, Akanuma K, Meguro M, Kasai M, et al. Incidence of dementia and associated risk factors in Japan: The Osaki-Tajiri Project. *Journal of the Neurological Sciences.* 2007;**260**(1-2):175-82. [PMID: 17553526]

Modrego 2005 *{published data only}*

Modrego PJ, Fayed N, Pina MA. Conversion from mild cognitive impairment to probable Alzheimer's disease predicted by brain magnetic resonance spectroscopy. *American Journal of Psychiatry* 2005;**162**(4):667-75. [PMID: 15800137]

Modrego 2013 *{published data only}*

Modrego PJ, Gazulla J. The predictive value of the memory impairment screen in patients with subjective memory complaints: a prospective study. The Primary Care Companion for CNS Disorders 2013;**15**(1). [2155-7772: (Print)]

Palmqvist 2012 *{published data only}*

* Palmquist S, Hertze J, Minthon L, Wattmo C, Zetterberg H, Blennow K, et al. Comparison of brief cognitive tests and CSB biomarkers in predicting Alzheimer's disease in mild cognitive impairment: six-year follow-up study. *PLoS One* 2012;**7**(6):e38639. [PMID: 22761691]

Pozueta 2011 *{published data only}*

* Pozueta A, Rodriguez-Rodriguez E, Vazquez-Higuera J L, Mateo I, Sanchez-Juan P, Gonzalez-Perez S, et al. Detection of early Alzheimer's disease in MCI patients by the combination of MMSE and an episodic memory test. *BMC Neurology* 2011;**11**:78. [PMID: 21702929]

Xu 2002 *{published data only}*

Xu G, Meyer JS, Thornby J, Chowdhury M, Quach M. Screening for mild cognitive impairment (MCI) utilizing combined mini-mental-cognitive capacity examinations for identifying dementia prodromes. *International Journal of Geriatric Psychiatry.* 2002;**17**(11):1027-33. [PMID: 12404652]

References to studies excluded from this review

Aevarsson 2000 *{published data only}*

Aevarsson O, Skoog I. A longitudinal population study of the mini-mental state examination in the very old: relation to dementia and education. *Dementia and Geriatric Cognitive Disorders* 2000;**11**(3):166-75. [PMID: 10765048]

Apostolova 2006 *{published data only}*

Apostolova LG, Dutton RA, Dinov ID, Hayashi KM, Toga AW, Cummings JL, et al. Conversion of mild cognitive impairment to Alzheimer disease predicted by hippocampal atrophy maps. *Archives of Neurology* 2006;**63**(5):693-9. [PMID: 16682538]

Armas 2009 *{published data only}*

Armas J. Clinical and neuropsychological risk factors to conversion from mild cognitive impairment to Alzheimer disease. *Alzheimer's & Dementia: The Journal of the Alzheimer's Association* 2009;**5**(4):P382.

Brodaty 2011 *{published data only}*

Brodaty H, Woodward M, Boundy K, Ames D, Balshaw R. Patients in australian memory clinics: baseline characteristics and predictors of decline at six months. *International Psychogeriatrics* 2011;**23**(7):1086-96. [PMID: 21489344]

Bruck 2013 *{published data only}*

Bruck A, Virta JR, Koivunen J, Koikkalainen J, Scheinin NM, Helenius H, et al. [11C]PIB, [18F]FDG and MR imaging in patients with mild cognitive impairment. European Journal of

Nuclear Medicine and Molecular Imaging 2013;**40**(10):1567-72. [1619-7089: (Electronic)]

Chan 2011 *{published data only}*

Chan WC, Lam LC, Tam CW, Lui VW, Leung GT, Lee AT, et al. Neuropsychiatric symptoms are associated with increased risks of progression to dementia: A 2-year prospective study of 321 Chinese older persons with mild cognitive impairment. *Age and Ageing* 2011;**40**(1):30-5. [PMID: 21106558]

Chilovi 2011 *{published data only}*

Chilovi BV, Caratozzolo S, Mombelli G, Zanetti M, Rozzini L, Padovani A. Does reversible mci exist? *Alzheimer's & Dementia* 2011;**7**(4):S548.

Choi 2013 *{published data only}*

Choi HJ, Lee DY, Seo EH, Sohn BK, Choe YM, Woo JI. Pib-negative amnestic mild cognitive impairment related with low plasma apolipoprotein a1 level. *Alzheimer's & Dementia* 2013;**9**(4):P698.

Cruz 2012 *{published data only}*

Cruz DM, Allen CM, Malmstrom TK, Tumosa N, Morley JE. Does the veterans affairs saint louis university mental status (SLUMS) exam predict the course of cognitive impairment? *Journal of the American Geriatrics Society* 2012;**60**:S180.

Devanand 2010 *{published data only}*

Devanand DP, Van Heertum RI, Kegeles LS, Liu X, Jin ZH, Pradhaban G, et al. (99m)Tc hexamethyl-propylene-aminoxime single-photon emission computed tomography prediction of conversion from mild cognitive impairment to Alzheimer disease. *American Journal of Geriatric Psychiatry* 2010;**18**(11):959-72. [PMID: 20808143]

Devier 2009 *{published data only}*

Devier DJ, Pelton GH, Tabert MH, Liu X, Cuasay K, Eisenstadt R, et al. The impact of anxiety on conversion from mild cognitive impairment to Alzheimer's disease. *International Journal of Geriatric Psychiatry* 2009;**24**(12):1335-42. [PMID: 19319929]

Devier 2010 *{published data only}*

Devier DJ, Villemarette-Pittman N, Brown P, Pelton G, Stern Y, Sano M, et al. Predictive utility of type and duration of symptoms at initial presentation in patients with mild cognitive impairment. *Dementia and Geriatric Cognitive Disorders* 2010;**30**(3):238-44. [PMID: 20847554]

Ehrensperger 2010 *{published data only}*

Ehrensperger MM, Berres M, Taylor KI, Monsch AU. Screening properties of the German IQCODE with a two-year time frame in MCI and early Alzheimer's disease. *International Psychogeriatrics* 2010;**22**(1):91-100. [PMID: 19747425]

Ewers 2007 *{published data only}*

Ewers M, Buerger K, Teipel SJ, Scheltens P, Schroder J, Zinkowski RP, et al. Multicenter assessment of CSF-phosphorylated tau for the prediction of conversion of MCI. *Neurology* 2007;**69**(24):2205-12. [PMID: 18071141]

Hampel 2004 *{published data only}*

Hampel H, Teipel SJ, Fuchsberger T, Andreasen N, Wiltfang J, Otto M, et al. Value of CSF beta-amyloid1-42 and tau as predictors of Alzheimer's disease in patients with mild cognitive impairment. *Molecular Psychiatry* 2004;**9**(7):705-10. [PMID: 14699432]

Ito 2013 *{published data only}*

Ito K, Mori E, Fukuyama H, Ishii K, Washimi Y, Asada T, et al. Prediction of outcomes in MCI with (123)I-IMP-CBF SPECT: a multicenter prospective cohort study. Annals of Nuclear Medicine 2013;**27**(10):898-906. [1864-6433: (Electronic)]

Koepsell 2012 *{published data only}*

Koepsell TD, Monsell SE. Reversion from mild cognitive impairment to normal or near-normal cognition: risk factors and prognosis. *Neurology* 2012;**79**(15):1591-8. [PMID: 23019264]

Korf 2004 *{published data only}*

Korf ES, Wahlund LO, Visser PJ, Scheltens P. Medial temporal lobe atrophy on MRI predicts dementia in patients with mild cognitive impairment. *Neurology* 2004;**63**(1):94-100. [PMID: 15249617]

Kruczyk 2012 *{published data only}*

Kruczyk M, Zetterberg H, Hansson O, Rolstad S, Minthon L, Wallin A, et al. Monte Carlo feature selection and rule-based models to predict Alzheimer's disease in mild cognitive impairment. *Journal of Neural Transmission* 2012;**119**(7):821-31. [PMID: 22573144]

Li 2011 *{published data only}*

Li J, Wang YJ, Zhang M, Xu ZQ, Gao CY, Fang CQ, et al. Vascular risk factors promote conversion from mild cognitive impairment to Alzheimer disease. *Neurology* 2011;**76**(17):1485-91. [PMID: 21490316]

Luck 2012 *{published data only}*

Luck T, Luppa M, Wiese B, Breitner J, Scherer M, Maier W, et al. Mild cognitive impairment: Determinant patterns of short and long time to incident dementia. *Alzheimer's & Dementia* 2012;**8**(4):P151-2.

Madureira 2010 *{published data only}*

Madureira S, Verdelho A, Moleiro C, Ferro JM, Erkinjuntti T, Jokinen H, et al. Neuropsychological predictors of dementia in a three-year follow-up period: data from the LADIS study. Dementia and Geriatric Cognitive Disorders 2010;**29**(4):325-34. [PMID: 20389074]

Mauri 2012 *{published data only}*

Mauri M, Sinforiani E, Corbetta S, Zucchella C, Giarracca V, Bono G. Progression to dementia of a population with aMCI: Clinical variables associated to conversion. *Journal of Alzheimer's Disease* 2011 Suppl 1:S66.

* Mauri M, Sinforiani E, Zucchella C, Cuzzoni MG, Bono G. Progression to dementia in a population with amnestic mild cognitive impairment: clinical variables associated with conversion. *Functional Neurology* 2012;**27**(1):49-54. [PMID: 22687167]

Meyer 2002 *{published data only}*

Meyer J, Xu G, Thornby J, Chowdhury M, Quach M. Longitudinal analysis of abnormal domains comprising mild cognitive impairment (MCI) during aging. *Journal of the Neurological Sciences* 2002;**201**(1-2):19-25. [PMID: 12163189]

Ott 2013 *{published data only}*

Ott B, Grace J, Frakey L, Kelley P, Tremont G. Prediction of functional decline and conversion from mild cognitive impairment with the telephone-administered Minnesota Cognitive Acuity Screen. *Alzheimer's & Dementia* 2013;**9**(4):P449.

Ouchi 2012 *{published data only}*

Ouchi Y, Akanuma K, Meguro M, Kasai M, Ishii H, Meguro K. Impaired instrumental activities of daily living affect conversion from mild cognitive impairment to dementia: the Osaki-Tajiri Project. *Psychogeriatrics* 2012;**12**(1):34-42. [PMID: 22416827]

Paajanen 2014 *{published data only}*

Paajanen T, Hanninen T, Tunnard C, Hallikainen M, Mecocci P, Sobow T, et al. Cerad memory composite score is an accurate neuropsychological predictor of progression to Alzheimer's disease: AddNeuroMed study. *Alzheimer's & Dementia* 2011;**7**(4):S259.

* Paajanen T, Hanninen T, Tunnard C, Hallikainen M, Mecocci P, Sobow T, et al. CERAD neuropsychological compound scores are accurate in detecting prodromal alzheimer's disease: a prospective AddNeuroMed study. Journal of Alzheimer's Disease 2014;**39**(3):679-90. [1875-8908: (Electronic)]

Rosenberg 2013 *{published data only}*

Rosenberg PB, Mielke MM, Appleby BS, Oh ES, Geda YE, Lyketsos CG. The Association of Neuropsychiatric Symptoms in MCI with Incident Dementia and Alzheimer Disease. *American Journal of Geriatric Psychiatry* 2013;**21**(7):685-95. [PMID: 23567400]

Serrano 2007 *{published data only}*

Serrano CM, Taragano F, Allegri RF, Krupitzki H, Martelli M, Feldman M, et al. Conversion predictors factors in mild cognitive impairment [Factores predictores de conversion en deterioro cognitivo leve (cohorte de seguimiento en CEMIC)]. *Revista Neurologica Argentina* 2007;**32**(2):75-93.

Tardif 2013 *{published data only}*

Tardif M, Roy M, Remi B, Laforce R, Verret L, Fortin M-P, et al. Months backward test as a reliable predictor of cognitive decline in mild Alzheimer's disease. *Alzheimer's & Dementia* 2013;**9**(4):P741-2.

Van Rossum 2011 *{published data only}*

Van Rossum I, Burns L, Scheltens P, Soininen H, Wahlund LO, Hampel H, et al. High csf tau predicts rapid decline to Alzheimer's type dementia in mci subjects with abnormal CSF As1-42. *Alzheimer's & Dementia* 2011;**7**(4):S548.

van Rossum 2012 *{published data only}*

van Rossum IA, Visser PJ, Knol DL, van der Flier WM, Teunissen CE, Barkhof F, et al. Injury markers but not amyloid markers are associated with rapid progression from mild cognitive impairment to dementia in Alzheimer's disease. *Journal of Alzheimer's Disease* 2012;**29**(2):319-27.

