



Cochrane
Library

Cochrane Database of Systematic Reviews

Structural magnetic resonance imaging for the early diagnosis of dementia due to Alzheimer's disease in people with mild cognitive impairment (Review)

Lombardi G, Crescioli G, Cavedo E, Lucenteforte E, Casazza G, Bellatorre AG, Lista C, Costantino G, Frisoni G, Virgili G, Filippini G

Lombardi G, Crescioli G, Cavedo E, Lucenteforte E, Casazza G, Bellatorre AG, Lista C, Costantino G, Frisoni G, Virgili G, Filippini G. Structural magnetic resonance imaging for the early diagnosis of dementia due to Alzheimer's disease in people with mild cognitive impairment.

Cochrane Database of Systematic Reviews 2020, Issue 3. Art. No.: CD009628.

DOI: [10.1002/14651858.CD009628.pub2](https://doi.org/10.1002/14651858.CD009628.pub2).

www.cochranelibrary.com

Structural magnetic resonance imaging for the early diagnosis of dementia due to Alzheimer's disease in people with mild cognitive impairment (Review)

Copyright © 2020 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

WILEY

TABLE OF CONTENTS

HEADER	1
ABSTRACT	1
PLAIN LANGUAGE SUMMARY	2
SUMMARY OF FINDINGS	4
BACKGROUND	6
OBJECTIVES	7
METHODS	8
RESULTS	10
Figure 1.	12
Figure 2.	15
Figure 3.	17
Figure 4.	18
Figure 5.	20
Figure 6.	21
Figure 7.	23
DISCUSSION	24
AUTHORS' CONCLUSIONS	25
ACKNOWLEDGEMENTS	25
REFERENCES	27
CHARACTERISTICS OF STUDIES	49
DATA	156
Test 1. Hippocampus total.	157
Test 2. Hippocampus left.	157
Test 3. Hippocampus right.	157
Test 4. Medial temporal lobe total.	157
Test 5. Medial temporal lobe left.	157
Test 6. Medial temporal lobe right.	157
Test 7. Lateral ventricles.	157
Test 8. Entorhinal cortex total.	157
Test 9. Entorhinal cortex left.	157
Test 10. Entorhinal cortex right.	158
Test 11. Whole brain.	158
Test 12. Medial temporal gyrus.	158
Test 13. Lateral temporal lobe total.	158
Test 14. Lateral temporal lobe right.	158
Test 15. Amygdala total.	158
Test 16. Amygdala left.	158
Test 17. Amygdala right.	158
Test 18. Cortical grey matter.	158
ADDITIONAL TABLES	158
APPENDICES	172
CONTRIBUTIONS OF AUTHORS	183
DECLARATIONS OF INTEREST	183
SOURCES OF SUPPORT	184
DIFFERENCES BETWEEN PROTOCOL AND REVIEW	184

[Diagnostic Test Accuracy Review]

Structural magnetic resonance imaging for the early diagnosis of dementia due to Alzheimer's disease in people with mild cognitive impairment

Gemma Lombardi¹, Giada Crescioli¹, Enrica Cavedo², Ersilia Lucenteforte³, Giovanni Casazza⁴, Alessandro-Giacco Bellatorre⁵, Chiara Lista⁶, Giorgio Costantino⁷, Giovanni Frisoni⁸, Gianni Virgili¹, Graziella Filippini⁹

¹Department of Neurosciences, Psychology, Drug Research and Child Health (NEUROFARBA), University of Florence, Florence, Italy. ²Alzheimer Precision Medicine (APM), AP-HP, Pitie-Salpetriere Hospital, Sorbonne University, Paris, France. ³Department of Clinical and Experimental Medicine, University of Pisa, Pisa, Italy. ⁴Dipartimento di Scienze Biomediche e Cliniche "L. Sacco", Università degli Studi di Milano, Milan, Italy. ⁵Dipartimento di Medicina Interna, Azienda sanitaria universitaria integrata di Udine, Udine, Italy. ⁶Neuroepidemiology Unit, Fondazione I.R.C.C.S. Istituto Neurologico Carlo Besta, Milan, Italy. ⁷UOC Pronto Soccorso e Medicina D'Urgenza, Fondazione IRCCS Ca' Granda, Ospedale Maggiore Policlinico, Università degli Studi di Milano, Milan, Italy. ⁸University Hospitals and University of Geneva, Geneva, Switzerland. ⁹Scientific Director's Office, Carlo Besta Foundation and Neurological Institute, Milan, Italy

Contact address: Graziella Filippini, Scientific Director's Office, Carlo Besta Foundation and Neurological Institute, via Celoria, 11, Milan, 20133, Italy. graziella.filippini@istituto-besta.it, filippini.graziella@gmail.com.

Editorial group: Cochrane Dementia and Cognitive Improvement Group.

Publication status and date: New, published in Issue 3, 2020.

Citation: Lombardi G, Crescioli G, Cavedo E, Lucenteforte E, Casazza G, Bellatorre AG, Lista C, Costantino G, Frisoni G, Virgili G, Filippini G. Structural magnetic resonance imaging for the early diagnosis of dementia due to Alzheimer's disease in people with mild cognitive impairment. *Cochrane Database of Systematic Reviews* 2020, Issue 3. Art. No.: CD009628. DOI: [10.1002/14651858.CD009628.pub2](https://doi.org/10.1002/14651858.CD009628.pub2).

Copyright © 2020 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

ABSTRACT

Background

Mild cognitive impairment (MCI) due to Alzheimer's disease is the symptomatic prodementia phase of Alzheimer's disease dementia, characterised by cognitive and functional impairment not severe enough to fulfil the criteria for dementia. In clinical samples, people with amnesic MCI are at high risk of developing Alzheimer's disease dementia, with annual rates of progression from MCI to Alzheimer's disease estimated at approximately 10% to 15% compared with the base incidence rates of Alzheimer's disease dementia of 1% to 2% per year.

Objectives

To assess the diagnostic accuracy of structural magnetic resonance imaging (MRI) for the early diagnosis of dementia due to Alzheimer's disease in people with MCI versus the clinical follow-up diagnosis of Alzheimer's disease dementia as a reference standard (delayed verification).

To investigate sources of heterogeneity in accuracy, such as the use of qualitative visual assessment or quantitative volumetric measurements, including manual or automatic (MRI) techniques, or the length of follow-up, and age of participants.

MRI was evaluated as an add-on test in addition to clinical diagnosis of MCI to improve early diagnosis of dementia due to Alzheimer's disease in people with MCI.

Search methods

On 29 January 2019 we searched Cochrane Dementia and Cognitive Improvement's Specialised Register and the databases, MEDLINE, Embase, BIOSIS Previews, Science Citation Index, PsycINFO, and LILACS. We also searched the reference lists of all eligible studies identified by the electronic searches.

Selection criteria

We considered cohort studies of any size that included prospectively recruited people of any age with a diagnosis of MCI. We included studies that compared the diagnostic test accuracy of baseline structural MRI versus the clinical follow-up diagnosis of Alzheimer's disease dementia (delayed verification). We did not exclude studies on the basis of length of follow-up. We included studies that used either qualitative visual assessment or quantitative volumetric measurements of MRI to detect atrophy in the whole brain or in specific brain regions, such as the hippocampus, medial temporal lobe, lateral ventricles, entorhinal cortex, medial temporal gyrus, lateral temporal lobe, amygdala, and cortical grey matter.

Data collection and analysis

Four teams of two review authors each independently reviewed titles and abstracts of articles identified by the search strategy. Two teams of two review authors each independently assessed the selected full-text articles for eligibility, extracted data and solved disagreements by consensus. Two review authors independently assessed the quality of studies using the QUADAS-2 tool. We used the hierarchical summary receiver operating characteristic (HSROC) model to fit summary ROC curves and to obtain overall measures of relative accuracy in subgroup analyses. We also used these models to obtain pooled estimates of sensitivity and specificity when sufficient data sets were available.

Main results

We included 33 studies, published from 1999 to 2019, with 3935 participants of whom 1341 (34%) progressed to Alzheimer's disease dementia and 2594 (66%) did not. Of the participants who did not progress to Alzheimer's disease dementia, 2561 (99%) remained stable MCI and 33 (1%) progressed to other types of dementia. The median proportion of women was 53% and the mean age of participants ranged from 63 to 87 years (median 73 years). The mean length of clinical follow-up ranged from 1 to 7.6 years (median 2 years). Most studies were of poor methodological quality due to risk of bias for participant selection or the index test, or both.

Most of the included studies reported data on the volume of the total hippocampus (pooled mean sensitivity 0.73 (95% confidence interval (CI) 0.64 to 0.80); pooled mean specificity 0.71 (95% CI 0.65 to 0.77); 22 studies, 2209 participants). This evidence was of low certainty due to risk of bias and inconsistency.

Seven studies reported data on the atrophy of the medial temporal lobe (mean sensitivity 0.64 (95% CI 0.53 to 0.73); mean specificity 0.65 (95% CI 0.51 to 0.76); 1077 participants) and five studies on the volume of the lateral ventricles (mean sensitivity 0.57 (95% CI 0.49 to 0.65); mean specificity 0.64 (95% CI 0.59 to 0.70); 1077 participants). This evidence was of moderate certainty due to risk of bias.

Four studies with 529 participants analysed the volume of the total entorhinal cortex and four studies with 424 participants analysed the volume of the whole brain. We did not estimate pooled sensitivity and specificity for the volume of these two regions because available data were sparse and heterogeneous.

We could not statistically evaluate the volumes of the lateral temporal lobe, amygdala, medial temporal gyrus, or cortical grey matter assessed in small individual studies.

We found no evidence of a difference between studies in the accuracy of the total hippocampal volume with regards to duration of follow-up or age of participants, but the manual MRI technique was superior to automatic techniques in mixed (mostly indirect) comparisons. We did not assess the relative accuracy of the volumes of different brain regions measured by MRI because only indirect comparisons were available, studies were heterogeneous, and the overall accuracy of all regions was moderate.

Authors' conclusions

The volume of hippocampus or medial temporal lobe, the most studied brain regions, showed low sensitivity and specificity and did not qualify structural MRI as a stand-alone add-on test for an early diagnosis of dementia due to Alzheimer's disease in people with MCI. This is consistent with international guidelines, which recommend imaging to exclude non-degenerative or surgical causes of cognitive impairment and not to diagnose dementia due to Alzheimer's disease. In view of the low quality of most of the included studies, the findings of this review should be interpreted with caution. Future research should not focus on a single biomarker, but rather on combinations of biomarkers to improve an early diagnosis of Alzheimer's disease dementia.

PLAIN LANGUAGE SUMMARY

How accurate is magnetic resonance imaging for the early diagnosis of dementia due to Alzheimer's disease in people with mild cognitive impairment?

Why is improving Alzheimer's disease diagnosis important?

Cognitive impairment is when people have problems remembering, learning, concentrating and making decisions. People with mild cognitive impairment (MCI) generally have more memory problems than other people of their age, but these problems are not severe enough to be classified as dementia. Studies have shown that people with MCI and loss of memory are more likely to develop Alzheimer's disease dementia (approximately 10% to 15% of cases per year) than people without MCI (1% to 2% per year). Currently, the only reliable way of diagnosing Alzheimer's disease dementia is to follow people with MCI and assess cognitive changes over the years. Magnetic resonance imaging (MRI) may detect changes in the brain structures that indicate the beginning of Alzheimer's disease. Early diagnosis of MCI due to Alzheimer's disease is important because people with MCI could benefit from early treatment to prevent or delay cognitive decline.

What was the aim of this review?

To assess the diagnostic accuracy of MRI for the early diagnosis of dementia due to Alzheimer's disease in people with MCI.

What was studied in the review?

The volume of several brain regions was measured with MRI. Most studies (22 studies, 2209 participants) measured the volume of the hippocampus, a region of the brain that is associated primarily with memory.

What are the main results in this review?

Thirty-three studies were eligible, in which 3935 participants with MCI were included and followed up for two or three years to see if they developed Alzheimer's disease dementia. About a third of them converted to Alzheimer's disease dementia, and the others did not or developed other types of dementia.

We found that MRI is not accurate enough to identify people with MCI who will develop dementia due to Alzheimer's disease. The correct prediction of Alzheimer's disease would be missed in 81 out of 300 people with MCI (false negatives) and a wrong prediction of Alzheimer's disease would be made in 203 out of 700 people with MCI (false positives). As a result, people with a false-negative diagnosis would be falsely reassured and would not prepare themselves to cope with Alzheimer's disease, while those with a false-positive diagnosis would suffer from the wrongly anticipated diagnosis.

How reliable are the results of the studies?

The included studies diagnosed Alzheimer's disease dementia by assessing all participants with standard clinical criteria after two or three years' follow-up. We had some concerns about how the studies were conducted, since the participants were mainly selected from clinical registries and referral centres, and we also had concerns about how studies interpreted MRI. Moreover, the studies were conducted differently from each other, and they used different methods to select people with MCI and perform MRI.