Waldorff 2012 *{published data only}*

Waldorff FB, Siersma V, Vogel A, Waldemar G. Subjective memory complaints in general practice predicts future dementia: a 4-year follow-up study. International Journal of Geriatric Psychiatry 2012;**27**(11):1180-8. [1099-1166: (Electronic)]

Wong 2013 *{published data only}*

Wong CH, Leung GT, Fung AW, Chan WC, Lam LC. Cognitive predictors for five-year conversion to dementia in community-dwelling Chinese older adults. International Psychogeriatrics / IPA 2013;**25**(7):1125-34. [1741-203X: (Electronic)]

Zhang 2012 *{published data only}*

Zhang D, Shen D, Alzheimer's Disease Neuroimaging Initiative. Predicting future clinical changes of MCI patients using longitudinal and multimodal biomarkers. *PLoS One* 2012;**7**(3):e33182. [PMID: 22457741]

References to ongoing studies

Hall 2012 *{published data only}*

Hall J, Plassman BI, Steffens D. Utility of NPI Scores predicting progression of CIND to dementia. *American Journal of Geriatric Psychiatry* 2012;**20**(3 Suppl 1):S8-167.

Additional references

Albert 2011

Albert MS, DeKosky ST, Dickson D, Dubois B, Feldman HH, Fox NC, et al. The diagnosis of mild cognitive impairment 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 & Dementia 2011;**7**(3):270-9.

American Psychiatric Association 1994

American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders (DSM-IV). Washington DC: American Psychiatric Association, 1994.

Arevalo-Rodriguez 2013

Arevalo-Rodriguez I, Pedraza OL, Rodríguez A, Sánchez E, Gich I, Solà I, et al. Alzheimer's disease dementia guidelines for diagnostic testing: a systematic review. *American Journal of Alzheimer's Disease and Other Dementias* 2013;**28**(2):111-9. [PMID: 23288575]

Arevalo-Rodriguez 2013a

Arevalo-Rodriguez I, Smailagic N, Ciapponi A, Sanchez-Perez E, Giannakou A, Roqué i Figuls M, at al. Mini-Mental State Examination (MMSE) for the detection of Alzheimer's disease and other dementias in people with mild cognitive impairment (MCI). *Cochrane Database of Systematic Reviews* 2013, Issue 10. Art. No: CD010783. [DOI: 10.1002/14651858.CD010783]

Arevalo-Rodriguez 2014

Arevalo-Rodriguez I, Segura O, Sola I, Bonfill X, Sanchez E, Alonso-Coello P. Diagnostic tools for alzheimer´s disease dementia and other dementias: an overview of diagnostic test accuracy (DTA) systematic reviews. *BMC Neurology* 2014;**14**(1):183. [PMID: 25248284]

Birks 2006

Birks J. Cholinesterase inhibitors for Alzheimer's disease. *Cochrane Database of Systematic Reviews* 2006, Issue 1. Art. No: CD005593. [DOI: 10.1002/14651858.CD005593] [PMID: 16437532]

Bleecker 1988

Bleecker ML, Bolla-Wilson K, Kawas C, Agnew J. Age-specific norms for the Mini-Mental State Exam. *Neurology* 1988;**38**(10):1565-8. [PMID: 3419600]

Boustani 2003

Boustani M, Peterson B, Hanson L, Harris R, Lohr KN. Screening for dementia in primary care: a summary of the evidence for the U.S. Preventive Services Task Force. *Annals of Internal Medicine* 2003;**138**(11):927-37. [PMID: 12779304]

Brayne 1990

Brayne C, Calloway P. The association of education and socioeconomic status with the Mini Mental State Examination and the clinical diagnosis of dementia in elderly people. *Age and Ageing* 1990;**19**(2):91-6. [PMID: 2337015]

Brayne 2011

Brayne C, Stephan BC, Matthews FE. A European perspective on population studies of dementia. *Alzheimer's & Dementia* 2011;**7**(1):3-9. [PMID: 21255739]

Bruscoli 2004

Bruscoli M, Lovestone S. Is MCI really just early dementia? A systematic review of conversion studies. *International Psychogeriatrics* 2004;**16**(2):129-40. [PMID: 15318760]

Chan 2014

Chan Calvin CH, Fage Bruce A, Smailagic N, Gill Sudeep S, Herrmann N, Nikolaou V, et al. Mini-Cog for the diagnosis of Alzheimer's disease dementia and other dementias within a secondary care setting. *Cochrane Database of Systematic Reviews* 2014, Issue 12. Art. No: CD011414. [DOI: 10.1002/14651858.CD011414] [CD011414]

Clare 2003

Clare L, Woods RT, Moniz Cook ED, Orrell M, Spector A. Cognitive rehabilitation and cognitive training for early-stage Alzheimer's disease and vascular dementia. *Cochrane Database of Systematic Reviews* 2003, Issue 4. Art. No: CD003260. [DOI: 10.1002/14651858.CD003260] [PMID: 14583963]

Crum 1993

Crum RM, Anthony JC, Bassett SS, Folstein MF. Population-based norms for the Mini-Mental State Examination by age and educational level. *Journal of the American Medical Association* 1993;**269**(18):2386-91. [PMID: 8479064]

Davis 2013

Davis Daniel HJ, Creavin Sam T, Noel-Storr A, Quinn Terry J, Smailagic N, Hyde C, et al. Neuropsychological tests for the diagnosis of Alzheimer's disease dementia and other dementias: a generic protocol for cross-sectional and delayed-verification studies. *Cochrane Database of Systematic Reviews* 2013, Issue 3. Art. No: CD010460. [DOI: 10.1002/14651858.CD010460] [CD010460]

Davis 2013a

Davis DHJ, Creavin ST, Yip JLY, Noel-Storr AH, Brayne C, Cullum S. The Montreal Cognitive Assessment for the diagnosis of Alzheimer's disease and other dementia disorders. *Cochrane Database of Systematic Reviews* 2013, Issue 10. Art. No: CD010775. [DOI: 10.1002/14651858.CD010775] [CD010775]

de Silva 2010

de Silva V, Hanwella R. Why are we copyrighting science? *BMJ* 2010;**341**:c4738. [PMID: 20847026]

Diniz 2007

Diniz BS, Yassuda MS, Nunes PV, Radanovic M, Forlenza OV. Mini-Mental State Examination performance in mild cognitive impairment subtypes. *International Psychogeriatrics* 2007;**19**(4):647-56. [PMID: 17502007]

Dubois 2010

Dubois B, Feldman HH, Jacova C, Cummings JL, Dekosky ST, Barberger-Gateau P, et al. Revising the definition of Alzheimer's disease: a new lexicon. *Lancet Neurology* 2010;**9**(11):1118-27. [PMID: 20934914]

Fage 2013

Fage BA, Seitz DP, Gill SS, Herrmann N, Smailagic N, Chan CCH, et al. Mini-Cog for the diagnosis of Alzheimer's disease dementia and other dementias within a community setting. *Cochrane Database of Systematic Reviews* 2013, Issue 11. Art. No: CD010860. [DOI: 10.1002/14651858.CD010860] [CD010860]

Farina 2012

Farina N, Isaac MGEKN, Clark Annalie R, Rusted J, Tabet N. Vitamin E for Alzheimer's dementia and mild cognitive impairment. *Cochrane Database of Systematic Reviews* 2012, Issue 11. Art. No: CD002854. [DOI: 10.1002/14651858.CD002854.pub3] [CD002854]

Feldman 2008

Feldman HH, Jacova C, Robillard A, Garcia A, Chow T, Borrie M, et al. Diagnosis and treatment of dementia: 2. Diagnosis. *Canadian Medical Association Journal* 2008;**178**(7):825-36. [PMID: 18362376]

Ferri 2005

Ferri CP, Prince M, Brayne C, Brodaty H, Fratiglioni L, Ganguli M, et al. Global prevalence of dementia: a Delphi consensus study. *Lancet* 2005;**366**(9503):2112-7. [PMID: 16360788]

Folstein 1975

Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for

the clinician. *Journal of Psychiatric Research* 1975;**12**(3):189-98. [PMID: 1202204]

Glanville 2012

Glanville J, Cikalo M, Crawford F, Dozier M, McIntosh H. Handsearching did not yield additional unique FDG-PET diagnostic test accuracy studies compared with electronic searches: a preliminary investigation. *Research Synthesis Methods* 2012;**3**(3):202-13. [DOI: DOI: 10.1002/jrsm.1046]

Goetz 2008

Goetz CG, Emre M, Dubois B. Parkinson's disease dementia: definitions, guidelines, and research perspectives in diagnosis. *Annals of Neurology* 2008;**64 Suppl 2**:S81-92. [PMID: 19127578]

Harbord 2007

Harbord RM, Deeks JJ, Egger M, Whiting P, Sterne JA. A unification of models for meta-analysis of diagnostic accuracy studies. *Biostatistics* 2007;**8**(2):239-51. [PMID: 16698768]

Hort 2010

Hort J, O'Brien JT, Gainotti G, Pirttila T, Popescu BO, Rektorova I, et al. EFNS guidelines for the diagnosis and management of Alzheimer's disease. *European Journal of Neurology* 2010;**17**(10):1236-48. [PMID: 20831773]

Jack 2011

Jack CR Jr, Albert MS, Knopman DS, McKhann GM, Sperling RA, Carrillo MC, et al. Introduction to the recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's & Dementia* 2011;**7**(3):257-62. [PMID: 21514247]

Kokkinou 2014

Kokkinou M, Smailagic N, Noel-Storr AH, Hyde C, Ukoumunne O, Worrall RE, et al. Plasma and cerebrospinal fluid (CSF) Abeta42 for the differential diagnosis of Alzheimer's disease dementia in participants diagnosed with any dementia subtype in a specialist care setting. *Cochrane Database of Systematic Reviews* 2014, Issue 1. Art. No: CD010945. [DOI: 10.1002/14651858.CD010945]

Kulisevsky 2009

Kulisevsky J, Pagonabarraga J. Cognitive impairment in Parkinson's disease: tools for diagnosis and assessment. *Movement Disorders* 2009;**24**(8):1103-10. [PMID: 19353727]

Leeflang 2008

Leeflang MM, Moons KG, Reitsma JB, Zwinderman AH. Bias in sensitivity and specificity caused by data-driven selection of optimal cutoff values: mechanisms, magnitude, and solutions. *Clinical Chemistry* 2008;**54**(4):729-37. [PMID: 18258670]

Lund and Manchester Groups 1994

The Lund and Manchester Groups. Clinical and neuropathological criteria for frontotemporal dementia. *Journal of Neurology, Neurosurgery and Psychiatry* 1994;**57**:416-8. [PMID: 8163988]

Matthews 2008

Matthews FE, Stephan BC, McKeith IG, Bond J, Brayne C. Two-year progression from mild cognitive impairment to dementia: to what extent do different definitions agree? *Journal of the American Geriatrics Society* 2008;**56**(8):1424-33. [PMID: 18662209]

Matthews 2009

Matthews FE, Brayne C, Lowe J, McKeith I, Wharton SB, Ince P. Epidemiological pathology of dementia: attributable-risks at death in the Medical Research Council Cognitive Function and Ageing Study. *PLoS Medicine* 2009;**6**(11):e1000180. [PMID: 19901977]

McKeith 1996

McKeith IG, Galasko D, Kosaka K, Perry EK, Dickson DW, Hansen LA, et al. Consensus guidelines for the clinical and pathologic diagnosis of dementia with Lewy bodies (DLB): report of the consortium on DLB international workshop. *Neurology* 1996;**47**(5):1113-24. [PMID: 8909416]

McKeith 2005

McKeith IG, Dickson DW, Lowe J, Emre M, O'Brien JT, Feldman H, et al. Diagnosis and management of dementia with Lewy bodies: third report of the DLB Consortium. *Neurology* 2005;**65**(12):1863-72. [PMID: 16237129]

McKhann 1984

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):939-44. [PMID: 6610841]

McKhann 2011

McKhann GM, Knopman DS, Chertkow H, Hyman BT, Jack CR Jr, Kawas CH, 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 & Dementia* 2011;**7**(3):263-9. [PMID: 21514250]

McShane 2006

McShane R, Areosa Sastre A, Minakaran N. Memantine for dementia. *Cochrane Database of Systematic Reviews* 2006, Issue 2. Art. No: CD003154. [DOI: 10.1002/14651858.CD003154.pub5] [PMID: 16625572]

Mitchell 2008

Mitchell AJ, Shiri-Feshki M. Temporal trends in the long term risk of progression of mild cognitive impairment: a pooled analysis. *Journal of Neurology, Neurosurgery and Psychiatry* 2008;**79**(12):1386-91.