Who do the results of this review apply to?

The results do not apply to people with MCI in the community, but only to people with MCI who attend memory clinics or referral centres.

What are the implications of this review?

MRI, as a single test, is not accurate for the early diagnosis of dementia due to Alzheimer's disease in people with MCI since one in three or four participants received a wrong diagnosis of Alzheimer's disease. Future research should not focus on a single test (such as MRI), but rather on combinations of tests to improve an early diagnosis of Alzheimer's disease dementia.

How up to date is this review?

This evidence is up to date to 29 January 2019.

SUMMARY OF FINDINGS

Summary of findings 1. Whole brain volume or volume of specific brain regions for early Alzheimer's disease dementia diagnosis in people with mild cognitive impairment

Whole brain volume versus volume of specific brain regions for early Alzheimer's disease dementia diagnosis in people with mild cognitive impairment

Patient or population: people with mild cognitive impairment (MCI)

Setting: memory clinics or registry data (e.g. ADNI)

New test: volume of total hippocampus, medial temporal lobe, total entorhinal cortex, lateral ventricles, and whole brain. Volume measured with either quantitative manual or automated MRI technique

Cut-off value: not reported

Number of results per 1000 participants tested (95% CI)								
Prevalence 30%. Typically seen in participants with MCI after 2 to 3 years of follow-up								
Test	Number of participants (Number of studies)	True positives	False negatives	True negatives	False positives	Pooled sensitivity (95% CI)	Pooled specificity (95% CI)	Certainty of the evidence (GRADE)
Total hippocampus	2209 (22)	219 (192 to 240)	81 (60 to 108)	497 (455 to 539)	203 (161 to 245)	0.73 (0.64 to 0.80)	0.71 (0.65 to 0.77)	⊕⊕⊕⊖ Low ^{a,b}
Medial temporal lobe	1077 (7)	192 (159 to 219)	108 (81 to 141)	455 (357 to 532)	245 (168 to 343)	0.64 (0.53 to 0.73)	0.65 (0.51 to 0.76)	⊕⊕⊕⊕ Moderate ^{a,c}
Lateral ventricles	1077 (5)	171 (147 to 195)	129 (105 to 153)	448 (413 to 490)	252 (210 to 287)	0.57 (0.49 to 0.65)	0.64 (0.59 to 0.70)	⊕⊕⊕⊕ Moderate ^{a,c}
Total entorhinal cortex	529 (4)	Meta-analyses not conducted due to sparse and heterogeneous data				Range: 0.50 to 0.88	Range: 0.60 to 1.00	⊕⊕⊕⊖ Very low ^{a,d}
Whole brain	424 (4)	Meta-analyses not conducted due to sparse and heterogeneous data				Range: 0.33 to 0.92	Range: 0.41 to 1.00	⊕⊕⊕⊖ Very low ^{a,d}

The table displays normalised number of participants within a hypothetical cohort of 1000 people at a prevalence of Alzheimer's disease (pre-test probabilities) of 30%. We selected a prevalence value based on a prevalence observed in people with MCI after 2 to 3 years of follow-up. We estimated confidence intervals based on those around the point estimates for pooled sensitivity and specificity.

ADNI: Alzheimer's Disease Neuroimaging Initiative; **CI:** confidence interval; **MCI:** mild cognitive impairment; **MRI:** magnetic resonance imaging

GRADE Working Group GRADES of evidence

High certainty: we are very confident that the true effect lies close to the estimate of the effect.

Moderate certainty: we are moderately confident in the effect estimate: the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.

Low certainty: our confidence in the effect estimate is limited: the true effect may be substantially different from the estimate of the effect.

Very low certainty: we have very little confidence in the effect estimate: the true effect is likely to be substantially different from the estimate of effect.

^aRisk of bias: most studies were at high risk of bias for participant selection (registry data), or index test or both. We downgraded the certainty of the evidence by one level.

^bImprecision: wide 95% confidence intervals. We downgraded the certainty of the evidence by one level

^cImprecision: wide 95% confidence intervals, however upper limit for both sensitivity and specificity are below 0.75, which is a modest performance. We did not downgrade.

^dInconsistency and imprecision: sparse and inconsistent data. We downgraded the certainty of the evidence by one level both for inconsistency and imprecision.

BACKGROUND

The shift from normal aging to Alzheimer's Disease dementia is a continuous process where the transitional state between normal cognition and Alzheimer's disease dementia progressively involves, to a variable extent and in different stages, episodic memory (i.e. the ability to learn and retain new information), executive functions (e.g. set-shifting, reasoning, problem-solving, planning), language (e.g. naming, fluency, expressive speech, and comprehension), visuospatial skills, attention and perceptual speed (Bäckman 2004). The criteria of a clinical diagnosis of probable or possible Alzheimer's disease dementia were proposed in 1984 by the National Institute of Neurological and Communicative Disorders and Stroke (NINCDS) and the Alzheimer's Disease and Related Disorders Association (ADRDA) (the NINCDS-ADRDA criteria; Appendix 1). A diagnosis of definite Alzheimer's disease dementia requires clinical criteria for probable Alzheimer's disease and histopathologic evidence obtained from a biopsy or autopsy, but these are not applicable in daily clinical practice (McKhann 1984). The NINCDS-ADRDA criteria were updated in 2011 by the National Institute on Aging (NIA) and the Alzheimer's Association (AA), known as the NIA-AA criteria (McKhaan 2011). In agreement with the NINCDS-ADRDA criteria, the NIA-AA criteria require a significant interference in the ability to function at work or in usual daily activities. The presence of any positive biomarker (e.g. medial temporal lobe atrophy detected by MRI) is not essential for the diagnosis but is useful to investigate the "biomarker probability of AD [Alzheimer's disease] dementia etiology" (McKhaan 2011).

The NIA-AA criteria for the diagnosis of MCI due to Alzheimer's disease dementia define MCI as the symptomatic pre-dementia phase of Alzheimer's disease dementia, and include two sets of criteria: (1) core clinical criteria that comprise evidence of concern about a change in cognition, in comparison with the person's previous level; lower performance in one or more cognitive domains that is greater than would be expected for the patient's age and educational background; preservation of independence in functional abilities and no evidence of a significant impairment in social or occupational functioning; and (2) the use of biomarkers based on imaging and cerebrospinal fluid measures in clinical research settings (Albert 2011; Appendix 2). Single or multiple cognitive domains may be affected in a person with MCI. If memory only is affected, MCI is defined as 'amnesic'. When single or more cognitive domains different from memory are affected, MCI is defined as 'non-amnesic'. In clinical series, people with amnesic MCI are at high risk of developing Alzheimer's disease dementia, with annual rates of progression from amnesic MCI to Alzheimer's disease dementia estimated at 10% to 15% compared with the base incidence rates of Alzheimer's disease dementia of 1% to 2% per year (Petersen 2009). In the general population also, people with amnesic MCI are at high risk of progression to Alzheimer's disease dementia over three years (Palmer 2008). Progression is high in the first few years following MCI diagnosis (Mitchell 2009). However, people diagnosed with MCI may be stable or revert to normal condition over time, while some of them may develop non-Alzheimer's disease dementia (Palmer 2008).

People with early cognitive impairments are increasingly presenting to both primary and secondary care (NICE 2018). Since an early intervention could be more effective in delaying the development of dementia, these people may represent the

suitable target for addressing future disease-modifying therapies. A survey conducted among members of the American Academy of Neurology, who had an aging, dementia, or behavioural neurology practice focus, found that the majority of respondents recognised MCI as a clinical diagnosis and used its diagnostic code for billing purposes. When seeing these patients, most respondents routinely communicated the dementia risk and sometimes prescribed cholinesterase inhibitors (Roberts 2010). While our protocol was under development, the ClinicalTrials.gov (clinicaltrials.gov), registry contained 230 references referring to completed or ongoing trials of medication as well as non-medication approaches for treating MCI.

In 2018 the NIA-AA published a "Research framework: towards a biological definition of Alzheimer's disease", which defined Alzheimer's disease on the basis of biomarkers as a proxy for the neuropathology of Alzheimer's disease (Jack 2018). Recommended biomarkers are markers of amyloid deposition (A), markers of neurofibrillary tangles tau (T), and markers of neurodegeneration (N). For each category, both a cerebrospinal fluid and a neuroimaging biomarker were suggested. According to the biological definition of Alzheimer's disease, cognitive symptoms can be added to the ATN system but are not mandatory for the diagnosis. The NIA-AA emphasised that it is premature and inappropriate to use this research framework in clinical practice. In a published comment Cochrane Dementia and Cognitive Improvement reported that the biomarkers described in the NIA-AA framework are neither sensitive nor specific to the diagnosis of Alzheimer's disease dementia (McCleery 2019).

In this context, the objective of this review is to determine the accuracy of structural MRI for the early diagnosis of dementia due to Alzheimer's disease in people with MCI.

Target condition being diagnosed

The primary target condition is dementia due to Alzheimer's disease, a degenerative disease of the brain accounting for 60% to 80% of dementia cases. In 2019 the Alzheimer's Disease International (ADI) estimates that there were over 50 million people living with dementia globally, a number set to increase to 152 million by 2050, primarily driven by increased longevity (ADI 2019).

Index test(s)

This review assesses the diagnostic accuracy of structural MRI in detecting atrophy in the whole brain and in specific brain regions, such as the hippocampus, lateral ventricles, entorhinal cortex, amygdala, medial temporal lobe, lateral temporal lobe, medial temporal gyrus, and cortical grey matter. Structural MRI assesses the structure of the brain tissues as opposed to functional MRI, which assesses functional brain activity. Atrophy is a decrease in volume of tissues.

MRI does not involve X-rays or the use of ionising radiation. It is non-invasive and has no significant adverse health effects. Patients are at risk of injury from MRI if they have metal objects in their bodies, such as pacemakers, clips or metallic prostheses. Since individuals with fear of confined spaces may become anxious during MRI, the test is contraindicated in people with claustrophobia.

Clinical pathway

Alzheimer's disease dementia shows an insidious onset characterised by progressive decline of cognitive functions such as memory, thinking, comprehension, calculation, language, learning capacity and judgement that are sufficient to impair personal activities of daily living (McKhann 1984). This disease needs to be clearly differentiated from age-related cognitive decline. The onset of Alzheimer's disease dementia is usually after 65 years of age, though earlier onset is not uncommon. As age advances, the incidence increases rapidly (it roughly doubles every five years). Since life expectancy increases in the population, the total number of individuals affected by dementia is expected to rise. Dementia due to Alzheimer's disease has economic as well as quality of life-related consequences, not only for the patients but also for their families.

People who present with symptoms of cognitive decline generally are evaluated first by the general practitioner who obtains information from the patient or a family member. The National Institute for Health and Clinical Excellence (NICE) guideline recommends that if dementia is suspected after the initial clinical judgement, a physical examination, appropriate blood and urine tests to exclude reversible causes of cognitive decline, and a cognitive assessment should be undertaken. Moreover, if dementia is still suspected, the person should be referred to a specialist dementia diagnostic service (such as a memory clinic or community old age psychiatry service). Specialists have to confirm cognitive decline, rule out reversible causes and, when possible, diagnose the dementia subtype. Brain computed tomography (CT) or MRI should be used to rule out reversible causes of cognitive decline and to assist the subtype diagnosis, unless dementia is well established and the subtype is clear (NICE 2018).

Role of index test(s)

We evaluated the potential role of structural MRI in improving an early diagnosis of dementia due to Alzheimer's disease in people with MCI when MRI is used in addition to clinical judgement or cognitive test performance or both (add-on test). Hippocampal atrophy measured by MRI has been qualified by the European Medicines Agency (EMA) for enrichment in regulatory clinical trials in the pre-dementia stage of Alzheimer's disease (European Medicines Agency 2011). The Food and Drug Administration (FDA) issued a letter supporting the role of a low baseline hippocampal volume as a prognostic biomarker for enrichment (US Food and Drug Administration 2015). Hippocampal or medial temporal lobe atrophy measured on MRI has been included as a marker of neuronal injury in the recommendations of the NIA-AA on the diagnosis of MCI due to Alzheimer's disease (Albert 2011). Although no treatment is currently available to cure MCI due to Alzheimer's disease, an early diagnosis of Alzheimer's disease dementia could be of significant support for patients and their families. For example, lifestyle interventions to prevent or postpone the onset of dementia or inclusion in clinical trials might be suggested to people with a diagnosis of MCI at risk of progression to Alzheimer's disease dementia.