Mitchell 2009

Mitchell AJ. A meta-analysis of the accuracy of the Mini-Mental State Examination in the detection of dementia and mild cognitive impairment. *Journal of Psychiatric Research* 2009;**43**(4):411-31. [PMID: 18579155]

Morris 1993

Morris JC. The Clinical Dementia Rating (CDR): current version and scoring rules. *Neurology* 1993;**43**(11):2412-4. [PMID: 8232972]

Moyer 2014

Moyer VA. Screening for cognitive impairment in older adults: U.S. Preventive Services Task Force recommendation statement. *Annals of Internal Medicine* 2014;**160**(11):791-7. [PMID: 24663815]

MRC CFAS 2001

Neuropathology Group of the Medical Research Council Cognitive Function and Ageing Study (MRC CFAS). Pathological correlates of late-onset dementia in a multicentre, community-based population in England and Wales. *Lancet* 2001;**357**(0140-6736 (Print)):169-75. [PMID: 11213093]

Nieuwenhuis-Mark 2010

Nieuwenhuis-Mark RE. The death knoll for the MMSE: has it outlived its purpose? *Journal of Geriatric Psychiatry and Neurology* 2010;**23**(3):151-7. [PMID: 20231732]

Petersen 1999

Petersen RC, Smith GE, Waring SC, Ivnik RJ, Tangalos EG, Kokmen E. Mild cognitive impairment: clinical characterization and outcome. *Archives of Neurology* 1999;**56**(3):303-8. [PMID: 10190820]

Petersen 2001

Petersen RC, Stevens JC, Ganguli M, Tangalos EG, Cummings JL, DeKosky ST. Practice parameter: early detection of dementia: mild cognitive impairment (an evidence-based review). Report of the Quality Standards Subcommittee of the American Academy of Neurology. *Neurology* 2001;**56**(9):1133-42.

Petersen 2004

Petersen RC. Mild cognitive impairment as a diagnostic entity. *Journal of Internal Medicine* 2004;**256**(3):183-94. [PMID: 15324362]

Petersen 2011

Petersen RC. Clinical practice: mild cognitive impairment. *The New England Journal of Medicine* 2011;**364**(23):2227-34.

Plassman 2008

Plassman BL, Langa KM, Fisher GG, Heeringa SG, Weir DR, Ofstedal MB, et al. Prevalence of cognitive impairment without dementia in the United States. *Annals of Internal Medicine* 2008;**148**(6):427-34.

Prince 2013

Prince M, Bryce R, Albanese E, Wimo A, Ribeiro W, Ferri CP. The global prevalence of dementia: a systematic review and metaanalysis. *Alzheimer's & Dementia* 2013;**9**(1):63-75.e2. [PMID: 23305823]

Reitsma 2005

Reitsma JB, Glas AS, Rutjes AW, Scholten RJ, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *Journal of Clinical Epidemiology* 2005;**58**(10):982-90. [PMID: 16168343]

Ritchie 2014

Ritchie C, Smailagic N, Noel-Storr Anna H, Takwoingi Y, Flicker L, Mason Sam E, et al. Plasma and cerebrospinal fluid amyloid beta for the diagnosis of Alzheimer's disease dementia and other dementias in people with mild cognitive impairment (MCI). *Cochrane Database of Systematic Reviews* 2014, Issue 6. Art. No: CD008782. [DOI: 10.1002/14651858.CD008782.pub4]

Roman 1993

Roman GC, Tatemichi TK, Erkinjuntti T, Cummings JL, Masdeu JC, Garcia JH, et al. Vascular dementia: diagnostic criteria for research studies. Report of the NINDS-AIREN International Workshop. *Neurology* 1993;**43**(2):250-60. [PMID: 8094895]

Russ 2012

Russ TC, Morling JR. Cholinesterase inhibitors for mild cognitive impairment. *Cochrane Database of Systematic Reviews* 2012, Issue 9. Art. No: CD009132. [DOI: 10.1002/14651858.CD009132.pub2] [CD009132]

Savva 2009

Savva GM, Wharton SB, Ince PG, Forster G, Matthews FE, Brayne C. Age, neuropathology, and dementia. *The New England Journal of Medicine* 2009;**360**(22):2302-9. [PMID: 19474427]

Scheltens 2011

Scheltens P, Rockwood K. How golden is the gold standard of neuropathology in dementia? *Alzheimer's & Dementia* 2011;**7**(4):486-9. [PMID: 21784357]

Seitz 2014

Seitz DP, Fage BA, Chan CCH, Gill SS, Herrmann N, Smailagic N, et al. Mini-Cog for the diagnosis of Alzheimer's disease dementia and other dementias within a primary care setting. *Cochrane Database of Systematic Reviews* 2014, Issue 12. Art. No: CD011415. [DOI: 10.1002/14651858.CD011415] [CD011415]

Shiroky 2007

Shiroky JS, Schipper HM, Bergman H, Chertkow H. Can you have dementia with an MMSE score of 30? *American Journal of Alzheimer's Disease and Other Dementias* 2007;**22**(5):406-15. [PMID: 17959876]

Sperling 2011

Sperling RA, Aisen PS, Beckett LA, Bennett DA, Craft S, Fagan AM, et al. Toward defining the preclinical stages of Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimer's & Dementia 2011;**7**(3):280-92. [1552-5279: (Electronic)]

Stephan 2007

Stephan BC, Matthews FE, McKeith IG, Bond J, Brayne C. Early cognitive change in the general population: how do different definitions work? *Journal of the American Geriatrics Society* 2007;**55**(10):1534-40.

Tombaugh 1992

Tombaugh TN, McIntyre NJ. The Mini-Mental State Examination: a comprehensive review. *Journal of the American Geriatrics Society* 1992;**40**(9):922-35. [PMID: 1512391]

Trenkle 2007

Trenkle DL, Shankle WR, Azen SP. Detecting cognitive impairment in primary care: performance assessment of three screening instruments. *Journal of Alzheimer's Disease* 2007;**11**(3):323-35. [PMID: 17851183]

Vacante 2013

Vacante M, Smailagic N, Sachpekidis C, Hyde C, Martin S, Ukoumunne O. The accuracy of 18FDG-PET in the early diagnosis of Alzheimer's disease dementia and other dementias in people with MCI. *Cochrane Database of Systematic Reviews* 2013, Issue 7. Art. No: CD010632. [DOI: 10.1002/14651858.CD010632]

Whiting 2011

Whiting P, Westwood M, Beynon R, Burke M, Sterne JA, Glanville J. Inclusion of methodological filters in searches for diagnostic test accuracy studies misses relevant studies. *Journal of Clinical Epidemiology* 2011;**64**(6):602-7. [PMID: 21075596]

Whiting 2011a

Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Annals of Internal Medicine* 2011;**155**(8):529-36. [PMID: 22007046]

Wimo 2010

Wimo A, Winblad B, Jonsson L. The worldwide societal costs of dementia: estimates for 2009. *Alzheimer's & Dementia* 2010;**6**(2):98-103. [PMID: 20298969]

Winblad 2004

Winblad B, Palmer K, Kivipelto M, Jelic V, Fratiglioni L, Wahlund LO, et al. Mild cognitive impairment—beyond controversies, towards a consensus: report of the International Working Group on Mild Cognitive Impairment. *Journal of Internal Medicine* 2004;**256**(3):240-6. [PMID: 15324367]

Yue 2012

Yue J, Dong BR, Lin X, Yang M, Wu HM, Wu T. Huperzine A for mild cognitive impairment. *Cochrane Database of Systematic Reviews* 2012, Issue 12. Art. No: CD008827. [DOI: 10.1002/14651858.CD008827.pub2] [CD008827]

Zhang 2014

Zhang S, Smailagic N, Hyde C, Noel-Storr AH, Takwoingi Y, McShane R, Feng J. 11C-PIB-PET for the early diagnosis of Alzheimer's disease dementia and other dementias in people with mild cognitive impairment (MCI). *Cochrane Database of Systematic Reviews* 2014, Issue 7. Art. No: CD010386. [DOI: 10.1002/14651858.CD010386.pub2]

* Indicates the major publication for the study

CHARACTERISTICS OF STUDIES

Characteristics of included studies *[ordered by study ID]*

Buchhave 2008

Study characteristics	
Patient Sampling	148 MCI patients who performed the cube copying test were recruited from a memory clinic in a prospective manner. Sampling procedure is not fully described (mostly referred by general practitioners). Exclusion criteria involves: patients with other causes of cognitive impairment, including brain tumour, subdural haematoma, CNS infection, major depressive episode, schizophrenia and current alcohol abuse
Patient characteristics and setting	This study included 148 MCI participants who performed the cube copying test and were diagnosed by the Petersen 1999 criteria at baseline. Physicians specialised in cognitive disorders performed a thorough physical, neurological and psychiatric examination, as well as a clinical interview, of each patient at baseline. Furthermore, analysis of apolipoprotein E (APOE) genotype and computed tomography of the brain were done Gender: 32 M, 30 F in MCI-stable group; 20 M, 43 F in MCI-AD group; 14 M, 8 F in MCI-other dementias group Age: 66.2±8.8 years in MCI-stable; 74.6±6.1 years in MCI-AD; 72.7±9.0 years in MCI-other dementia APOEε 4 carrier: 32 in MCI-stable; 20 in MCI-AD; 14 in MCI-other dementia MMSE at baseline: MCI-stable 27.5±1.8; MCI-AD 26.8±1.4; MCI-other dementia 27.0 ± 1.7

Buchhave 2008 *(Continued)*

	Resources of referrals: "The majority of the subjects with MCI (74.1%) were referred by their general practitioners. The remaining MCI subjects were referred by other physicians (19.0%) or the patients themselves (6.8%)." Page 545 Resources of recruitment: memory disorder clinic, Malmo University Hospital, Sweden
Index tests	Mini-Mental state Examination (MMSE): no details about version, who administered and interpreted the test and scoring systems are provided. No threshold is pre-specified. After analysis, authors estimated threshold in 28/29 score
Target condition and reference standard(s)	Main target condition for conversion: conversion from MCI to Alzheimer's disease dementia, vascular dementia, dementia by Lewy bodies or frontotemporal dementia Reference standard: for ADD = NINCDS-ADRDA + DSM-IIIR criteria; for VaD = DSM-IIIR criteria + NINDS-AIREN; for DLB = consensus criteria by McKeith 1999; for frontotemporal dementia = Brun 1994
Flow and timing	Duration of follow-up: average 5.2 years (range 4 to 6.8 years) At baseline:148 participants with MCI At follow up: 147 participants: 63 MCI-AD; 22 MCI-other dementia (15 VD; 4 LBD; 1 SD; 1 FTD; 1 dementia induced by traumatic brain injury); 62 stable-MCI Conversion to AD N = 85; disease positive = 63; disease negative = 62 (p546); sensitivity = 89%; specificity = 33% (Table 2, p548) TP = 56; FP = 42; FN=7; TN=20 (Calculated in RevMan5) Loss to follow-up: one participants died before 4 years follow-up (≅ 1%)
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 3: Reference Standard			

Buchhave 2008 *(Continued)*

Is the assessment used for clinical diagnosis of dementia acceptable?	Yes		
Was clinical assessment for dementia performed without knowledge of the MMSE results?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Unclear
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between MMSE and the reference standard?	Yes		
Did all participants receive the same reference standard?	Yes		
Were all participants included in the final analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Chopard 2009

Study characteristics	
Patient Sampling	A cohort of 106/362 patients with QD (presence of cognitive complaints+ no daily activities affected + CDR = 0,5 + no DSM-IV criteria for dementia) who attended local memory impairment consultation centers and recruited through the database of the Regional Network for Diagnostic aids and Management of patients with cognitive impairment in the Franche-Comtê geographical area was included Exclusion criteria: craniocerebral trauma, stroke within 3 months from the beginning of the study, acute neurological or somatic pathologies, progressive psychiatric illness except major depression
Patient characteristics and setting	A cohort of 106 participants were followed-up for 6 to 24 months (mean = 14.9 ± 4.5), from an original sample of 362 patients with QD (59% without follow-up) Demographic data reported by total sample and by conversion at follow-up (Tables 1 and 2): Gender: 41.5% M, 58.5% F Age: 75.7 ± 5.0 years APOE 4 carrier: not reported Education more than 12 years = 8.5%

Chopard 2009 *(Continued)*

	MMSE at baseline: 25.2 ± 2.7 Sources of referral: the database of the Regional Network for Diagnostic aids and Management of patients with cognitive impairment in the Franche-Comtê geographical area Sources of recruitment: the database of the Regional Network for Diagnostic aids and Management of patients with cognitive impairment in the Franche-Comtê geographical area
Index tests	Mini-Mental state Examination (MMSE): no details about version and scoring systems are provided. It is unclear if threshold was pre-specified or not (26/27) and who administered and interpreted the test
Target condition and reference standard(s)	Target condition:Conversion from MCI to Dementia and later defined as ADD, VaD, DLB or FTD Reference standard: dementia was defined according with the following criteria: progressive worsening of cognitive function at follow-up severe enough to affect IADLs and progression of CDR score form 0.5 to 1. Standard criteria were used for diagnosis of ADD (NINCDS-ADRDA), vascular dementia (NINDS-AIREN), dementia for LB (McKeith criteria) and frontotemporal dementia (Lund and Manchester criteria). Dementia was defined independently of MMSE scores
Flow and timing	Duration of follow-up: 6 to 24 months (mean 14.9± 4.5 months) At baseline: 106 patients with QD At follow-up: 38 converted to dementia; 68 remained free of dementia TP = 29; FP = 41; FN = 9; TN = 27 Sensitivity: 76%; Specificity: 40% (calculated in RevMan5); cut-off: ≤ 26 (26/27) (page 705) Loss to follow-up: none
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 3: Reference Standard			

Chopard 2009 *(Continued)*

Is the assessment used for clinical diagnosis of dementia acceptable?	Yes		
Was clinical assessment for dementia performed without knowledge of the MMSE results?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between MMSE and the reference standard?	Yes		
Did all participants receive the same reference standard?	Yes		
Were all participants included in the final analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Conde-Sala 2012

Study characteristics	
Patient Sampling	A longitudinal study (nested in a case-control study) including 109 MCI patients was analysed. In addition, no details about recruitment of MCI patients were reported Exclusion criteria: not reported
Patient characteristics and setting	A total of 109 participants with MCI, diagnosed with the Portet criteria (2006), were recruited from the Memory and Dementia Assessment Unit. Demographic data reported for total sample (N = 342) (Page 950). There were no important differences between the study groups in terms of age, gender, marital status or schooling Gender: 235 (69%) F; 107 (31%) M Age: 74.3 ± 6.2 years APOE 4 carrier: not reported MMSE at baseline: 21.7 ± 3.3 MCI-converters; 21.0 ± 2.9 MCI-stable Education: 21% illiterate or no formal schooling; 18% one to five, 50% six to eight and 11% more than eight years of education respectively.

Conde-Sala 2012 *(Continued)*

	Sources of referral: not reported Sources of recruitment: outpatient Memory and Dementia Assessment Unit, Santa Catarina Hospital, Girona, Spain
Index tests	Mini-Mental state Examination (MMSE): basic details provided. The test was administered by a team of neuropsychologist from the hospital unit. No details about who interpreted the test. Used threshold was pre-specified (21/22)
Target condition and reference standard(s)	Target condition: conversion from MCI to Alzheimer's disease dementia Reference standards: NINCDS-ADRDA criteria for probable AD and DSM-IV-TR criteria for AD
Flow and timing	Duration of follow-up: 5 years At baseline:109 participants with MCI At follow up: 43 MCI-AD converters; 66 MCI stable (non-converters) Number included in analyses: 109 Conversion to AD Information provided from the author (unpublished): TP = 19; FP = 36; FN = 24; TN = 30 Sensitivity: 44.19%; Specificity: 45.45% (calculated in RevMan5); cut-off: ≤ 21 (21/22) (p949) Loss to follow-up: none
Comparative	
Notes	The trial investigators was contacted; they provided requested data tor the 2 X 2 table to be completed; e-mail from Dr Conde-Sala on 20 February 2014

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 3: Reference Standard			

Conde-Sala 2012 *(Continued)*

Is the assessment used for clinical diagnosis of dementia acceptable?	Yes		
Was clinical assessment for dementia performed without knowledge of the MMSE results?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between MMSE and the reference standard?	Yes		
Did all participants receive the same reference standard?	Yes		
Were all participants included in the final analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Devanand 2008

Study characteristics	
Patient Sampling	A total of 148 outpatients with memory complaints (MCI broadly defined) from a Memory Disorders Clinic were consecutive recruited Exclusion criteria: not reported
Patient characteristics and setting	A total of 148 patients were included in analysis. Sample size estimation = 150 patients; two patients were excluded by misdiagnosis. Baseline MCI was determined post hoc due to the study began before MCI criteria were published Demographic data presented for 39 MCi-converters and 109 MCI non-converters was provided on Table 1 (Page 14): Gender: Converters: 56.4% F; Non-converters: 55.1% F Age: Converters: 73.2± 7.1 years; Non-converters: 64.9 ± 9.9 years Education (years): Converters: 14 ± 4.7 years; Non-converters: 15.4 ± 4.1 years APOE ε4 carrier: Converters: 34.3%; Non-converters: 23.6% MMSE at baseline: Converters: 26.3± 2.2; Non-converters: 27.9 ± 2.0 Sources of referral: the majority (52%) were physician referred, 25% were self-referred, and 23% were referred by family or friends or other sources (Devanand 2007)

Devanand 2008 *(Continued)*

	Sources of recruitment: Memory Disorders Clinic (jointly run by NY Psychiatric Institute, Columbia University)
Index tests	Mini-Mental state Examination (MMSE): no details about version and scoring systems are provided. It is unclear who administered the baseline test. Two expert raters reviewed the neuropsychological information to determine inclusion No threshold was used. Quote: "Based on the fitted logistic regression and dichotomizing estimated risk at 0.5, likelihood ratios with 95% CI were calculated"
Target condition and reference standard(s)	Target condition: Conversion from MCI to Alzheimer`s disease dementia Reference standard: DSM- IV criteria + NINCDS-ADRDA criteria. The endpoint of conversion to AD required diagnosis at two consecutive annual outcome criteria Two expert raters (DPD and YS) made a consensus diagnosis at each follow-up, while remaining blind to data from previous visits (page 3)
Flow and timing	At baseline = 148 patients were included During 1 to 9 years of follow-up: 39 converters, 109 non-converters, 63 healthy controls For analysis of sensitivity and specificity, authors selected a sub-sample of 125 patients with 3-years of follow-up: At 3-years of follow-up: 33 converters, 92 non-converters TP = 9; FN = 9; FP = 24; TN = 83 Sensitivity = 26.8%; Specificity = 90% (calculated in RevMan5) Loss to follow up: 22 patients dropped out before the 3-year follow up time-point
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 3: Reference Standard			

Devanand 2008 *(Continued)*

Is the assessment used for clinical diagnosis of dementia acceptable?	Yes		
Was clinical assessment for dementia performed without knowledge of the MMSE results?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between MMSE and the reference standard?	Yes		
Did all participants receive the same reference standard?	Yes		
Were all participants included in the final analysis?	No		
Could the patient flow have introduced bias?		High risk	

Meguro 2007a

Study characteristics	
Patient Sampling	This reference have two incidence studies (called Meguro 2007a and Meguro 2007b for this review). In the 2003 study (Meguro 2007a) 54 participants with CDR = 0.5 with MRI results were analysed. Whole sample included participants with CDR = 0 (204 patients) Exclusion criteria: not reported
Patient characteristics and setting	A total of 54 participants with MCI ('questionable dementia') diagnosed with CDR = 0.5 (Morris 1993) were included in study 2003. Those participants were recruited from the same community based population. Demographic data presented for the MCI-converters (n = 133; Table 1, p178) was provided for both studies (Meguro 2007a and Meguro 2007b): Gender: 45 M; 88 F Age: 9 aged 65 to 69 years; 77 aged 70 to 79 years; 47 aged 80^{+} Sources of recruitment: community of Tajiri, Japan
Index tests	Mini-Mental state Examination (MMSE): Folstein version (1975, reference 25). A team of trained psychologist performed MMSE blindly to the diagnosis and CDR Threshold: “Abnormal MMSE data are assessed with reference to the participants’ educational levels based on the old Japanese education system: 17/18 for 6 (or less) years of schooling,

Meguro 2007a *(Continued)*

	20/21 for 8 years of schooling, and 23/24 for 10 (or more) years of schooling" (information provided from the author) Cut off: < 17 or < 20 or < 23
Target condition and reference standard(s)	Target condition: conversion from MCI to Alzheimer's disease dementia, VaD, DLB or FTD Reference standard: NINCDS-ADRDA for probable AD; NINDS-AIREN for possible AD with CVD; NINDS-AIREN for probable VaD; the consensus guidelines for diagnosis of dementia with Lewy bodies (DLB) (McKeith 2005), and the Lund and Manchester Groups criteria for frontotemporal dementia (McKhann 2001). A clinical team comprising medical doctors and public health nurses determined the follow-up CDR blindly to the previous CDR stages, baseline cognitive test scores, and MRI data
Flow and timing	Participants 2003 at baseline: N = 54 Duration of follow-up: 5 years At follow-up: 20 MCI-dementia: 7 with abnormal MMSE (at baseline); 34 MCI-stable: 2 with abnormal MMSE (at baseline) (information from the author) TP=7; FN=13; FP=2; TN=32 Sensitivity=35%; Specificity=94% (Calculated in RevMan5)
Comparative	
Notes	The trial investigators were contacted; they provided requested data tor the 2 X 2 table to be completed; e-mail from Dr Meguro on 4 March 2014

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the assessment used for clinical diagnosis of dementia acceptable?	Yes		
Was clinical assessment for dementia performed without knowledge of the MMSE results?	Yes		

Meguro 2007a *(Continued)*

Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between MMSE and the reference standard?	Yes		
Did all participants receive the same reference standard?	Yes		
Were all participants included in the final analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Meguro 2007b

Study characteristics	
Patient Sampling	This reference have two incidence studies (called Meguro 2007a and Meguro 2007b for this review). In the 2005 study (Meguro 2007b), 281 patients without MRI results were analysed Exclusion criteria: not reported
Patient characteristics and setting	A total of 281 participants with MCI ('questionable dementia') diagnosed with CDR = 0.5 (Morris 1993) were included in 2005 study. Those participants were recruited from the same community based population. Demographic data presented for the MCI-converters (n=133; Table 1, p178) was provided for both studies (Meguro 2007a and Meguro 2007b): Gender: 45 M; 88 F Age: 9 aged 65 to 69 years; 77 aged 70 to 79 years; 47 aged 80^{+} Sources of recruitment: community of Tajiri, Japan
Index tests	Mini-Mental state Examination (MMSE): Folstein version (1975, reference 25). A team of trained psychologist performed MMSE blindly to the diagnosis and CDR Threshold: “Abnormal MMSE data are assessed with reference to the participants’ educational levels based on the old Japanese education system: 17/18 for 6 (or less) years of schooling, 20/21 for 8 years of schooling, and 23/24 for 10 (or more) years of schooling” (information provided from the author) Cut off: < 17 or < 20 or < 23
Target condition and reference standard(s)	Target condition: conversion from MCI to dementia

Meguro 2007b *(Continued)*

	Reference standard: DSM-IV and CDR 1 +. A clinical team comprising medical doctors and public health nurses determined the follow-up CDR blindly to the previous CDR stages, baseline cognitive test scores, and MRI data
Flow and timing	Participants 2005 (at baseline): N = 281 Duration of follow-up: 7 years At follow-up: 113 MCI-dementia: 26 with abnormal MMSE (at baseline); 168 MCI-stable: 14 with abnormal MMSE (at baseline) (information from the author) TP = 26; FN = 87; FP = 14; TN = 154 Sensitivity = 23%; Specificity = 92% (calculated in RevMan5) Loss to follow-up: none
Comparative	
Notes	The trial investigators were contacted; they provided requested data tor the 2 X 2 table to be completed; e-mail from Dr Meguro on 4 March 2014

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the assessment used for clinical diagnosis of dementia acceptable?	Yes		
Was clinical assessment for dementia performed without knowledge of the MMSE results?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference			Low concern