Alternative test(s)

We did not include an alternative test in the review. An initial single-test review is preliminary to conducting comparative reviews or reviews of test combinations. The accuracy of other biomarkers (cerebrospinal fluid (CSF) biomarkers, plasma biomarkers, amyloid

positron imaging tomography (PET), fluorodeoxyglucose (FDG) PET for the longitudinal prediction of dementia due to Alzheimer's disease and other dementias in people with cognitive decline but no dementia are presented in other Cochrane Reviews (Ritchie 2014; Ritchie 2017, Martínez 2017; Smailagic 2015).

Rationale

MCI is considered either a risk factor or a symptomatic pre-dementia phase. MCI represents a target to better understand mechanisms underlying dementia onset and progression, and a clinical condition to test preventive strategies or early intervention. The Lancet Commission on prevention and management of dementia reported a large body of research evidence showing that interventions for improving modifiable risk factors might have the potential to delay or prevent a third of dementia cases (Livingston 2017). Early diagnosis of dementia due to Alzheimer's disease would facilitate timely referral to education, counselling and support services for people with cognitive impairment and their carers, and would likely allow input from the patients about their care plans. An early differential diagnosis is also important to identify treatable medical causes of cognitive impairment, such as depression, metabolic conditions, cardiovascular or cerebrovascular disease. Moreover, early diagnosis would allow people to participate in treatment trials preventing or delaying cognitive decline (Livingston 2017). Currently there are more than 200 treatments under investigation (www.clinicaltrials.gov), and a large consensus exists on the hypothesis that the earlier the intervention takes place, the greater will be the protection against further neuronal damage. Disease-modifying approaches for people with MCI require better knowledge of the accuracy of diagnostic tests that are used in clinical trials. The new criteria for the diagnosis of MCI due to Alzheimer's disease (Albert 2011), incorporate biomarkers based on imaging and CSF measures in order to increase the probability to identify MCI due to Alzheimer's disease. These biomarkers used with clinical judgement might increase the sensitivity or specificity of a testing strategy. However, biomarkers must be preliminarily assessed for individual accuracy before starting to use them as add-on tests in clinical practice.

OBJECTIVES

To assess the diagnostic accuracy of structural MRI for the early diagnosis of dementia due to Alzheimer's disease in people with MCI versus the clinical follow-up diagnosis of Alzheimer's disease dementia as a reference standard (delayed verification).

Secondary objectives

To investigate sources of heterogeneity in the diagnostic accuracy of structural MRI for the early diagnosis of dementia due to Alzheimer's disease in people with MCI. Potential sources of heterogeneity included the following.

1. Setting: referral centres versus population cohorts
2. Patient spectrum: mean or median age and amnesic versus non-amnesic MCI
3. Mean or median duration of follow-up: less than three years versus three years or longer
4. MRI region of interest: medial temporal lobe versus other structures and, if possible, hippocampus versus other structures, entorhinal cortex versus other structures, and temporoparietal regions versus other structures

5. MRI technology: magnetic field strength less than 1 Tesla versus 1 Tesla or higher
6. MRI techniques: visual versus manual versus automatic and semiautomatic computer-based methods.

METHODS

Criteria for considering studies for this review

Types of studies

We included studies if they:

1. were prospective cohort studies with a clinical follow-up as a reference standard for diagnosis of dementia due to Alzheimer's disease. In the cohort design, participants are enrolled and undergo the index test before the final outcome (presence or absence of Alzheimer's disease dementia) is known
2. contained sufficient data to construct 2 x 2 contingency tables expressing MRI results by disease status
3. were conducted in any healthcare setting, that is, population-based studies or clinical settings
4. were published in any language.

We excluded case series or case-control studies, which lead to inflated estimates of disease prevalence and test accuracy (Lijmer 1999; Whiting 2004). We excluded retrospective studies when participants were selected through a retrospective review of clinical records. We also excluded studies reported only in abstract form or in conference proceedings for which the full text was not available and study authors did not respond to our request to clarify study eligibility.

Participants

Study participants included people with a diagnosis of MCI, based on a decline in memory objectively verified by neuropsychological tests in combination with a precise history, referred by the patient, an informant, or both (Petersen 2004). We included participants with a decline in other cognitive domains and not meeting the criteria for dementia, as defined by the Diagnostic and Statistical Manual of Mental Disorders (American Psychiatric Association 2000; American Psychiatric Association 2013). We included all subtypes of MCI (amnesic single domain, amnesic multiple domain, non-amnesic single domain, non-amnesic multiple domain). Since clinical criteria for the diagnosis of MCI have changed over the past 20 years, we accepted the diagnostic criteria reported by the study authors, for example, a Clinical Dementia Rating (CDR) score of 0.5 (Morris 1993), a Global Deterioration Scale score of 3 (Reisberg 1982), "questionable dementia" (Galton 2005) or "minimal dementia" (Visser 1999), "cognitive impairment, no dementia, as the presence of objective cognitive impairment in any tested domain, with performance falling between the two extremes of normality and dementia" (Graham 1997). We accepted only studies in which the MCI diagnosis was based exclusively on clinical judgement or cognitive test performance, or both. We included all people with MCI for whom clinicians would suspect initial dementia and who would undergo MRI in clinical practice (Differences between protocol and review).

We excluded studies reporting results of MRI on healthy people, or subjective cognitive decline in the absence of objective cognitive dysfunction. We ruled out papers that based MCI definition

on biomarker results. Eventually, in order to avoid participants overlapping, if more studies were performed on the same database (e.g. ADNI, AddNeuroMed) and reported results for the same brain regions, we included only the paper reporting the highest number of participants.

Index tests

We assessed the diagnostic accuracy of structural MRI in detecting atrophy in the whole brain or in specific brain regions: hippocampus, medial temporal lobe, lateral ventricles, entorhinal cortex, medial temporal gyrus, lateral temporal lobe, amygdala, and cortical grey matter.

For the interpretation of the atrophy patterns, Scheltens 1992 and Ten Kate 2017b validated and reported visual rating scales and Frisoni 2017a quantitative volumetric measures. Methods of image quantification vary among research groups and are constantly being refined. A minimum set of MRI criteria for the evaluation of memory clinic patients consists of 3D T1-weighted imaging, fluid-attenuated inversion recovery (FLAIR), turbo-spin or fast-spin T2-weighted images, diffusion-weighted images (DWI) and T2-weighted gradient-recalled echo (GRE) imaging (Vernooij 2019).

We included studies that used either visual assessment or quantitative volumetric measurements, including manually outlining the brain structure and computer-based, semi-automated or automated segmentation methods that allow anatomical identification of areas of the brain. We included studies that used an 'automatic classifier' of MRI data only when accuracy results were based on the volume of individual brain regions. We included any strength of magnetic field, that is, 0.5, 1 or 3 Tesla.

We considered studies only if they reported diagnostic accuracy estimates per number of participants ('participant-level' analysis) and reported data in sufficient detail for construction of 2 x 2 contingency tables.

We excluded studies that reported a single index of MRI accuracy estimate derived from multiple volumetric measures (e.g. multiple regions of interest (ROI)), because of a wide heterogeneity in the number and areas of the brain considered in such studies, or studies that reported MRI-derived index as the spatial pattern of abnormalities for recognition of early Alzheimer's disease (SPARE-AD).

In order to estimate accuracy of a pure volumetric index test, we excluded studies that assessed:

1. a "mixed index test", i.e. comprehensive of both volumetric and cortical thickness measures of the brain
2. an "MRI-derived index", in which volumetric data were summed or divided for other values
3. sub-volumes of brain regions
4. a voxel-based-morphometry (VBM) test which allows to detect important information about regions of atrophy across groups, but cannot provide reliable information about single-subject diagnosis (Teipel 2013).

We excluded studies that used more than one volumetric technique, that is, manual and automated, without reporting separated results per technique. We excluded studies that reported accuracy results of combined MRI test with other

methods to diagnose dementia due to Alzheimer's disease (e.g. neuropsychological tests or genetic data). We did not include longitudinal changes of the volumes of brain regions.

Target conditions

Dementia due to Alzheimer's disease was the target condition. We excluded studies in which the diagnosis of Alzheimer's disease dementia was not the primary outcome of the study and separate data for Alzheimer's disease dementia were not available. We also excluded studies in which findings of the baseline MRI index test formed the basis of selection for the reference standard because this was likely to distort any assessment of the diagnostic value of MRI.

Reference standards

The reference standard for this review was clinical diagnosis of Alzheimer's disease dementia during the follow-up using the NINCDS-ADRDA criteria (McKhann 1984; delayed verification). The gold standard for the diagnosis of Alzheimer's disease dementia is biopsy or autopsy, however this clinical diagnosis represents the best available reference standard in clinical practice. In recent years several biomarkers have been proposed in order to support the diagnosis of dementia due to Alzheimer's disease. In this regard, the updated NIA-AA diagnostic criteria for Alzheimer's disease dementia (McKhann 2011), are an acceptable reference standard if only the Alzheimer's disease-core clinical criteria were used, because they substantially correspond to the NINCDS-ADRDA criteria (McKhann 1984).

Search methods for identification of studies

We developed the search strategy in collaboration with Cochrane Dementia and Cognitive Improvement's Information Specialists, according to recommendations provided in the *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy* (De Vet 2008).

Electronic searches

We searched:

1. Cochrane Dementia and Cognitive Improvement's Specialised Register for diagnostic test accuracy reviews;
2. MEDLINE (Ovid SP; Ovid MEDLINE Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Ovid MEDLINE Daily and Ovid MEDLINE) 1946 to 29 January 2019;
3. Embase (Ovid SP) 1974 to 29 January 2019;
4. BIOSIS Citation Index (ISI Web of Science) 1926 to 29 January 2019;
5. Web of Science Core Collection (ISI Web of Science) 1945 to 29 January 2019;
6. PSYCINFO (Ovid SP) 1806 to 29 January 2019;
7. LILACS (Bireme) 29 January 2019.

See [Appendix 3](#) for a list of the sources searched and the search strategies used.

We did not apply any language restriction to the electronic searches. In addition, we did not use any methodological search filters, aimed to increase specificity, as filters currently published have been shown to lack sensitivity and can therefore miss potentially relevant studies (De Vet 2008; Doust 2005). We

performed the most recent search for this review on 29 January 2019.

Searching other resources

We handsearched reference lists of all relevant publications (retrieved full-texts of key articles and reviews).

Data collection and analysis

Selection of studies

Four teams of two review authors each (GF and GCa; AGB and CL; GC and GL; EC and GCo) independently reviewed titles and abstracts of articles identified by our search to select potentially relevant studies for inclusion. If the study eligibility was unclear from the abstract, or if no abstract was available but the title suggested a potentially relevant study, we obtained the full-text of the article to assess eligibility for inclusion on the basis of criteria listed above under [Criteria for considering studies for this review](#). We also screened the reference lists of systematic reviews and included articles to identify any studies missed by the electronic database search. We prepared a manual to assist review authors in the selection of studies. We solved any disagreement in selection of abstracts or full-text articles by discussion. We stored all abstracts and full-text articles in a database designed for the review. When articles reported on a cohort (same database) that overlapped with a cohort in another paper, we used the study with the higher sample size. We included articles reporting results on different MRI techniques on the same study population or separate results for relevant participants' subgroups. If we included more than one article from the same study authors, we assessed the absence of overlap by using the reported recruitment periods or directly contacting the study authors to clarify study eligibility. For excluded studies, we documented reasons for exclusion ([Characteristics of excluded studies](#)).

Data extraction and management

Two teams of two review authors each (AGB and EC; GL and GC) independently extracted data and solved disagreements by consensus. If required, we contacted study authors for missing data. We designed for the review a data extraction form and pilot-tested it on five studies. The extraction form was uploaded in a Microsoft Access 2003 database. Review authors who extracted data were not blind to publishing journal, names of study authors, and institutions. We prepared a manual to assist review authors in data extraction and management. We extracted the following data from eligible studies.

1. Study characteristics: identity number (ID), first author, country, language, year of publication, journal name, additional bibliographic references linked to the study
2. Characteristics of study participants: multicentric study (item A0), inclusion and exclusion criteria (item A1 to A11), clinical characteristics (item A12 to A22), co-pathologies and treatments (item A23 to A25)
3. Features of the index test (item B1 to B5)
4. Features of the reference standard, including the follow-up length (item C1 to C8)
5. Data tables and missing data (item D1-D19)
6. Numbers of true positives (TP), false negatives (FN), true negatives (TN) and false positives (FP) were used to construct a 2

x 2 table for the index test. If studies did not report these values, we contacted the study authors or attempted to reconstruct 2 x 2 tables from the accuracy estimates reported in the article.

7. Notes

Assessment of methodological quality

We used QUADAS-2, a modified version of the QUADAS (Quality Assessment of Diagnostic Accuracy Studies) tool, to assess the methodological quality of each included study (Whiting 2011). We have presented the review-specific QUADAS-2 tool and an explanatory document in Appendix 4. We judged each paper as having a 'low', 'high' or 'unclear' risk of bias for each of the following four domains: patient selection; index test; reference standard; flow and timing. We assessed concerns about applicability in three domains: patient characteristics and setting; index test; reference standard. We judged low-quality studies as having high or unclear risk of bias in at least one QUADAS-2 domain. Two review authors (GL and GC) independently assessed each included study and solved disagreement by reaching consensus. Any disagreement that could not be solved by consensus was referred to a third author (GF).

Statistical analysis and data synthesis

We used data from 2 x 2 tables of structural MRI performance (TP, FN, FP, TN) to summarise accuracy estimates of each primary study. We estimated sensitivity, specificity, positive and negative likelihood ratios (LR+ and LR-), with their 95% confidence intervals (CI). We provided graphical representation of the studies by plotting sensitivity and specificity estimates with their 95% CIs in both a forest plot and a receiver operating characteristic (ROC) space. We used the hierarchical summary ROC curve (HSROC) model proposed by Rutter and Gatsonis (Rutter 2001) and in chapter 10 of the *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy* (Macaskill 2010), to estimate pooled accuracy measures in the absence of specified thresholds as well as to investigate relative diagnostic odds ratios (DORs) in subgroup analyses (assuming parallel ROC curves in logits). We used this technique to plot the summary ROC curve, and we also calculated pooled point estimates of sensitivity and specificity, since we found studies yielded heterogeneous estimates and no clear threshold effects were apparent both graphically and statistically in analyses with more data. We used the *metadas* user-written command in Statistical Analysis System (SAS) (version 9.4. SAS Institute Inc., Cary, NC, USA) statistical package for the analyses (Takwoingi 2010).

Very few studies reported MRI data extracted with both manual and automated methods and we decided post-hoc to use manual methods in order to be consistent with the majority of the studies.

If an individual study reported results for more than one follow-up period, we reported accuracy estimates for all the periods, but selected just the estimate from the three years' (or longer) follow-up for the meta-analysis. This choice is based on the assumption that the conversion rate to Alzheimer's disease dementia is higher in the first few years following MCI diagnosis and declines thereafter, and that short-term MRI accuracy is therefore the most relevant information for patients and clinicians. Moreover, most of the data were available at two or three years of follow-up and very few studies reported a follow-up period of more than three years.

If estimates of sensitivity and specificity or sufficient data to construct a 2 x 2 table of test performance were not available,

we wrote to the authors of the primary study requesting the individual participant data. If we received the individual participants' data, we calculated the estimates of sensitivity and specificity corresponding to the threshold nearer to the upper left point of the ROC curve. We were aware that this data-driven method for threshold selection could lead to an overestimate of diagnostic accuracy (Leeflang 2008). However, there are no accepted thresholds to a priori define a positive MRI, and published accuracy estimates are likely to be based on data-driven threshold selection.

If the primary study authors did not provide data (e.g. we were not able to locate contact details of study authors, we received no reply from study authors, study authors replied that the requested information was unavailable), we excluded the study from the review.

Investigations of heterogeneity

We initially assessed heterogeneity by visually examining forest plots of sensitivities and specificities and ROC plots. We planned to formally explore heterogeneity by a likelihood ratio test comparing the model without covariate with the model including the test type covariate. We stated potential sources of heterogeneity under *Secondary objectives*.

Sensitivity analyses

We planned to conduct sensitivity analyses to assess the impact of the methodological quality of included studies on MRI accuracy estimates excluding studies at high risk of bias (see *Assessment of methodological quality*). However, we were not able to do this because we judged almost all studies at high risk of bias.

Thus we decided to perform specific analysis according to the brain region and considering MRI techniques, duration of follow-up length and age of participants as covariate.

'Summary of findings' table

We presented the main results of the review in a 'Summary of findings' table, according to recommendations described in Chapter 11 of the *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy* (version 0.9; Bossuyt 2013). We graded the quality of evidence according to the GRADE system for diagnostic tests, considering study limitations (risk of bias), indirectness, inconsistency, imprecision, and risk of publication bias (Schunemann 2008; Schunemann 2016). According to the software GRADEpro GDT, we assigned four levels of quality of evidence: high, moderate, low, and very low.

RESULTS

Results of the search

A flow chart describes the results of the selection process (Figure 1). The literature search identified 29,335 references. We screened titles and abstracts to exclude duplicates (n = 5064) and irrelevant studies (n = 23,962). We retrieved the full texts of the remaining 309 references and assessed them for eligibility. Ultimately, 33 studies that were eligible according to the inclusion criteria provided data for the review; we excluded 276 studies. We reported the list and descriptions of excluded studies under *Characteristics of excluded studies*. For 112 studies the index test was outside the inclusion criteria, reporting data for a combination of multiple

volumetric measures, or a test comprehensive of both volumetric and cortical thickness measures of the brain, or a voxel-based-morphometry test. We excluded an additional 43 studies as they were of retrospective, case-control, or cross-sectional design. Thirty-one studies were not diagnostic test accuracy studies and focused on technical aspects of the test. We excluded another 24 studies as they reported on a cohort that overlapped with a cohort in another included paper. We excluded 23 studies as they enrolled healthy participants or participants with dementia. Twenty studies presented insufficient descriptions of study results needed to construct 2 x 2 tables and we were unable to contact study authors. We could not extract data for 2 x 2 tables from 14

studies and authors did not reply to our request ([Frisoni 2010a \[pers comm\]](#); [Frisoni 2010d \[pers comm\]](#); [Frisoni 2010f \[pers comm\]](#); [Frisoni 2010g \[pers comm\]](#); [Frisoni 2010k \[pers comm\]](#); [Frisoni 2010m \[pers comm\]](#); [Frisoni 2010n \[pers comm\]](#); [Frisoni 2010o \[pers comm\]](#); [Frisoni 2016a \[pers comm\]](#); [Frisoni 2016b \[pers comm\]](#); [Frisoni 2016c \[pers comm\]](#); [Frisoni 2016e \[pers comm\]](#); [Frisoni 2016f \[pers comm\]](#)) or answered but provided no information ([Frisoni 2012 \[pers comm\]](#)). In four excluded studies, the reference standard was outside the inclusion criteria. We excluded four unpublished studies. We excluded one additional study that reported outcomes for number of MRI, not number of participants.

Figure 1. Figure 1. Flow of studies identified in literature search for systematic review on structural magnetic resonance imaging for an early diagnosis of dementia due to Alzheimer's disease in people with mild cognitive impairment

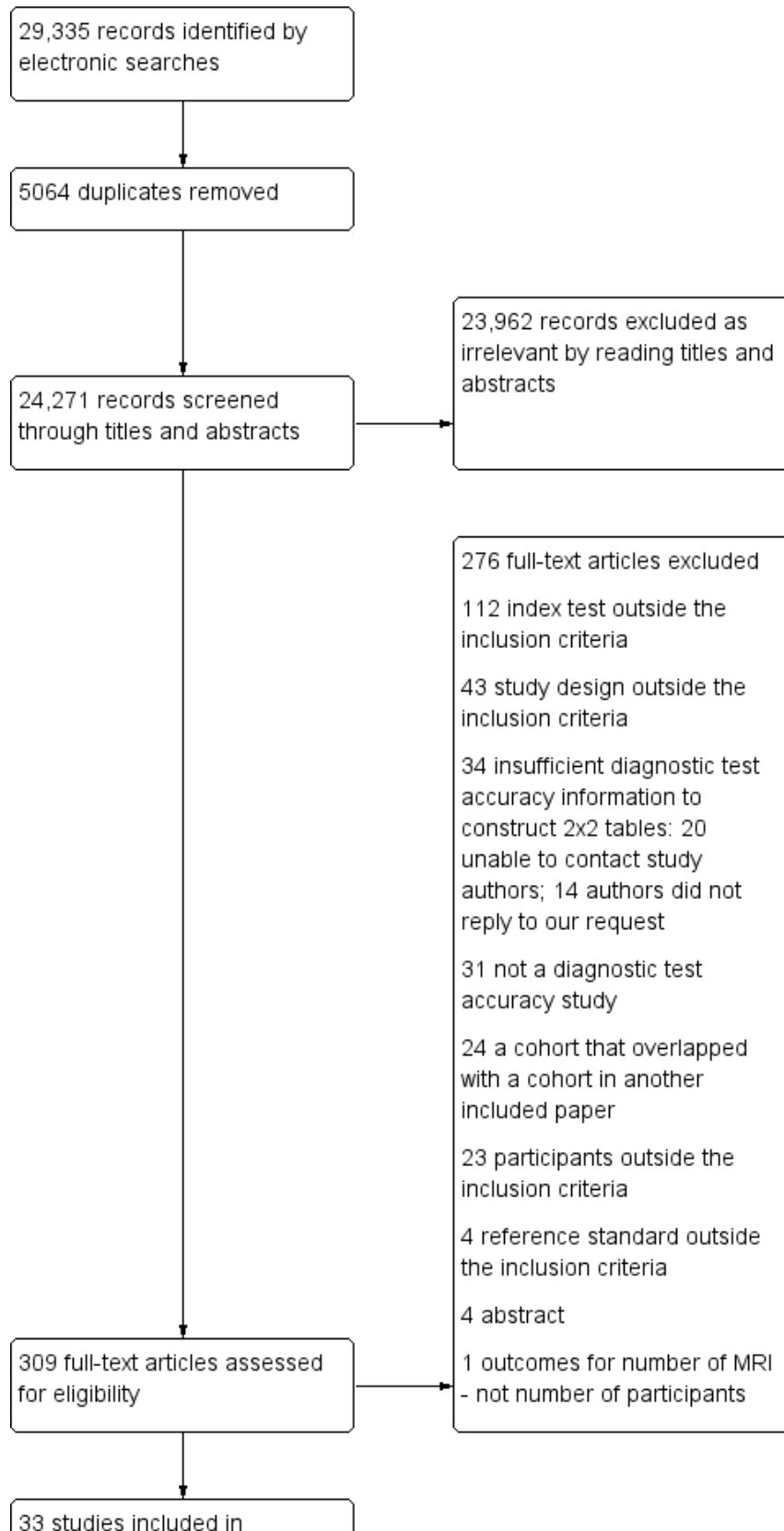


Figure 1. (Continued)

↑
33 studies included in quantitative synthesis

We reported the list and details of the included studies under [Characteristics of included studies](#) and [Table 1](#). The 33 included studies involved 3935 participants, with a median of 43 participants per study (range 13 to 480). All the included studies were conducted at tertiary referral centres and 16 (48%) were multicentric. Nineteen studies were conducted in Europe, nine in North America, three in North America and Europe, one in Taiwan and one in Australia. The articles were published from 1999 to 2019. The median proportion of women was 53% (range 26% to 71%), and the mean age of participants ranged from 63 to 87 years (median 73 years). At baseline, participants had a median Mini Mental State Examination (MMSE) score of 27 (range 22 to 29) and a mean level of education (years of schooling) of 12 years (range 5 to 16 years). Of the 3935 participants, 1341 (34%) progressed to Alzheimer's disease dementia and 2594 (66%) did not progress. Of the participants who did not progress to Alzheimer's disease dementia, 2561 (99%) remained stable MCI and 33 (1%) converted to other types of dementia. The percentages of participants who remained stable with MCI and those who converted to other types of dementia varied among the included studies from 31% to 81%, and from 1% to 19% respectively. The mean length of follow-up ranged from 1 to 7.6 years (median 2 years) and the percentages of participants who progressed to Alzheimer's disease dementia during follow-up ranged from 19% to 69%. All included studies reported accuracy estimates for one follow-up period, except [Gaser 2013](#), who reported data at one and three years' follow-up. We included data from the three years' follow-up.

During the literature screening we assessed the eligibility of several Alzheimer's Disease Neuroimaging Initiative (ADNI) papers. ADNI is a multicenter project ongoing in 50 medical centres and university sites across the USA and Canada. The primary objective of ADNI is to collect, validate and utilise data, acquired serially over two to three years of follow-up, including structural MRI and positron emission tomography (PET) images, genetic data, cognitive tests, cerebrospinal fluid (CSF) and blood biomarkers as predictors of Alzheimer's disease dementia. In order to avoid participants overlapping, we excluded all studies performed on the ADNI database in the same period and focusing on the same brain region ([Table 2](#)). Thus, we included ADNI studies with the larger sample size. Among the 33 included papers, we identified seven eligible ADNI studies, from which we extracted sixteen 2 x 2 contingency tables ([Gaser 2013](#); [Jang 2018](#); [Khan 2015](#); [Ledig 2018](#); [Pereira 2014](#); [Prestia 2013 \(ADNI\)](#); [Wolz 2011](#)). We also applied the same selection criteria to studies performed on other databases belonging to DESCRIPA study (Development of Screening Guidelines and Clinical Criteria for Predementia Alzheimer's disease), VUmc study (University Medical Center Amsterdam) and AddNeuroMed study. Review authors focused on one to eight different brain regions (hippocampus, entorhinal cortex, amygdala, medial temporal lobe, lateral temporal lobe, lateral ventricles, medial temporal gyrus, cortical grey matter) or whole brain, using several MRI techniques.