Meguro 2007b *(Continued)*

standard does not match the question?		
DOMAIN 4: Flow and Timing		
Was there an appropriate interval between MMSE and the reference standard?	Yes	
Did all participants receive the same reference standard?	Yes	
Were all participants included in the final analysis?	Yes	
Could the patient flow have introduced bias?		Low risk

Modrego 2005

Study characteristics	
Patient Sampling	Cohort of 59 consecutive amnestic MCI patients referred from the community by family physicians for recent memory complaints Exclusion criteria: not detailed
Patient characteristics and setting	A total of 59 consecutive patients with memory complaints and criteria for MCI were recruited from a outpatient clinic. The authors asked the general practitioners to refer to them all elderly patients with memory complaints and then they included all who fulfilled the criteria for amnestic mild cognitive impairment (Petersen 1999). The final diagnosis was made by taking into account as much information as possible from all sources. Baseline demographic data are reported inside the text (Page 670) and in the table 1: Gender: 34 F, 21 M Age: 72.7 years ± 5.3 APOE 4 carrier: Not reported. MMSE at baseline: 27.6 ± 3.6 Sources of referral: primary practitioners Sources of recruitment: Outpatient clinic at Spain
Index tests	Mini-Mental state Examination (MMSE): Spanish version which has a maximum of 35 points and a cut-off point of 23 for elderly subjects. No details about who administered and interpreted the test. Used threshold was not pre-specified (at 1 year, 26 points or less; at 3 years, 29 points or less)
Target condition and reference standard(s)	Main target condition: conversion from MCI to dementia and ADD (unclear information) Reference standard: NINCDS-ADRDA criteria
Flow and timing	Duration of follow-up: 1 year and 3 years

Modrego 2005 *(Continued)*

	In 1 year: Patients at baseline: 53/55 patients; one patient was lost and another non-demented patient had died from cardio-embolic stroke 13 patients (53%) converted to AD MMSE cut-off ≤ 26: TP = 8; FP = 7; FN = 5; TN = 33 (calculated in Revman5) In 3 years: Patients at baseline: 53/53 patients; no losses to follow-up 29 patients (55%) converted to AD MMSE cut-off ≤ 29: TP = 22; FP = 8; FN = 7; TN = 16 (calculated in Revman5)
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Unclear
DOMAIN 3: Reference Standard			
Is the assessment used for clinical diagnosis of dementia acceptable?	Unclear		
Was clinical assessment for dementia performed without knowledge of the MMSE results?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Unclear
DOMAIN 4: Flow and Timing			

Modrego 2005 *(Continued)*

Was there an appropriate interval between MMSE and the reference standard?	Yes	
Did all participants receive the same reference standard?	Unclear	
Were all participants included in the final analysis?	Yes	
Could the patient flow have introduced bias?		Unclear risk

Modrego 2013

Study characteristics	
Patient Sampling	Cohort of 105 patients with subjective memory complaints and referred by family physicians were assessed between December 2007 to April 2011 in an outpatient clinic of a university hospital in Zaragoza, Spain Exclusion criteria: patients with hydrocephalus, chronic psychiatric conditions, large infarcts of the brain or those who met the NINCDS-ADRDA criteria for dementia
Patient characteristics and setting	A total of 105 patients with subjective memory complaints corroborated by an informant (relative or caregiver), preservation of daily living activities, no behavioral symptoms (BDRS < 4), and normality in global cognitive function (MEC > 23 or > 26 by age; CDR 0,5) were included Demographic data reported for 105 participants (Table 1, p7) Age: converters 76 ± 9 years; non-converters: 72.4 ± 6.6 years APOE4 carrier: not provided MMSE: converters 27.6 ± 2.7; non-converters 29.8 ± 2.8 Education: University: converters = 3.5%; non-converters = 8.3% Sources of referral: family physicians Sources of recruitment: unclear
Index tests	Mini-Mental state Examination (MMSE): Spanish version which has a maximum of 35 points. No details about who administered and interpreted the test at baseline. Used threshold was no pre-specified
Target condition and reference standard(s)	Main target condition: conversion from MCI to probable AD Reference standard: NINCDS-ADRDA criteria
Flow and timing	Duration of follow-up: median of 2 years (range 1 to 4 years) Patients at baseline: 105/110 patients; 5 patients were excluded at baseline according with exclusion criteria 57 patients (54.2%) converted to AD MMSE cut-off ≤ 30/35: TP = 26; FP = 6; FN = 31; TN = 42 (calculated in Revman5)

Modrego 2013 *(Continued)*

Comparative
Notes

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the assessment used for clinical diagnosis of dementia acceptable?	Yes		
Was clinical assessment for dementia performed without knowledge of the MMSE results?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Unclear
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between MMSE and the reference standard?	Yes		
Did all participants receive the same reference standard?	Yes		
Were all participants included in the final analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Palmqvist 2012

Study characteristics	
Patient Sampling	A MCI cohort of 133 MCI patients referred to the clinic between 2000 to 2006 were assessed. Most patients were referred from primary care units, but some referrals came from other clinics at the hospital. No further details about recruitment were reported Exclusion criteria: patients with diagnosis of haematoma, brain tumour, CNS infection, schizophrenia, major depressive episode or current alcohol abuse were not included
Patient characteristics and setting	A total of 133 participants with MCI, diagnosed with the Petersen 2004 criteria, were recruited in the memory clinic. Baseline demographic data reported in table 1 (p4) Gender: MCI-MCI: 34 F, 28 M; MCI-AD: 36 F, 16 M; MCI-other dementias: 8 F, 11 M Age: MCI-MCI: 69.8 years (55 to 85); MCI-AD: 75.3 years (55 to 87); MCI-other dementias: 71.2 years (59 to 83) APOE4 carrier: MCI-MCI: 28; MCI-AD: 39; MCI-other dementias: 12 MMSE: MCI-MCI: mean 28.1 ± 1.2; MCI-AD: mean 26.1 ± 1.5; MCI-other dementias: mean 27.1 ± 2.0 MMSE (O & R): MCI-MCI: mean 11.4 ± 1.1; MCI-AD: mean 9.6 ± 1.4; MCI-other dementias: mean 10.9 ± 1.3 Education: not reported Sources of referral: most patients were referred from primary care units, but some referrals came from other clinics at the hospital Sources of recruitment: memory clinic of Skane University Hospital, Malmö, Sweden
Index tests	Mini-Mental state Examination (MMSE): Full details provided. The test was administered by physicians experienced in dementia disorders. also, the items about orientation and recall (MMSE O&R) are assessed. Used thresholds were not pre-specified
Target condition and reference standard(s)	Target condition: conversion from MCI to Alzheimer's disease dementia, vascular dementia, dementia with Lewy bodies Reference standard: AD:NINCDS-ADRDA criteria, Vascular dementia (VaD); either probable VaD according to NINDS-AIREN (Roman 1993) or subcortical VaD according to Erkinjuntti 2000, Dementia with Lewy bodies (DLB): according to the McKeith 2005 criteria A consensus group of three study physicians experienced in dementia disorders (OH, JH and LM) later determined all diagnoses. The physicians were blinded to the CSF and cognitive test data collected on the initial visit
Flow and timing	Duration of follow-up: mean 5.9 years (range 3.2 to 8.8 years) At baseline: 133 MCI At follow up: 71 MCI converters (52 MCI-AD; 10 MCI-VD; 4 MCI-DLB; 3 MCI-PSP; 1 MCI-SD; 1 MCI-dementia due to brain tumour); 62 MCI stable Table 2: ROC curve analysis: predicting follow-up AD diagnosis 1) MMSE (cut-off < 27): sensitivity = 62%; specificity = 84% Conversion to AD dementia: N = 133: 52 AD ('disease positive') and 81 'disease negative' (non-AD) TP = 32; FP = 13; FN = 20; TN = 68 (calculated in Revman5) 2) MMSE (O & R) (cut-off < 10): sensitivity = 54%; specificity = 94%

Palmqvist 2012 *(Continued)*

	Conversion to AD dementia: N = 133: 52 AD ('disease positive') and 81 'disease negative' (non-AD) TP = 28; FP = 5; FN = 24; TN = 78 (calculated in RevMan 5)
Comparative	
Notes	The trial investigators contacted; they confirmed that there was no overlapping between samples in two papers, Palmquist 2012 and Buchhave 2008. E-mail from Dr Palmqvist on 28 February 2014

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the assessment used for clinical diagnosis of dementia acceptable?	Yes		
Was clinical assessment for dementia performed without knowledge of the MMSE results?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			

Palmqvist 2012 *(Continued)*

Was there an appropriate interval between MMSE and the reference standard?	Yes	
Did all participants receive the same reference standard?	Yes	
Were all participants included in the final analysis?	Yes	
Could the patient flow have introduced bias?		Low risk

Pozueta 2011

Study characteristics	
Patient Sampling	Consecutive sample of 115 MCI participants was recruited from a memory clinic between 2007 and 2008 Exclusion criteria: participants who met criteria for dementia (DSM-IV), AD (NINCDS-ADRDA), depressive episode (IDC-10), subjects with significant cerebrovascular disease (Hachinski scale score 0.4), and those with any other medical or psychiatric identifiable cause accounting for their complaints
Patient characteristics and setting	A total of 115 participants with MCI, diagnosed initially with the Petersen 2004 criteria were recruited from a memory clinic Demographic data reported for 105 participants (Table 1, p3) Gender: total: 65 F, 40 M; converters: 30 F, 20 M; non-converters: 35 F, 20 M Age: converters 75.94 ± 6.05 years; non-converters: 72.93 ± 7.3 years APOE4 carrier: converters 27; non-converters 16 MMSE: converters 25.92 ± 1.88; non-converters 27.78 ± 1.55 Education: not reported Sources of referral: not reported Sources of recruitment: memory clinic of the University Hospital
Index tests	Mini-Mental state Examination (MMSE): no details about version, who administered and interpreted the test and scoring systems are provided. Used threshold was not pre-specified
Target condition and reference standard(s)	Target condition: conversion from MCI to Alzheimer's disease dementia Reference standards for identifying the target conditions at follow-up were not reported
Flow and timing	Duration of follow-up: 2 years At follow-up: 50 AD dementia; 2 LBD; 2 VaD, 55 MCI-S (MCI stable, non-converters) Number included in analyses: 105 (excluding 2 patients with LBD; 2 with VD)

Pozueta 2011 *(Continued)*

	Number included in analyses: 105 Conversion to AD Sensitivity: 64%; Specificity: 80%; cut-off: ≤ 26 (Table 2, p3) (disease positive = 59; disease negative = 292) TP = 32; FP = 11; FN = 44; TN = 18 (calculated in RevMan5) Loss to follow-up: 6 participants did not complete the first year of follow-up, so they were not included in analyses
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the assessment used for clinical diagnosis of dementia acceptable?	Unclear		
Was clinical assessment for dementia performed without knowledge of the MMSE results?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Unclear
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between MMSE and the reference standard?	Yes		

Pozueta 2011 *(Continued)*

Did all participants receive the same reference standard?	Unclear	
Were all participants included in the final analysis?	No	
Could the patient flow have introduced bias?		High risk

Xu 2002

Study characteristics	
Patient Sampling	Retrospective cohort of 351 consecutive patients attending outpatient research clinic, admitted between 1992 and 1997 and referred by specialists, primary care physicians or self-referred Exclusion criteria: subjects with pre-existing dementia, non-dementing organic brain disorders, epilepsy, previous strokes and infectious central nerve system diseases
Patient characteristics and setting	351 participants with subjective memory complaints Gender:140 M; 211 F Age: mean age 67 years (SD 11.23) APOE4 carrier: not provided Resources of referral: specialists and primary care physicians; in addition, many relatives, friends and caregivers volunteered to participate (self-referrals) Resources of recruitment: outpatient research clinic, Houston and Southwest United States
Index tests	Mini-Mental state Examination (MMSE): full details about version. Unclear who administered and interpreted the test. Used threshold was not pre-specified
Target condition and reference standard(s)	Target condition: conversion from SMC to Alzheimer's disease dementia, vascular dementia, dementia by Lewy bodies or frontotemporal dementia Reference standard: for AD = NINCDS-ADRDA; VaD = NINDS-AIREN; DLB = McKeith criteria; FTD = Lund and Manchester criteria
Flow and timing	Duration of follow-up: 3 to 6 years (mean 3.89 ± 2.17 years) Number included in analyses: 351 1) Conversion to probable AD: N = 47 probable AD and 304 non-AD (37 non-AD dementias and 267 MCI stable); disease positive = 47; disease negative = 304 Sensitivity: 61%; Specificity: 82.9%; cut-off: ≤ 26 (27/26 scores) (Table 2, p1030) TP = 29; FP = 52; FN = 18; TN = 252 (calculated in RevMan5) 2) Conversion to probable VaD: N = 22 probable VaD and 329 non-VaD (62 non-VaD dementias and 267 MCI stable): disease positive = 22; disease negative = 329 Sensitivity: 36.4%; Specificity: 80%; cut-off: ≤ 25 (26/25 scores) (Table 3, p1030) TP = 8; FP = 65; FN = 14; TN = 264 (calculated in RevMan5)