In response to our request ([Frisoni 2010b \[pers comm\]](#); [Frisoni 2010c \[pers comm\]](#); [Frisoni 2010e \[pers comm\]](#); [Frisoni 2010h \[pers comm\]](#); [Frisoni 2010i \[pers comm\]](#); [Frisoni 2010j \[pers comm\]](#); [Frisoni 2010l \[pers comm\]](#); [Frisoni 2010p \[pers comm\]](#); [Frisoni 2010q \[pers comm\]](#); [Frisoni 2010r \[pers comm\]](#); [Frisoni 2010s \[pers comm\]](#); [Frisoni 2016d \[pers comm\]](#); [Frisoni 2017b \[pers comm\]](#)), the authors of 12 included studies sent us the data needed to complete 2 x 2 tables ([Carmichael 2007](#); [Caroli 2007](#); [deToledo-Morell 2004](#); [Devanand 2007](#); [Eckerstrom 2008](#); [Erten-Lyons 2006](#); [Herukka 2008](#); [Jack 2000](#); [Prestia 2013](#); [VanderFlier 2005](#); [Visser 2002](#); [Wang 2006](#)).

Fifteen studies analysed one brain region, six studies considered two regions, and the remaining studies considered three or more than three regions ([Table 3](#)). Twenty-four studies measured volume of brain regions with quantitative manual or automated methods, four studies used visual and quantitative methods, five studies used only the visual method ([Table 3](#)). Studies generally measured the volume of the hippocampus and the entorhinal cortex with a quantitative manual method, whereas they mainly used a visual method based on the Scheltens scale ([Scheltens 1992](#); [Scheltens 1997](#)), to measure medial temporal lobe atrophy. The choice of a threshold value for the medial temporal lobe atrophy was heterogeneous between studies. Three study authors ([Caroli 2007](#); [Clersx 2013a](#); [Visser 2002](#)), did not specify a cut-off value, while [Pereira 2014](#) and [Rhodius-Meester 2016](#) used an averaged left and right medial temporal lobe cut-off value of 1.5 or more. One study ([Monge Argilés 2014](#)), considered a cut-off based on the sum of left and right medial temporal lobe atrophy scores (≥ 3.0). The most used software for the manual assessment of the brain region volume was DISPLAY ([Caroli 2007](#); [Prestia 2013 \(ADNI\)](#); [VanderFlier 2005](#)) and for the automated assessment it was Freesurfer ([Gaser 2013](#); [Khan 2015](#); [Nesteruk 2016](#); [Prestia 2013](#); [Prieto del Val 2016](#)). Some study authors used software developed in house. The main manufacturers of MRI scanners used in the included studies were Philips, Siemens and General Electric (29 of the 33 included studies). Several studies, such as ADNI studies, used all three manufacturers. Two studies used Toshiba and Technicare; two studies did not report manufacturers. Twenty-six (79%) of the included studies performed the MRI at 1.5 Tesla, one study at 3.0 Tesla ([Jang 2018](#)), and two studies at 0.5 Tesla ([Eckerstrom 2008](#); [Eckerstrom 2013](#)). Only [Visser 1999](#) used MRI at 0.6 Tesla. One study did not report this information.

The majority of the included studies used the NINCDS ADRA criteria as a reference standard ([McKhann 1984](#)). Three studies ([Nesteruk 2016](#); [Rhodius-Meester 2016](#); [Wood 2016](#)), used the NIA-AA diagnostic criteria ([McKhann 2011](#)).

Methodological quality of included studies

We present the details on the quality of included studies in the QUADAS-2 results summary ([Figure 2](#)). We judged all studies as low

quality because we rated them all as having at least one domain with high or unclear risk of bias.

Figure 2. Risk of bias and applicability concerns summary: review authors' judgements about each domain for each included study

	Risk of Bias				Applicability Concerns		
	Patient Selection	Index Test	Reference Standard	Flow and Timing	Patient Selection	Index Test	Reference Standard
Carmichael 2007	+	+	+	+	+	+	+
Caroli 2007	+	+	+	+	+	+	+
Clerx 2013a	+	+	+	+	+	+	+
deToledo-Morell 2004	+	+	+	+	+	+	+
Devanand 2007	+	+	+	+	+	+	+
Eckerstrom 2008	+	+	+	+	+	+	+
Eckerstrom 2013	+	+	+	?	+	+	+
Erten-Lyons 2006	+	+	?	?	+	+	?
Frolich 2017	+	+	+	+	+	+	+
Galton 2005	+	+	+	+	+	+	+
Gaser 2013	+	+	+	+	+	+	+
Herukka 2008	+	+	+	+	+	+	+
Jack 2000	+	+	+	+	+	+	+
Jang 2018	+	+	+	+	+	+	+
Khan 2015	+	?	+	+	+	+	+
Ledig 2018	+	+	+	+	+	+	+
Liu 2010	+	+	+	+	+	+	+
Monge Argilés 2014	+	+	+	+	+	+	+
Nesteruk 2016	+	+	?	+	+	+	?
Ong 2015	+	?	+	+	+	+	+
Pereira 2014	+	+	+	+	+	+	+
Platero 2019	+	+	+	+	+	+	+
Prestia 2013	+	?	+	+	+	+	+
Prestia 2013 (ADNI)	+	?	?	+	+	+	+
Prieto del Val 2016	+	+	+	+	+	+	+
Rhodijs-Meester 2016	+	?	+	+	+	+	+

Figure 2. (Continued)

Rhodijs-Meester 2016	⊖	?	+	+	+	+	+	+
VanderFlier 2005	+	⊖	+	⊖	+	+	+	+
Visser 1999	?	⊖	+	⊖	+	+	+	+
Visser 2002	⊖	⊖	+	⊖	+	+	+	+
Wang 2006	⊖	⊖	+	+	+	+	+	+
Westman 2011	⊖	⊖	+	+	+	+	+	+
Wolz 2011	⊖	?	+	+	+	+	+	+
Wood 2016	⊖	⊖	?	+	+	+	+	?

⊖ High ? Unclear ⊕ Low

Participant selection

Only one study (VanderFlier 2005), demonstrated low risk of participant selection bias, we judged one study (Visser 1999), at unclear risk of bias, and the other included studies demonstrated high risk. Non-consecutive enrolment or the use of registry data, such as in ADNI or AddNeuroMed studies, which, despite being prospective registries, imposed specific participant selection criteria, such as the availability of multiple biomarkers, were the main reasons for assessment of high risk of bias. Moreover, in prospective registries it was unclear if inappropriate exclusions (e.g. depression, vascular lesions on MRI) were avoided. Absence of a clear definition of inclusion and exclusion criteria was the reason for assessment of high risk of bias in other included cohort studies.

Index test

We judged 24 (73%) studies at high risk of bias, six (18%) at unclear risk, and three (9%) (Galton 2005; Monge Argilés 2014; Pereira 2014), at low risk of bias for this domain. Overall, 30 studies did not provide sufficient details on the index test, either because of a lack of a clear, pre-specified definition of what was considered to be a 'positive' result of the MRI or lack of blinding of radiologists to the reference standard, or both. Twenty-four studies had unclear criteria for a positive MRI result. Only nine studies reported a clear, pre-specified definition of a positive MRI result (Galton 2005; Khan 2015; Monge Argilés 2014; Ong 2015; Pereira 2014; Prestia 2013; Prestia 2013 (ADNI); Rhodijs-Meester 2016; Wolz 2011).

We specified a threshold mainly for MRI assessed by the visual method and the automated method, and chose a cut-off value for the visual method according to the Scheltens scale (Scheltens 1992; Scheltens 1997), but cut-offs were set at different levels between studies, for example, a cut-off of 1.5 or higher, based on the mean medial temporal lobe atrophy scores of both hemispheres (Pereira 2014; Rhodijs-Meester 2016), or a cut-off of 3 or higher, based on the sum of right and left medial temporal lobe atrophy (Monge Argilés 2014). Clerx 2013a, Frolich 2017, Prieto del Val 2016 and Wood 2016 used criteria for a positive manual MRI test based on the Youden's index, which has the advantage of being a single

measure, but it loses the distinction between false positives and false negatives (Hilden 1996).

Blinding of radiologists to the clinical diagnosis of Alzheimer's disease dementia was unclear in 18 (54%) studies. To evaluate blinding, we applied the same criteria for visual, manual, or automated methods. However, we acknowledge that the absence of blinding of radiologist or the interpretation of MRI by different radiologists for different participants may be considered less severe when the automated method was used rather than the visual or manual methods. In three studies (Clerx 2013a; Erten-Lyons 2006; Jang 2018), more than one radiologist interpreted MRI scans for different participants, whereas in sixteen studies it was unclear if one or more radiologists interpreted MRI separately or in a joint session. Thirteen studies assessed interobserver or intraobserver variability in the whole cohort or in a subset of randomly selected participants (Clerx 2013a; deToledo-Morell 2004; Devanand 2007; Eckerstrom 2008; Eckerstrom 2013; Erten-Lyons 2006; Herukka 2008; Jang 2018; Monge Argilés 2014; Rhodijs-Meester 2016; VanderFlier 2005; Visser 2002; Westman 2011).

Reference standard

Twenty-nine (88%) studies were at low risk of bias in the 'reference standard' domain and we classified four as unclear risk. We judged one study (Prestia 2013 (ADNI)), at unclear risk of bias because baseline biomarker results of participants were available to clinicians who diagnosed Alzheimer's disease dementia. As specified in the Methods section, we accepted the new diagnostic criteria for dementia due to Alzheimer's disease (McKhaan 2011), if only the Alzheimer's disease core clinical criteria were used, as in Rhodijs-Meester 2016. When this information was not available, we judged the included study at unclear risk of bias and unclear concern about the incorporation of MRI into the diagnosis of Alzheimer's disease during follow-up (incorporation risk) (Nesteruk 2016; Wood 2016). We judged one study (Erten-Lyons 2006), at unclear risk of bias because of insufficient information regarding the reference standard.

Flow and timing

Nine (27%) studies (Carmichael 2007; Caroli 2007; Devanand 2007; Galton 2005; Platero 2019; Prestia 2013 (ADNI); VanderFlier 2005; Visser 1999; Visser 2002), were at high risk of bias in the 'flow and timing' domain, two (6%) were at unclear risk, and 22 (67%) were at low risk. We classified a study as having high risk of bias when study authors did not adequately explain withdrawals or losses to follow-up, or study authors excluded from the analysis participants who progressed to non-Alzheimer's disease dementia. Erten-Lyons 2006 did not report if all participants received the same reference standard and we judged it at unclear risk. Eckerstrom 2013 was at unclear risk of bias because it did not specify if non-Alzheimer's disease dementia cases were included in the analysis. The median interval between MRI test and reference standard was two years (range 1 to 7.6 years).

Concerns regarding applicability

We had no concerns for any studies about applicability in the 'patient selection' and 'index test' domains. Participants and the index test in the included studies did not differ from those targeted by the review question. Three studies (Erten-Lyons 2006; Nesteruk 2016; Wood 2016), demonstrated unclear concern for the 'reference standard' domain and the remaining 30 studies demonstrated low concern.

Findings

We have presented findings under five main brain regions (Summary of findings 1).

Hippocampus

Twenty-two studies (Caroli 2007; Clerx 2013a; deToledo-Morell 2004; Devanand 2007; Eckerstrom 2008; Erten-Lyons 2006; Frolich 2017; Herukka 2008; Jack 2000; Khan 2015; Liu 2010; Ong 2015; Platero 2019; Prestia 2013; Prestia 2013 (ADNI); VanderFlier 2005; Visser 1999; Visser 2002; Wang 2006; Westman 2011; Wolz 2011; Wood 2016), which included a total of 2209 participants (687 (31%) of whom progressed to Alzheimer's disease dementia), measured the total hippocampal volume. The studies used different MRI techniques: manual (11 studies, 512 participants); semiautomatic or automatic (nine studies, 1334 participants); manual and semiautomatic or automatic (two studies; 421 participants). The overall sample size ranged from 13 participants (Visser 1999), to 447 participants (Khan 2015). Sensitivity ranged from 0.28 to 1.00, while specificity ranged from 0.43 to 0.94. Forest plots demonstrated a high degree of heterogeneity and wide confidence intervals for estimates of both sensitivity and specificity between the included studies (Figure 3). Two studies used two techniques for total hippocampal volume, thus we chose the manual technique for these studies (Clerx 2013a; Prestia 2013 (ADNI)), for consistency with other studies. The mean sensitivity and specificity (summary operating point) were 0.73 (95% CI 0.64 to 0.80) and 0.71 (95% CI 0.65 to 0.77) respectively (Figure 4). Positive likelihood ratio was 2.53 (95% CI 2.09 to 3.06) while negative likelihood ratio was 0.38 (95% CI 0.29 to 0.50). The certainty of the evidence (Summary of findings 1) was low for both sensitivity and specificity due to risk of bias (-1) and inconsistency due to heterogeneous study results (-1).