Xu 2002 *(Continued)*

	3) Conversion to all dementias: N = 84 all dementia and 267 non-converters (267 MCI stable): disease positive = 84; disease negative = 267 Sensitivity: 57.1%; Specificity: 85.8%; cut-off: ≤ 26 (27/26 scores) (Table 4, p1030) TP = 48; FP = 38; FN = 36; TN = 229 (calculated in RevMan5) Loss to follow-up: all participants who completed at least three years of longitudinal cognitive assessment were included in the analysis
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the assessment used for clinical diagnosis of dementia acceptable?	Yes		
Was clinical assessment for dementia performed without knowledge of the MMSE results?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between MMSE and the reference standard?	Yes		

Xu 2002 *(Continued)*

Did all participants receive the same reference standard?	Yes	
Were all participants included in the final analysis?	Yes	
Could the patient flow have introduced bias?		Low risk

Characteristics of excluded studies *[ordered by study ID]*

Study	Reason for exclusion
Aevarsson 2000	Change in MMSE scores from general population (no MCI patients) to dementia are detailed (logistic regression results)
Apostolova 2006	MMSE is assessed as a predictor of conversion to dementia for MCI patients (logistic regression results), but the authors do not provide information about the accuracy of any threshold
Armas 2009	Descriptive information about MMSE is provided without data about its accuracy
Brodaty 2011	Study focused on cognitive decline (based on MMSE); no accuracy information is provided
Bruck 2013	Descriptive information about MMSE is provided without data about its accuracy
Chan 2011	Baseline MMSE is assessed as a predictor of conversion to dementia for MCI patients (logistic regression results), but the authors do not provide information about the accuracy of any threshold
Chilovi 2011	Descriptive information about MMSE is provided without data about its accuracy
Choi 2013	Descriptive information about MMSE is provided without data about its accuracy
Cruz 2012	MMSE is included in a Cox proportional hazard model; no data about accuracy are provided
Devanand 2010	Baseline MMSE scores are evaluated as a part of a model; individual accuracy data are not provided
Devier 2009	MMSE is included in a Cox proportional hazard model; no data about accuracy are provided
Devier 2010	MMSE is included in a Cox proportional hazard model; no data about accuracy are provided
Ehrensperger 2010	MMSE is not evaluated as a factor of conversion (no follow-up)
Ewers 2007	MMSE is included in a Cox proportional hazard model; no data about accuracy are provided
Hampel 2004	MMSE is included in a Cox proportional hazard model; no data about accuracy are provided
Ito 2013	MMSE is included in a logistic regression model; no data about accuracy are provided
Koepsell 2012	MMSE is included in a logistic regression model; no data about accuracy are provided
Korf 2004	MMSE is included in a Cox proportional hazard model; no data about accuracy are provided
Kruczyk 2012	MMSE is included in a rule-based model; no data about accuracy are provided

Study	Reason for exclusion
Li 2011	We made request to authors to obtain useable data, but we did not receive response. At the time to publish this review, we did not had enough information to include this study
Luck 2012	MMSE is included in a Cox proportional hazard model; no data about accuracy are provided
Madureira 2010	MMSE is included in a logistic regression model
Mauri 2012	We made request to authors to obtain useable data, but we did not receive response. At the time to publish this review, we did not had enough information to include this study
Meyer 2002	MMSE is not evaluated as a factor of conversion
Ott 2013	Descriptive information about MMSE is provided without data about its accuracy
Ouchi 2012	MMSE is included in a logistic regression model; no data about accuracy are provided
Paajanen 2014	MMSE is not evaluated as a factor of conversion
Rosenberg 2013	MMSE is included in a Cox proportional hazard model; no data about accuracy are provided
Serrano 2007	MMSE is included in a Cox proportional hazard model; no data about accuracy are provided
Tardif 2013	MMSE is not evaluated as a factor of conversion
Van Rossum 2011	MMSE is included in a Cox proportional hazard model; no data about accuracy are provided
van Rossum 2012	MMSE is included in a Cox proportional hazard model; no data about accuracy are provided
Waldorff 2012	MMSE is included in a Cox proportional hazard model; no data about accuracy are provided
Wong 2013	Participants not individuals with MCI at baseline. MMSE is not evaluated as a factor of conversion
Zhang 2012	MMSE is included in a sparse linear regression model; no data about accuracy are provided

Characteristics of ongoing studies *[ordered by study ID]*

Hall 2012

Study name	Utility of NPI Scores predicting progression of CIND to dementia
Target condition and reference standard(s)	Target condition: dementia Reference standard: unclear
Index and comparator tests	Neuropsychiatric Inventory (NPI)
Starting date	Not provided
Contact information	J Hall; Bl Plassman; D Steffens
Notes	Information from a conference abstract

DATA

Presented below are all the data for all of the tests entered into the review.

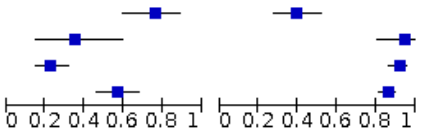
Table Tests. Data tables by test

Test	No. of studies	No. of participants
1 MMSE Conversion to All-cause Dementia	4	792
2 MMSE Conversion to AD dementia	8	1128
3 MMSE Conversion to Vascular Dementia	1	351

Test 1. MMSE Conversion to All-cause Dementia

MMSE Conversion to All-cause Dementia

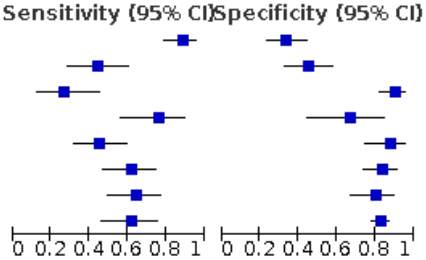
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Chopard 2009	29	41	9	27	0.76 [0.60, 0.89]	0.40 [0.28, 0.52]		
Meguro 2007a	7	2	13	32	0.35 [0.15, 0.59]	0.94 [0.80, 0.99]		
Meguro 2007b	26	14	87	154	0.23 [0.16, 0.32]	0.92 [0.86, 0.95]		
Xu 2002	48	38	36	229	0.57 [0.46, 0.68]	0.86 [0.81, 0.90]		



Test 2. MMSE Conversion to AD dementia

MMSE Conversion to AD dementia

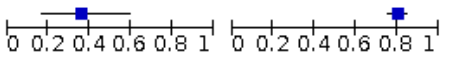
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Buchhave 2008	56	56	7	28	0.89 [0.78, 0.95]	0.33 [0.23, 0.44]		
Conde-Sala 2012	19	36	24	30	0.44 [0.29, 0.60]	0.45 [0.33, 0.58]		
Devanand 2008	9	9	24	83	0.27 [0.13, 0.46]	0.90 [0.82, 0.95]		
Modrego 2005	22	8	7	16	0.76 [0.56, 0.90]	0.67 [0.45, 0.84]		
Modrego 2013	26	6	31	42	0.46 [0.32, 0.59]	0.88 [0.75, 0.95]		
Palmqvist 2012	32	13	20	68	0.62 [0.47, 0.75]	0.84 [0.74, 0.91]		
Pozueta 2011	32	11	18	44	0.64 [0.49, 0.77]	0.80 [0.67, 0.90]		
Xu 2002	29	52	18	252	0.62 [0.46, 0.75]	0.83 [0.78, 0.87]		



Test 3. MMSE Conversion to Vascular Dementia

MMSE Conversion to Vascular Dementia

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Xu 2002	8	65	14	264	0.36 [0.17, 0.59]	0.80 [0.76, 0.84]		



APPENDICES

Appendix 1. Search strategies

Source	Search strategy
1. MEDLINE In-process and other non-indexed citations and MEDLINE 1946-present (OvidSP) up to May 2014	1. MMSE*.ti,ab. 2. sMMSE.ti,ab. 3. Folstein*.ti,ab. 4. MiniMental.ti,ab. 5. "mini mental stat*".ti,ab. 6. or/1-5
2. EMBASE 1980-2014 May 16 (OvidSP)	1. MMSE*.ti,ab. 2. sMMSE.ti,ab. 3. Folstein*.ti,ab. 4. MiniMental.ti,ab. 5. "mini mental stat*".ti,ab. 6. 3MS.ti,ab. 7. *mini mental state examination/ 8. or/1-7 9. dement*.ti,ab. 10. alzheimer*.ti,ab. 11. exp *dementia/ 12. "vascular cognitive impair*".ti,ab. 13. ("lewy bod*" or DLB or LBD).ti,ab. 14. (AD or VaD or FTLD or FTD or DLB or LDB).ti,ab. 15. delirium/ 16. deliri*.ti,ab. 17. or/9-16 18. exp *mild cognitive impairment/ 19. "cognit* impair*".ti,ab. 20. (forgetful* or confused or confusion).ti,ab. 21. MCI.ti,ab. 22. ACMI.ti,ab. 23. ARCD.ti,ab. 24. SMC.ti,ab.

(Continued)

	25. CIND.ti,ab.
	26. BSF.ti,ab.
	27. AAMI.ti,ab.
	28. LCD.ti,ab.
	29. QD.ti,ab.
	30. AACD.ti,ab.
	31. MNCD.ti,ab.
	32. MCD.ti,ab.
	33. (nMCI or aMCI or mMCI).ti,ab.
	34. ("N-MCI" or "A-MCI" or "M-MCI").ti,ab.
	35. "Petersen criteria".ab.
	36. ((CDR adj2 "0.5") or ("clinical dementia rating" adj3 "0.5")).ab.
	37. "cognit* declin*".ti,ab.
	38. "cognit* deficit*".ti,ab.
	39. or/18-38
	40. 17 or 39
	41. 8 and 40
3. PSYCINFO 1806-May week 3 2014 (OvidSP)	1. exp Dementia/
	2. exp Delirium/
	3. exp Huntingtons Disease/
	4. exp Kluver Bucy Syndrome/
	5. exp Wernickes Syndrome/
	6. exp Cognitive Impairment/
	7. dement*.mp.
	8. alzheimer*.mp.
	9. (lewy* adj2 bod*).mp.
	10. deliri*.mp.
	11. (chronic adj2 cerebrovascular).mp.
	12. ("organic brain disease" or "organic brain syndrome").mp.
	13. "supranuclear palsy".mp.
	14. ("normal pressure hydrocephalus" and "shunt*").mp.
	15. "benign senescent forgetfulness".mp.
	16. (cerebr* adj2 deteriorat*).mp.
	17. (cerebral* adj2 insufficient*).mp.

(Continued)

18. (pick* adj2 disease).mp.

19. (creutzfeldt or jcd or cjd).mp.

20. huntington*.mp.

21. binswanger*.mp.

22. korsako*.mp.

23. ("parkinson* disease dementia" or PDD or "parkinson* dementia").mp.

24. or/1-23

25. "cognit* impair*".mp.

26. exp Cognitive Impairment/

27. MCI.ti,ab.

28. ACMI.ti,ab.

29. ARCD.ti,ab.

30. SMC.ti,ab.

31. CIND.ti,ab.

32. BSF.ti,ab.

33. AAMI.ti,ab.

34. MD.ti,ab.

35. LCD.ti,ab.

36. QD.ti,ab.

37. AACD.ti,ab.

38. MNCD.ti,ab.

39. MCD.ti,ab.

40. ("N-MCI" or "A-MCI" or "M-MCI").ti,ab.

41. ((cognit* or memory or cerebr* or mental*) adj3 (declin* or impair* or los* or deteriorat* or degenerat* or complain* or disturb* or disorder*)).ti,ab.

42. "preclinical AD".mp.

43. "pre-clinical AD".mp.

44. ("preclinical alzheimer*" or "pre-clinical alzheimer*").mp.

45. (aMCI or MCIa).ti,ab.