Figure 3. Forest plot of total hippocampal volume measured by structural magnetic resonance imaging (MRI) for early diagnosis of dementia due to Alzheimer's disease in people with mild cognitive impairment. Plot shows study-specific estimates of sensitivity and specificity (squares) with 95% confidence interval (black line) and study. Studies are ordered according to the estimates of sensitivity. TP: true positive; FP: false positive; FN: false negative; TN: true negative

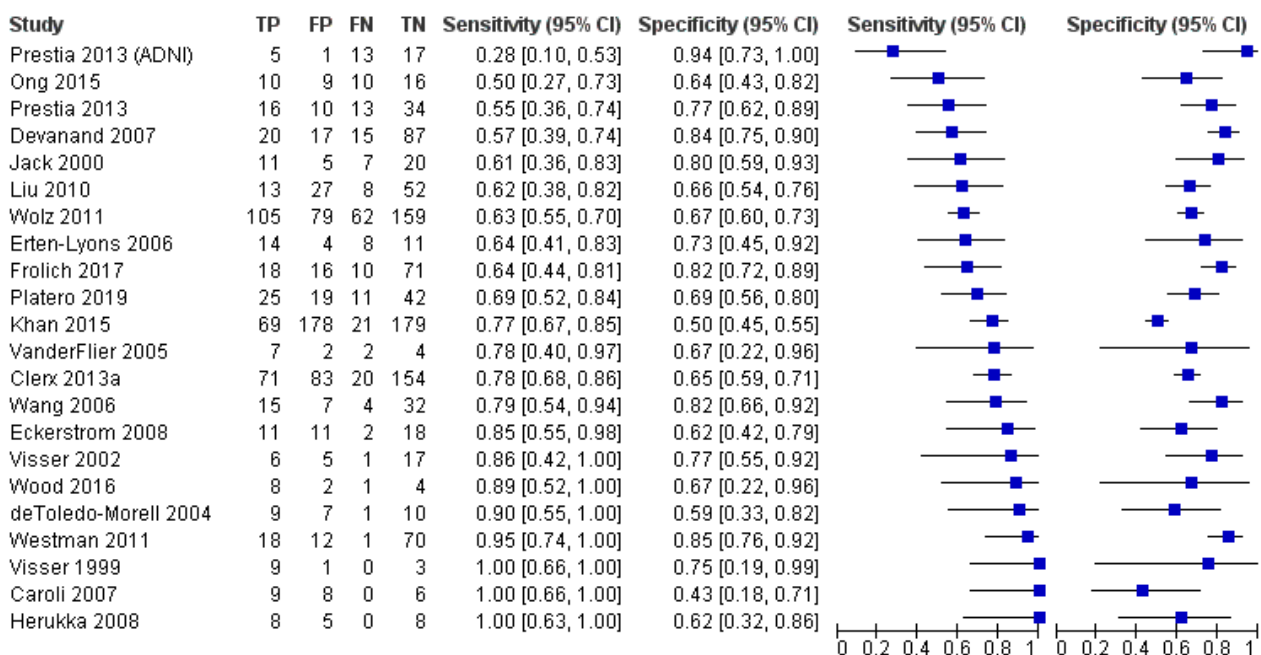
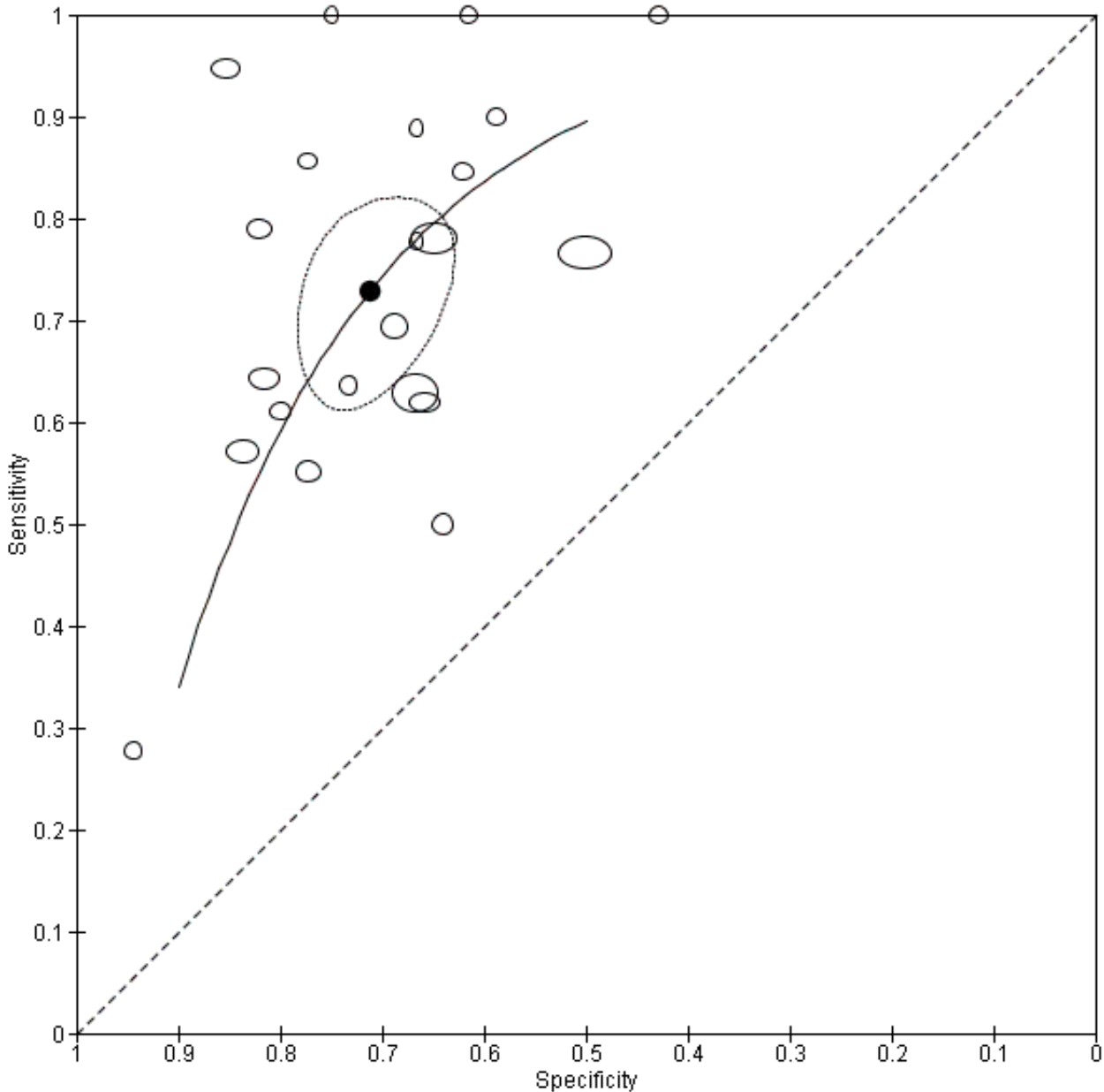


Figure 4. Summary receiver operating characteristic (ROC) plot of total hippocampus volume measured by structural magnetic resonance imaging (MRI) for early diagnosis of dementia due to Alzheimer's disease in people with mild cognitive impairment. Each point represents the pair of sensitivity and specificity from a study. The solid black circle represents the pooled sensitivity and specificity, which is surrounded by a 95% confidence region (dashed line)



Investigation of heterogeneity

Potential sources of heterogeneity are outlined under the section [Secondary objectives](#). We were able to assess the impact of heterogeneity on the MRI diagnostic accuracy for MRI technique, duration of follow-up and age of participants. We have reported the numbers of participants in the subgroup analyses in [Table 4](#). Because very few studies reached a follow-up of more than three years, we used the cut-off value for subgroup analyses on follow-up

time at 'less than three years' versus 'at least three years'. All these comparisons were between studies or indirect.

Sensitivity was 0.82 (95% CI 0.69 to 0.90) for manual technique (13 studies) and 0.59 (95% CI 0.48 to 0.70) for automatic or semiautomatic technique (11 studies); specificity was 0.74 (95% CI 0.67 to 0.81) and 0.66 (95% CI 0.56 to 0.74), respectively ([Table 4](#)). The relative DOR was 4.83 (95% CI 1.82 to 12.8), suggesting better accuracy with the manual compared with the automatic technique.

Sensitivity was 0.74 (95% CI 0.65 to 0.81) for less than three years' follow-up (14 studies), and 0.71 (95% CI 0.54 to 0.84) for at least three years' follow-up (8 studies); specificity was 0.69 (95% CI 0.61 to 0.76) and 0.76 (95% CI 0.67 to 0.82), respectively. No difference in accuracy was found (relative DOR 0.80, 95% CI 0.34 to 1.92) for the longer versus shorter follow-up).

Sensitivity was 0.72 (95% CI 0.54 to 0.84) for studies including patients with a mean age of less than 70 years (6 studies), and 0.73 (95% CI 0.64 to 0.81) for a mean age of 70 years or more (16 studies); specificity was 0.77 (95% CI 0.67 to 0.84) and 0.69 (95% CI 0.62 to 0.75), respectively. We found no difference in accuracy (relative DOR 1.38, 95% CI 0.57 to 3.31) for younger versus older age.

We were not able to explore the effects of amnestic versus non-amnestic MCI, medial temporal lobe versus other structures, hippocampus versus other structures, entorhinal cortex versus other structures, or temporoparietal regions versus other structures, as potential sources of heterogeneity, because studies providing these data were too few to make this a meaningful analysis. Furthermore, we performed no assessment of heterogeneity for setting and MRI Tesla because all included participants had been referred to tertiary centres and the majority of the included studies used a magnetic field strength of 1.5 Tesla.

Direct comparisons of left and right hippocampal volumes

Seven studies, including 298 participants, made a direct comparison of the left and right hippocampal volumes that were

measured by manual technique (Caroli 2007; Devanand 2007; Eckerstrom 2013; Herukka 2008; VanderFlier 2005; Wang 2006) or automatic technique (Nesteruk 2016). Galton 2005, including 29 participants, used a visual technique and found sensitivity and specificity of 0.64 (95% CI 0.31 to 0.89) and 0.89 (95% CI 0.65 to 0.99) for the right hippocampus and 0.91 (95% CI 0.59 to 1.00) and 0.89 (95% CI 0.65 to 0.99) for the left hippocampus, respectively. We considered that the visual technique should not be pooled with quantitative manual or automatic techniques and excluded this study from direct comparisons of right versus left hippocampus, as follows.

1. Left hippocampal volume: sensitivity varied from 0.44 to 0.89, while specificity varied from 0.64 to 1.00 (Data table 2; Data table 3). The mean sensitivity and specificity (summary operating point) were, respectively, 0.71 (95% CI 0.62 to 0.79) and 0.76 (95% CI 0.67 to 0.83; Figure 5). Positive likelihood ratio was 2.95 (95% CI 2.14 to 4.06) while negative likelihood ratio was 0.38 (95% CI 0.28 to 0.51).
2. Right hippocampal volume: sensitivity ranged from 0.61 to 1.00, specificity from 0.43 to 0.81 (Figure 6). The mean sensitivity and specificity were, respectively, 0.81 (95% CI 0.73 to 0.88) and 0.71 (95% CI 0.61 to 0.80; Figure 5). Positive likelihood ratio was 2.82 (95% CI 2.01 to 3.96) while negative likelihood ratio was 0.23 (95% CI 0.11 to 0.46).