46. ("CDR 0.5" or "clinical dementia rating scale 0.5").ti,ab.

47. ("GDS 3" or "stage 3 GDS").ti,ab.

48. ("global deterioration scale" and "stage 3").mp.

49. "Benign senescent forgetfulness".ti,ab.

50. "mild neurocognit* disorder*".ti,ab.

51. (prodrom* adj2 dement*).ti,ab.

(Continued)

	52. "age-related symptom*".mp. 53. (episodic adj2 memory).mp. 54. ("pre-clinical dementia" or "preclinical dementia").mp. 55. or/25-54 56. 24 or 55 57. mini mental state examination/ 58. "mini mental stat*".ti,ab. 59. MiniMental.ti,ab. 60. Folstein*.ti,ab. 61. sMMSE.ti,ab. 62. MMSE*.ti,ab. 63. or/57-62 64. 56 and 63
4. Biosis previews 1926 to present (ISI Web of Science) to 20 May 2014	Topic=(MMSE OR sMMSE OR "mini mental stat*" OR folstein* OR MiniMental) AND Year Published=(1975-2012) AND Topic=(detect* OR diagnos* OR predict* OR identify OR validity OR validation OR validate OR utility OR sensitivity OR specificity OR screen* OR preval* OR incidence) AND Topic=(dement* OR alzheimer* OR cognitive OR cognition OR memory OR MCI OR petersen) Databases=BIOSIS Previews. Lemmatization=On
5. Web of Science and conference proceedings (1945-present) to 20 May 2014	Topic=(MMSE OR sMMSE OR "mini mental stat*" OR folstein* OR MiniMental) AND Year Published=(1975-2012) AND Topic=(detect* OR diagnos* OR predict* OR identify OR validity OR validation OR validate OR utility OR sensitivity OR specificity OR screen* OR preval* OR incidence) AND Topic=(dement* OR alzheimer* OR cognitive OR cognition OR memory OR MCI OR petersen) Databases=SCI-EXPANDED, SSCI, CPCI-S, CPCI-SSH. Lemmatization=On
6. LILACS (BIREME) to 20 May 2014	MMSE OR flostein OR "mini mental stat$" OR sMMSE OR MiniMental [Words]
7. ALOIS (CDCIG Specialized Register) to 20 May 2014 [Search strategy run in Medline (OvidSP) used to populate ALOIS (Cochrane Dementia and Cognitive Improvement specialized register)]	1. "word recall".ti,ab. 2. ("7-minute screen" OR “seven-minute screen”).ti,ab. 3. ("6 item cognitive impairment test" OR “six-item cognitive impairment test”).ti,ab. 4. "6 CIT".ti,ab. 5. "AB cognitive screen".ti,ab. 6. "abbreviated mental test".ti,ab. 7. "ADAS-cog".ti,ab. 8. AD8.ti,ab. 9. "inform* interview".ti,ab. 10. "animal fluency test".ti,ab.

(Continued)

11. "brief alzheimer* screen".ti,ab.
12. "brief cognitive scale".ti,ab.
13. "clinical dementia rating scale".ti,ab.
14. "clinical dementia test".ti,ab.
15. "community screening interview for dementia".ti,ab.
16. "cognitive abilities screening instrument".ti,ab.
17. "cognitive assessment screening test".ti,ab.
18. "cognitive capacity screening examination".ti,ab.
19. "clock drawing test".ti,ab.
20. "deterioration cognitive observee".ti,ab.
21. ("Dem Tect" OR DemTect).ti,ab.
22. "object memory evaluation".ti,ab.
23. "IQCODE".ti,ab.
24. "mattis dementia rating scale".ti,ab.
25. "memory impairment screen".ti,ab.
26. "minnesota cognitive acuity screen".ti,ab.
27. "mini-cog".ti,ab.
28. "mini-mental state exam*".ti,ab.
29. "mmse".ti,ab.
30. "modified mini-mental state exam".ti,ab.
31. "3MS".ti,ab.
32. “neurobehavio?ral cognitive status exam*”.ti,ab.
33. "cognistat".ti,ab.
34. "quick cognitive screening test".ti,ab.
35. "QCST".ti,ab.
36. "rapid dementia screening test".ti,ab.
37. "RDST".ti,ab.
38. "repeatable battery for the assessment of neuropsychological status".ti,ab.
39. "RBANS".ti,ab.
40. "rowland universal dementia assessment scale".ti,ab.
41. "rudas".ti,ab.
42. "self-administered gerocognitive exam*".ti,ab.
43. ("self-administered" and "SAGE").ti,ab
44. "self-administered computerized screening test for dementia".ti,ab.

(Continued)

45. "short and sweet screening instrument".ti,ab.

46. "sassi".ti,ab.

47. "short cognitive performance test".ti,ab.

48. "syndrome kurztest".ti,ab.

49. ("six item screener" OR “6-item screener”).ti,ab.

50. "short memory questionnaire".ti,ab.

51. ("short memory questionnaire" and "SMQ").ti,ab.

52. "short orientation memory concentration test".ti,ab.

53. "s-omc".ti,ab.

54. "short blessed test".ti,ab.

55. "short portable mental status questionnaire".ti,ab.

56. "spmsq".ti,ab.

57. "short test of mental status".ti,ab.

58. "telephone interview of cognitive status modified".ti,ab.

59. "tics-m".ti,ab.

60. "trail making test".ti,ab.

61. "verbal fluency categories".ti,ab.

62. "WORLD test".ti,ab.

63. "general practitioner assessment of cognition".ti,ab.

64. "GPCOG".ti,ab.

65. "Hopkins verbal learning test".ti,ab.

66. "HVLT".ti,ab.

67. "time and change test".ti,ab.

68. "modified world test".ti,ab.

69. "symptoms of dementia screener".ti,ab.

70. "dementia questionnaire".ti,ab.

71. "7MS".ti,ab.

72. ("concord informant dementia scale" or CIDS).ti,ab.

73. (SAPH or "dementia screening and perceived harm*").ti,ab.

74. or/1-73

75. exp Dementia/

76. Delirium, Dementia, Amnestic, Cognitive Disorders/

77. dement*.ti,ab.

78. alzheimer*.ti,ab.

(Continued)

79. AD.ti,ab.

80. ("lewy bod*" or DLB or LBD or FTD or FTLD or “frontotemporal lobar degeneration” or “frontal-temporal dement*).ti,ab.

81. "cognit* impair*".ti,ab.

82. (cognit* adj4 (disorder* or declin* or fail* or function* or degenerat* or deteriorat*)).ti,ab.

83. (memory adj3 (complain* or declin* or function* or disorder*)).ti,ab.

84. or/75-83

85. exp "sensitivity and specificity"/

86. "reproducibility of results"/

87. (predict* adj3 (dement* or AD or alzheimer*)).ti,ab.

88. (identif* adj3 (dement* or AD or alzheimer*)).ti,ab.

89. (discriminat* adj3 (dement* or AD or alzheimer*)).ti,ab.

90. (distinguish* adj3 (dement* or AD or alzheimer*)).ti,ab.

91. (differenti* adj3 (dement* or AD or alzheimer*)).ti,ab.

92. diagnos*.ti.

93. di.fs.

94. sensitivit*.ab.

95. specificit*.ab.

96. (ROC or "receiver operat*").ab.

97. Area under curve/

98. ("Area under curve" or AUC).ab.

99. (detect* adj3 (dement* or AD or alzheimer*)).ti,ab.

100. sROC.ab.

101. accura*.ti,ab.

102. (likelihood adj3 (ratio* or function*)).ab.

103. (conver* adj3 (dement* or AD or alzheimer*)).ti,ab.

104. ((true or false) adj3 (positive* or negative*)).ab.

105. ((positive* or negative* or false or true) adj3 rate*).ti,ab.

106. or/85-105

107. exp dementia/di

108. Cognition Disorders/di [Diagnosis]

109. Memory Disorders/di

110. or/107-109

111. *Neuropsychological Tests/

112. *Questionnaires/

(Continued)

113. Geriatric Assessment/mt

114. *Geriatric Assessment/

115. Neuropsychological Tests/mt, st

116. "neuropsychological test*".ti,ab.

117. (neuropsychological adj (assess* or evaluat* or test*)).ti,ab.

118. (neuropsychological adj (assess* or evaluat* or test* or exam* or battery)).ti,ab.

119. Self report/

120. self-assessment/ or diagnostic self evaluation/

121. Mass Screening/

122. early diagnosis/

123. or/111-122

124. 74 or 123

125. 110 and 124

126. 74 or 123

127. 84 and 106 and 126

128. 74 and 106

129. 125 or 127 or 128

130. exp Animals/ not Humans.sh.

131. 129 not 130

Appendix 2. Information for extraction to proforma

Bibliographic details of primary paper.

- Author, title of study, year and journal.

Details of index test.

- Method of MMSE administration, including who administered and interpreted the test and their training.
- Thresholds used to define positive and negative tests.

Reference standard.

- Reference standard used.
- Method of reference standard administration, including who administered the test and their training.

Study population.

- Number of participants.
- Age.
- Gender.
- Other characteristics (e.g. APOE status).
- Settings: (i) community; (ii) primary care; (iii) secondary care outpatients; (iv) secondary care inpatients and residential care.
- Participant recruitment.
- Sampling procedures.

- Time between index test and reference standard.
- Proportion of people in sample with dementia.
- Subtype and stage of dementia if available.
- MCI definition used (if applicable).
- Duration of follow-up.
- Attrition and missing data.

Results of the 2 × 2 tables cross-relating index test results of the reference standards.

Table 1: Conversion from MCI to Alzheimer's disease dementia

Index test information	References standard information	
	ADD present	**ADD absent**
Index test positive	MMSE + who convert to ADD (TP)	MMSE + who remain MCI (FP) & MMSE+ who convert to ADD (FP)
Index test negative	MMSE - who convert to ADD (FN)	MMSE - who remain MCI (TN) & MMSE- who convert to ADD (TN)

Table 2: Conversion from MCI to non-Alzheimer's disease dementia

Index test information	References standard information	
	Non-ADD present	**Non-ADD absent**
Index test positive	MMSE + who convert to non-ADD (TP)	MMSE + who remain MCI (FP) & MMSE+ who convert to non-ADD (FP)
Index test negative	MMSE - who convert to non-ADD (FN)	MMSE - who remain MCI (TN) & MMSE- who convert to non-ADD (TN)

Table 3: Conversion from MCI to all-cause dementia

Index test information	References standard information	
	All-cause dementia present	**All-cause dementia absent**
Index test positive	MMSE + who convert to all-cause dementia (TP)	MMSE + who remain MCI (FP)
Index test negative	MMS - who convert to all-cause dementia (FN)	MMSE - who remain MCI (TN)

Appendix 3. Assessment of methodological quality QUADAS-2

Domain	Patient selection	Index test	Reference standard	Flow and timing

(Continued)

Description	Describe methods of patient selection Describe included participants (prior testing, presentation, intended use of index test and setting)	Describe the index test and how it was conducted and interpreted	Describe the reference standard and how it was conducted and interpreted	Describe any participants who did not receive the index test(s) and/or reference standard or who were excluded from the 2 × 2 table (refer to flow diagram) Describe the time interval and any interventions between index test(s) and reference standard
Signalling questions (yes, no, unclear)	Was a consecutive or random sample of participants enrolled? Was a case-control design avoided? Did the study avoid inappropriate exclusions?	Were the index test results interpreted without knowledge of results of the reference standard? If a threshold was used, was it prespecified?	Is the reference standard likely to correctly classify the target condition? Were the reference standard results interpreted without knowledge of results of the index test?	Did all participants receive a reference standard? Did all participants receive the same reference standard? Were all participants included in the analysis?
Risk of bias (high, low, unclear)	Could the selection of participants have introduced bias?	Could the conduct or interpretation of the index test have introduced bias?	Could the reference standard, its conduct or its interpretation have introduced bias?	Could the participant flow have introduced bias?
Concerns regarding applicability (high, low, unclear)	Are there concerns that included participants do not match the review question?	Are there concerns that the index test, its conduct or its interpretation differs from the review question?	Are there concerns that the target condition as defined by the reference standard does not match the review question?	

Appendix 4. Anchoring statements for quality assessment of MMSE diagnostic studies

We provide some core anchoring statements for quality assessment of diagnostic test accuracy of MMSE. These statements are designed for use with the QUADAS-2 tool and were derived during a 2-day, multidisciplinary focus group in 2010. If a QUADAS-2 signalling question for a specific domain is answered 'yes', then the risk of bias can be judged to be 'low'. If a question is answered 'no', this indicates risk of potential bias.