Figure 5. Summary receiver operating characteristic curve (ROC) presenting direct comparisons of hippocampus left and hippocampus right

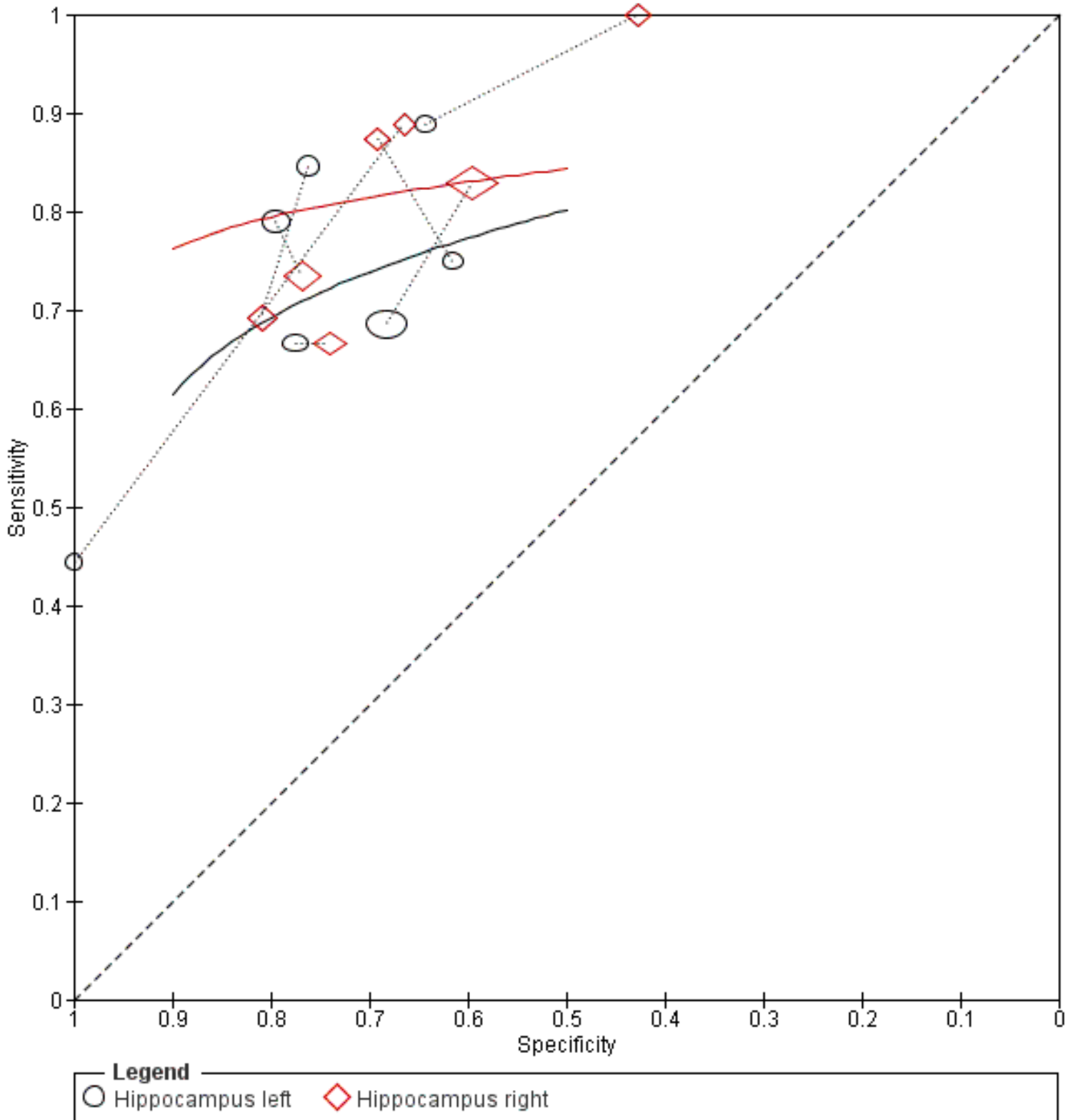


Figure 6. Summary receiver operating characteristic (ROC) plot of total medial temporal lobe volume measured by structural MRI for early diagnosis of dementia due to Alzheimer's disease in people with mild cognitive impairment. Each point represents the pair of sensitivity and specificity from a study. The solid black circle represents the pooled sensitivity and specificity, which is surrounded by a 95% confidence region (dashed line)

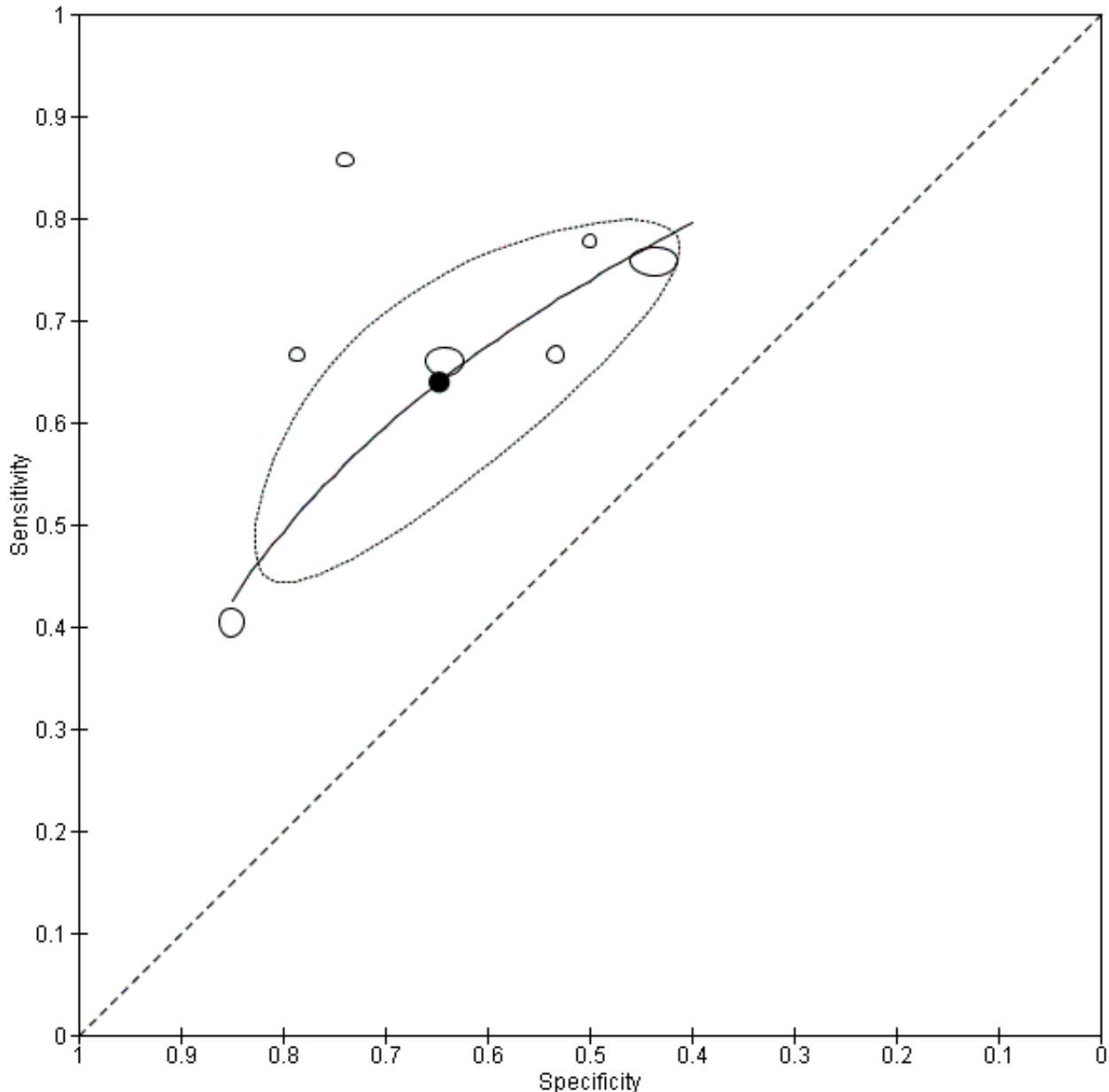


Figure 5 shows the paired ROC plot with studies directly comparing right versus left hippocampus. The relative DOR suggested no overall difference in accuracy (1.37, 95% CI 0.63 to 2.99).

Medial temporal lobe

Seven studies (Caroli 2007; Clerx 2013a; Monge Argilés 2014; Pereira 2014; Rhodius-Meester 2016; VanderFlier 2005; Visser 2002), assessed the volume of the total medial temporal lobe for a total number of 1077 participants (330 (31%) of whom progressed to Alzheimer's disease dementia; Data table 4). Six studies used a

visual method and one (VanderFlier 2005), a quantitative manual method. The smallest study (VanderFlier 2005), recruited 15 participants, while the largest study (Pereira 2014), enrolled 480 participants. Sensitivities and specificities ranged from 0.40 to 0.86 and from 0.44 to 0.85, respectively. The mean sensitivity and specificity (summary operating point) were 0.64 (95% CI 0.53 to 0.73) and 0.65 (95% CI 0.51 to 0.76), respectively (Figure 6). Positive likelihood ratio was 1.81 (95% CI 1.41 to 2.32) while negative likelihood ratio was 0.56 (95% CI 0.46 to 0.67). The certainty of the evidence (Summary of findings 1) was moderate for both sensitivity

and specificity due to risk of bias (-1), but we did not downgrade for imprecision since confidence intervals were large, but still their upper limit was below 0.75 for both sensitivity and specificity, which is a modest performance.

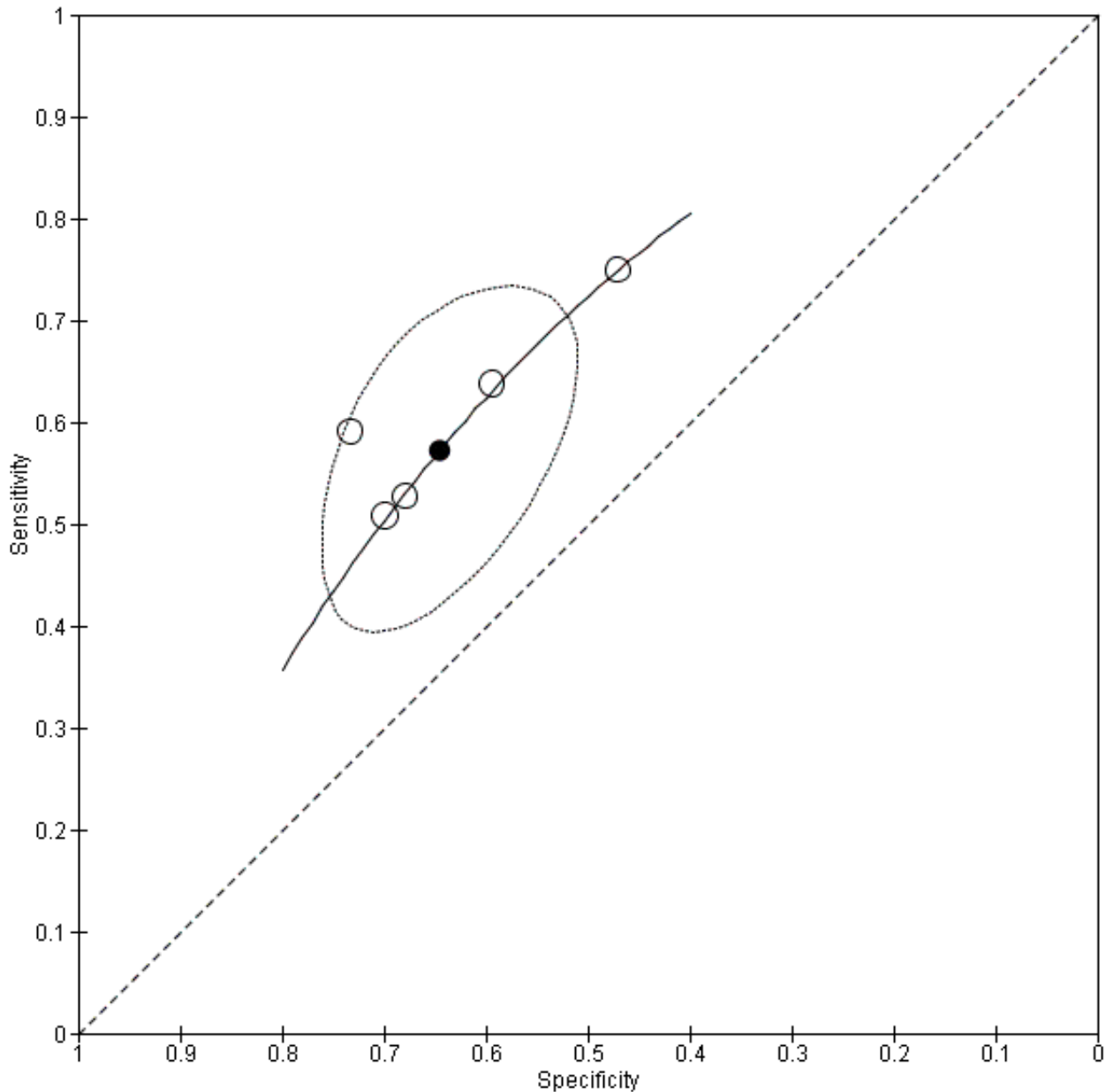
[VanderFlier 2005](#) analysed separately left and right medial temporal lobe and reported sensitivity and specificity of 0.89 (95% CI 0.52 to 1.00) and 0.33 (95% CI 0.04 to 0.78) for the left lobe, 0.22 (95% CI 0.03 to 0.60) and 1.00 (95% CI 0.54 to 1.00) for the right lobe ([Data table 5](#); [Data table 6](#)).