The focus group was tasked with judging the extent of the bias for each domain. During this process, it became clear that certain issues were key to assessing quality, whilst others were important to record but were less important for assessing overall quality. To assist, we describe a 'weighting' system. When an item is weighted 'high risk', that section of the QUADAS-2 results table is judged to have a high potential for bias if a signalling question is answered “no”. For example, in dementia, diagnostic test accuracy studies, ensuring that clinicians performing dementia assessment are blinded to results of the index test, are fundamental. If this blinding was not present, then the item on the reference standard should be scored 'high risk of bias', regardless of the other contributory elements. When an item is weighted 'low risk', it is judged to have a low potential for bias if a signalling question for that section of the QUADAS-2 results table is answered 'no'. Overall bias will be judged on whether other signalling questions (with a high risk of bias) for the same domain are also answered “no”. In assessing individual items, the score of "unclear" should be given only if there is genuine uncertainty. In these situations, review authors will contact the relevant study teams for additional information.

Anchoring statements to assist with assessment for risk of bias

Domain 1: patient selection

Risk of bias: could the selection of patients have introduced bias? (high, low, unclear)

Was a consecutive or random sample of patients enrolled?
When sampling is used, the methods least likely to cause bias are consecutive sampling and random sampling, which should be stated and/or described. Non-random sampling or sampling based on volunteers is more likely to be at high risk of bias.
Weighting: high risk of bias (no)

Was a case-control design avoided?

Case-control study designs have a high risk of bias, but sometimes they are the only studies available, especially if the index test is expensive and/or invasive. Nested case-control designs (systematically selected from a defined population cohort) are less prone to bias, but they will still narrow the spectrum of participants who receive the index test. Other study designs (both cohort and case-control) that may increase bias are those designs for which the study team deliberately increases or decreases the proportion of participants with the target condition, for example, a population study may be enriched with extra participants with dementia from a secondary care setting.
Weighting: high risk of bias (no)

Did the study avoid inappropriate exclusions?
The study will be automatically graded as unclear if exclusions are not detailed (pending contact with study authors). When exclusions are detailed, the study will be graded as "low risk" if exclusions are believed by the review authors to be appropriate. Exclusions common to many studies of dementia include the following: medical instability; terminal disease; alcohol/substance misuse; concomitant psychiatric diagnosis; and other neurodegenerative condition. However, if "difficult to diagnose" groups are excluded, this may introduce bias, so exclusion criteria must be justified. For a community sample, we would expect relatively few exclusions. Post hoc exclusions will be labelled "high risk" of bias.
Weighting: high risk of bias (no)

Applicability: are there concerns that included patients do not match the review question? (high, low, unclear)

Included patients should match the intended population as described in the review question. If not already specified in the review inclusion criteria, the setting will be particularly important—the review authors should consider population in terms of symptoms, pretesting and potential disease prevalence. Studies that use very selected participants or subgroups will be classified as having low applicability, unless they are intended to represent a defined target population, for example, people with memory problems referred to a specialist and investigated by lumbar puncture.

Domain 2: index test

Risk of bias: could the conduct or interpretation of the index test have introduced bias? (high, low, unclear)

Were MMSE results interpreted without knowledge of the reference standard?
Terms such as "blinded" or "independently and without knowledge of" are sufficient, and full details of the blinding procedure are not required. This item may be scored as "low risk" if explicitly described, or if a clear temporal pattern to the order of testing precludes the need for formal blinding (e.g. all MMSE assessments were performed before the dementia assessment). As most neuropsychological tests are administered by a third party, knowledge of the dementia diagnosis may influence ratings; tests that are self-administered, for example, use of a computerised version, may be associated with less risk of bias.
Weighting: high risk (no)

Were MMSE thresholds prespecified?
For neuropsychological scales, there is usually a threshold above which participants are classified as "test positive"; this may be referred to as the *threshold,* the *clinical cut-off* or the *dichotomisation point*. Different thresholds are used in different populations. A study is classified as having higher risk of bias if the study authors define the optimal cut-off post hoc on the basis of their own study data. Some papers use an alternative methodology for analysis that does not use thresholds; these papers should be classified as not applicable.
Weighting: high risk (no)

Were sufficient data on MMSE application given for the test to be repeated in an independent study?
Particular points of interest include method of administration (e.g. self-completed questionnaire vs direct questioning interview); nature of the informant; and language of the assessment. If a novel form of the index test is used, for example, a translated questionnaire, details of the scale should be included and a reference given to an appropriate descriptive text; evidence of validation should be provided.
Weighting: high risk (no)

Applicability: are there concerns that the index test, its conduct or its interpretation may differ from the review question? (high, low, unclear)

Variations in length, structure, language and/or administration of the index test may affect applicability if they vary from those specified in the review question.

Domain 3: reference standard

Risk of bias: could the reference standard, its conduct or its interpretation have introduced bias? (high, low, unclear)

Is the assessment used for clinical diagnosis of dementia acceptable?
Commonly used international criteria to assist with clinical diagnosis of dementia include those detailed in *DSM-IV* and *ICD-10*. Criteria specific to dementia subtypes include but are not limited to NINCDS-ADRDA criteria for Alzheimer's dementia, McKeith criteria for Lewy body dementia, Lund-Manchester criteria for fronto-temporal dementia and NINDS-AIREN criteria for vascular dementia. When the criteria used for assessment are not familiar to the review authors and the Cochrane Dementia and Cognitive Improvement Group, this item should be classified as "high risk of bias".
Weighting: high risk (no)

Was clinical assessment for dementia performed without knowledge of the MMSE results?
Terms such as “blinded” and “independent” are sufficient, and full details of the blinding procedure are not required. Interpretation of results of the reference standard may be influenced by knowledge of results of the index test.
Weighting: high risk (no)

Applicability: are there concerns that the target condition as defined by the reference standard does not match the review question? (high, low, unclear)

Some methods of dementia assessment, although valid, may diagnose a far smaller or larger proportion of individuals with the disease than in usual clinical practise. For example, currently the reference standard for vascular dementia may underdiagnose compared with usual clinical practise. In this instance, the item should be rated as having poor applicability.

Domain 4: participant flow and timing (note refer to, or construct, a flow diagram)

Risk of bias: could the participant flow have introduced bias? (high, low, unclear)

Was there an appropriate interval between MMSE and the reference standard?
As we test the accuracy of the MMSE test for MCI conversion to dementia, a delay will always be noted between the index test and the reference standard assessments. The time between reference standard and index test will influence the accuracy, and therefore we will note time as a separate variable (both within and between studies) and will test its influence on diagnostic accuracy. We have set a minimum mean time to follow-up assessment of 1 year. If more than 16% of participants undergo assessment for MCI conversion before 9 months, this item will score "no".
Weighting: high risk (no)

Did all participants receive the same reference standard?
In some scenarios, participants who score “test positive” on the index test may undergo a more detailed assessment for the target condition. When dementia assessment (or reference standard) differs between participants, this should be classified as high risk of bias.
Weighting: high risk (no)

Were all participants included in the final analysis?
If the number of participants enrolled differs from the number of participants included in the 2 × 2 table, the potential for bias exists. If participants lost to follow-up differ systematically from those who remain, then estimates of test performance may differ. If drop-outs are present, these should be accounted for; the maximum proportion of drop-outs for low risk of bias has been specified as 20%. Details of the causes of study drop-outs are crucial, and if such data are missing, the reliability of the conclusions must be questioned.
Weighting: high risk (no)

WHAT'S NEW

Date	Event	Description
15 June 2021	Amended	The title and objectives have been changed to make it clear that screening tests alone cannot give a diagnostic formulation. There have been no other changes. Similar changes have been made to other Cochrane diagnostic test accuracy reviews relating to dementia. These changes were made by Cochrane Dementia and Cognitive Improvement in conjunction with the Cochrane Mental Health and Neuroscience Network and the review authors, following feedback from a group of dementia researchers, expressing concern that the review titles implied that short screening tests could make a diagnosis of a dementia subtype like Alzheimer’s disease. This interpretation was not intended, but the revised titles and objectives clarify the reviews’ scope. Further details are available here: https://dementia.cochrane.org/our-reviews/feedback-about-reviews
15 June 2021	New citation required but conclusions have not changed	Title and objectives have changed for clarification

HISTORY

Protocol first published: Issue 10, 2013
Review first published: Issue 3, 2015

CONTRIBUTIONS OF AUTHORS

IA-R: designed and drafted protocol; overall responsibility for study selection and data extraction; completed characteristics of included studies; data entry check; QUADAS-2 assessment; set up data and analysis tables; completed SOF table and additional tables; drafted Results and Discussion section and finalised manuscript.

NS: designed and drafted protocol; study selection and data extraction; completed characteristics of included and excluded studies tables; entered data and data entry check; QUADAS-2 assessment; set up data and analysis tables; updated Methods and drafted the Results section; managed the review process and produced progress reports, attended progress meetings and worked with all review authors to ensure that the review met publication deadlines.

AC: designed and drafted protocol; study selection and data extraction; QUADAS-2 assessment; drafted Results and Discussion sections and finalised manuscript.

ES-P: designed and drafted protocol; study selection and data extraction; drafted Results and Discussion sections and finalised manuscript.

AG: designed and drafted protocol; set up data and analysis tables; drafted Results and Discussion sections and finalised manuscript.

MRF: designed and drafted protocol; QUADAS-2 assessment; set up data and analysis tables; drafted Results and Discussion sections and finalised manuscript.

OLP: designed and drafted protocol; study selection and data extraction; drafted Results and Discussion sections and finalised manuscript.

XBC: designed and drafted protocol; drafted Results and Discussion sections and finalised manuscript.

SC: designed and drafted protocol; drafted Results and Discussion sections and finalised manuscript.

DECLARATIONS OF INTEREST

None known

SOURCES OF SUPPORT

Internal sources

- Fundación Universitaria de Ciencias de la Salud, Hospital San José/Hospital Infantil de San José, Bogotá D.C., Colombia
- Institute for Clinical Effectiveness and Health Policy IECS, Buenos Aires, Argentina
- Iberoamerican Cochrane Centre, Barcelona, Spain

External sources

- Agencia de Calidad del Sistema Nacional de Salud, Ministry of Health, Madrid, Spain

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

1. Target condition was amended to reflect the group of dementia in general (defined by studies): "The target condition was conversion at follow-up from MCI to Alzheimer's disease dementia or other forms of dementia." change to "The target condition was conversion at follow-up from MCI to all-cause dementia, Alzheimer's disease dementia or other forms of dementia."

2. The objective was amended to reflect the group of dementia in general (defined by studies): "To determine the diagnostic accuracy of the MMSE at various thresholds for detecting individuals with MCI at baseline who would clinically convert to Alzheimer’s disease dementia or other forms of dementia at follow-up" changed to "To determine the diagnostic accuracy of the MMSE at various thresholds for detecting individuals with MCI at baseline who would clinically convert to all-cause dementia, Alzheimer’s disease dementia or other forms of dementia at follow-up"

3. We planned (Arevalo-Rodriguez 2013a) to investigate the following but these investigations were not undertaken due to the scarcity of data.

Index test

- Thresholds

- Technical features (including different versions of the test)
- Operator characteristics (e.g. training)

Target disorder

- Reference standards used: DSM definition, ICD definition, NINDS-ARDRA or other classification, including pathological definitions, and operationalisation of these classifications (e.g. individual clinician, algorithm, consensus group)
- Spectrum of target disorder

Target population

- Age, sex, education, sociocultural variables (social network/social engagement)
- Other characteristics (e.g. APOE status, definition and duration of MCI at baseline (if applicable))
- Prevalence in different settings
- Treatment: previous or current interventions

Study quality

- Types of studies
- Prior clinical information to increase the accuracy of the index test
- Duration of follow-up (measured in years for delayed-verification studies)
- Loss to follow-up

INDEX TERMS

Medical Subject Headings (MeSH)

Alzheimer Disease [diagnosis]; Cognitive Dysfunction [*complications]; Dementia [*diagnosis] [etiology]; Dementia, Vascular [diagnosis] [etiology]; Disease Progression; Early Diagnosis; Frontotemporal Dementia [diagnosis] [etiology]; Lewy Body Disease [diagnosis] [etiology]; *Mental Status and Dementia Tests; Neuropsychological Tests; Sensitivity and Specificity

MeSH check words

Humans