Lateral ventricles

Five studies ([Carmichael 2007](#); [Clerx 2013a](#); [Erten-Lyons 2006](#); [Jang 2018](#); [Ledig 2018](#)), measured the volume of the lateral ventricles for a total of 1077 participants (371 (34%) of whom progressed

to Alzheimer's disease dementia; [Data table 7](#)). Four studies used an automatic or semi-automatic technique and one study ([Jang 2018](#)), used a visual method. The smallest study ([Carmichael 2007](#)), recruited 29 participants and the largest study ([Ledig 2018](#)), recruited 343 participants. Sensitivities and specificities ranged from 0.51 to 0.75 and from 0.47 to 0.73, respectively. The mean sensitivity and specificity (summary operating point) were 0.57 (95% CI 0.49 to 0.65) and 0.64 (95% CI 0.59 to 0.70) respectively ([Figure 7](#)). Positive likelihood ratio was 1.61 (95% CI 1.39 to 1.87) while negative likelihood ratio was 0.66 (95% CI 0.57 to 0.78). The certainty of the evidence ([Summary of findings 1](#)) was moderate for both sensitivity and specificity due to risk of bias (-1), but was not downgraded for imprecision since confidence intervals were large, but still their upper limit was below 0.75 for both sensitivity and specificity, which is a modest performance.

Figure 7. Summary receiver operating characteristic (ROC) plot of volume of lateral ventricles measured by structural magnetic resonance imaging (MRI) for early diagnosis of dementia due to Alzheimer's disease in people with mild cognitive impairment. Each point represents the pair of sensitivity and specificity from a study. The solid black circle represents the pooled sensitivity and specificity, which is surrounded by a 95% confidence region (dashed line)



Entorhinal cortex

Four studies (deToledo-Morell 2004; Devanand 2007; Herukka 2008; Ledig 2018), measured the volume of the total entorhinal cortex for a total of 529 participants (229 (43%) of whom progressed to Alzheimer's disease dementia). Three studies used a manual method and one study (Ledig 2018), used an automated method. The smallest study (Herukka 2008) recruited 21 participants, while the largest study (Ledig 2018) recruited 343 participants. Sensitivities and specificities ranged from 0.50 to 0.88 and from

0.60 to 1.00, respectively. We did not do a meta-analysis because of sparse and heterogeneous data, which were suggestive of moderate accuracy. The certainty of the evidence (Summary of findings 1) was very low due to risk of bias (-1), imprecision (-1) and inconsistency due to heterogenous study results (-1).

Whole brain

Four studies (Carmichael 2007; Erten-Lyons 2006; Ledig 2018; VanderFlier 2005) measured the whole brain volume for a total of

424 participants (220 (52%) of whom progressed to Alzheimer's disease dementia). The four studies used an automatic or semi-automatic technique. The smallest study (VanderFlier 2005) recruited 15 participants, while the largest study (Ledig 2018) recruited 343 participants. Sensitivities and specificities ranged from 0.33 to 0.92 and from 0.41 to 1.00, respectively. We did not do a meta-analysis because of sparse and heterogeneous data, which were suggestive of moderate accuracy, particularly the largest study (Ledig 2018). The certainty of the evidence (Summary of findings 1) was very low for both sensitivity and specificity due to risk of bias (-1), imprecision (-1) and inconsistency due to heterogeneous study results (-1).

Medial temporal gyrus, lateral temporal lobe, amygdala, cortical grey matter

Due to the limited number of studies, we did not calculate summary estimates for these regions (Data table 12; Data table 13; Data table 14; Data table 15; Data table 16; Data table 17; Data table 18). Visser 1999 studied the total lateral temporal lobe, and Galton 2005 analysed the right lateral temporal lobe. Wang 2006 and Ledig 2018 evaluated the total, left and right amygdala, Prieto del Val 2016 evaluated only the right amygdala. The amygdala volume was measured manually (Wang 2006), and with an automated method (Ledig 2018). Ledig 2018 evaluated the medial temporal gyrus and cortical grey matter.

DISCUSSION

Summary of main results

This review analysed the diagnostic accuracy of structural MRI for the early diagnosis of dementia due to Alzheimer's disease in people with MCI. We used clinical diagnosis of dementia due to Alzheimer's disease at follow-up as the reference standard. Data from 3935 participants with a diagnosis of MCI at baseline, who undertook a structural MRI and were followed for at least one year, were analysed in 33 primary studies published from 1999 to 2019. Sociodemographic and clinical characteristics of participants are presented in Table 1 and key results in Summary of findings 1.

We assumed an add-on role of structural MRI, that is, a test used in addition to the clinical judgement or cognitive test performance or both to improve a timely diagnosis of Alzheimer's disease dementia in people with MCI. The results of this review show that structural MRI did not meet the sensitivity and specificity criteria that are needed for an add-on test, which should be highly specific and sensitive. This evidence was of low certainty for total hippocampus volume, which the largest number of studies reported, moderate for the volumes of the medial temporal lobe and lateral ventricles, and very low for the entorhinal cortex and the whole brain volumes. False positives should be low because a false diagnosis of Alzheimer's disease dementia can lead to a heavy burden for the patient and their family, inappropriate treatment of patients with medications for Alzheimer's disease, or lack of a proper therapy for potentially treatable causes of cognitive impairment. Moreover, false positives have a significant impact on health and social care costs. False negatives also should be low because a timely diagnosis of Alzheimer's disease dementia, at a time when people first seek for help being worried about changes in cognition, behaviour, or functioning, can allow them to receive counselling about lifestyle modifications that may help to slow down the progression of cognitive impairment. A timely diagnosis of Alzheimer's disease

dementia, moreover, allows people to participate in clinical trials of new drugs for dementia due to Alzheimer's disease.

The results of this review cannot be considered conclusive because the included studies were at high or unclear risk of bias and heterogeneous, and data were not sufficient to compare test accuracy between different brain regions or different types of MCI, for example, amnesic or non-amnesic MCI. We found no significant differences in sensitivity or specificity of total hippocampal volume between included studies with regards to follow-up length, or age of participants, but the overall accuracy was better for manual versus automatic MRI techniques in mixed (mostly indirect) comparisons.

In a qualitative review, the authors concluded that, for the early diagnosis of Alzheimer's disease dementia, volume of entorhinal cortex provided better diagnostic accuracies than volume of other brain regions, such as the hippocampus (Leandrou 2018). However, key aspects of this qualitative review undermine its conclusion. The results were based on two studies (deToledo-Morell 2004; Killiany 2000). We included the deToledo-Morell 2004 study and judged it at high risk of bias for patient selection and index test. Sensitivity and specificity of entorhinal cortex were 0.50 and 1.00 respectively. We excluded the study of Killiany 2000 because participants were people with normal cognition or "questionable AD dementia". Leandrou 2018 and colleagues did not assess the quality of evidence of the results arising from their review.

Strengths and weaknesses of the review

Strengths of this review include the following.

1. We conducted an extensive, comprehensive, and sensitive literature search, using different electronic databases, and assessed the eventuality of participants' overlapping in the eligible studies.
2. Two teams of two review authors each independently extracted data and two independent review authors used the QUADAS-2 tool for quality assessments of the included studies.
3. We included only prospective studies of participants who underwent structural MRI before diagnosis of dementia due to Alzheimer's disease, minimising the risk of bias in interpretation of the index test results.
4. We approached authors of studies in an attempt to obtain missing information.

Limitations of this review include the following.

1. Only heterogeneous, small studies were available, and few studies were available for some brain regions. This undermined our confidence in the pooled estimates of structural MRI diagnostic accuracy and likely contributed to the great variability in sensitivity and specificity observed in the included studies.
2. We judged most of the included studies at high or unclear risk of bias, which contributed to the low certainty of evidence we presented in this review for the region with the most studies, total hippocampal volume.
3. The studies varied with respect to the included participants and definition of MCI. Moreover, consecutive enrolling of participants and the method of recruitment used were seldom reported in most of the included studies. We considered

participant selection at high risk of bias in 31 out of 33 included studies.

4. Twenty-four studies did not provide sufficient information regarding the index test, and we had to judge them at high risk of bias in this domain.
5. The studies varied with respect to protocols for structural MRI. Most of the included studies (24 out of 33) described the MRI findings but did not provide a clear, pre-specified definition of what was considered a 'positive' result of structural MRI.
6. Only 13 studies addressed interobserver and intraobserver variability for MRI.
7. Diagnosis of dementia due to Alzheimer's disease would require a histopathological confirmation but this is not feasible in clinical practice. The clinical diagnosis of Alzheimer's disease dementia at follow-up is a delayed verification test, which is an imperfect reference standard and could have introduced bias. Furthermore, the experience of clinicians and the clinical pathway were poorly reported in most of the included studies.

Additional limitations of this review may be the following.

1. We excluded studies that reported MRI accuracy obtained from multiple volumetric brain regions. However, some studies reported the highest diagnostic accuracies when both entorhinal cortex and hippocampus were combined in the analysis (Leandrou 2018), or when hippocampal subvolumes and presubiculum volume were combined (Khan 2015).
2. We addressed accuracy of structural MRI alone and not as a component of a combination of tests. Other reviews reported that assessment of hippocampal volume or medial temporal lobe atrophy in isolation for the early diagnosis of dementia due to Alzheimer's disease in people with MCI is not supported by the current evidence (Frisoni 2013; Frisoni 2017a; Payton 2018; Ten Kate 2017b). These authors recommended that clinical research should focus on assessing the impact of combinations of biomarkers. Neuropsychological tests and multiple putative biomarkers including neuroimaging (MRI or PET) have been proposed, but the clinical usefulness of these biomarkers is still under evaluation.
3. We have presented pooled estimates of sensitivity and specificity, despite the fact that explicit volume cut-offs were not reported, which limits the clinical usefulness of summary estimates. Nonetheless, they are consistent with poor accuracy.

Applicability of findings to the review question

We did not judge any studies as having high concerns about applicability in 'patient selection' and 'index test' domains. Participants and the index text in the included studies did not differ from those targeted by the review question. Three studies presented unclear concern for the 'reference standard' domain and the remaining 30 studies demonstrated low concern. Structural imaging techniques and expertise needed to measure volume of brain areas, although potentially applicable, are not widely used in routine clinical practice. In an Italian study, the choice of neuroimaging technique (CT, MRI, or PET) in the clinical pathway of dementia was driven as much by test availability, physicians' familiarity with the technology, and waiting time for patients as by the patient's age, severity of cognitive impairment, or the diagnostic question (e.g. clinical suspicion of cerebrovascular disease; Frisoni 2017a).

AUTHORS' CONCLUSIONS

Implications for practice

Structural magnetic resonance imaging (MRI) in hippocampus or medial temporal lobe, the most studied brain regions, showed low sensitivity and specificity and did not reach the standard required to be a stand-alone, add-on test for an early diagnosis of dementia due to Alzheimer's disease in people with MCI. This is consistent with international guidelines, which recommend structural MRI to exclude non-degenerative or surgical causes of cognitive impairment but not to diagnose dementia due to Alzheimer's disease. Medial temporal lobe atrophy or hippocampal volume measured by structural MRI cannot be recommended in clinical practice for an early diagnosis of dementia due to Alzheimer's disease in people with MCI.

Implications for research

Research priorities include the definition of what is considered to be a 'positive' result of volumetric assessment of brain regions measured by structural MRI. Research is essential for the development of accurate criteria to address a timely diagnosis of Alzheimer's disease dementia. Frisoni and colleagues proposed a research framework to assess the analytical and clinical validity of biomarkers for Alzheimer's disease and their clinical utility. To achieve these objectives, research priorities include the standardisation of the readout of biomarker assays and thresholds for normality, the evaluation of their performance in detecting early disease, the development of diagnostic algorithms comprising combinations of biomarkers, and the development of clinical guidelines for the use of biomarkers in qualified memory clinics (Frisoni 2017a). Implementation of these proposed research topics are expected to provide useful results over the medium term.

We identified several weaknesses in the included studies using the QUADAS 2 quality assessment tool. We recommend that future studies consider:

1. including large prospective cohorts of consecutive or random samples of people with a definite diagnosis of MCI;
2. using a diagnostic accuracy study design that adheres to the recommendations of the STARDdem Initiative 'Reporting standards for studies of diagnostic test accuracy in dementia' (Noel-Storr 2014);
3. Incorporating the QUADAS 2 tool into the study design (Whiting 2011);
4. Providing a clear, pre-specified definition of what is a 'positive' result of the index test;
5. Assessing interobserver and intraobserver variability; and
6. evaluating long-term outcomes and cost effectiveness of the index text implementation.

ACKNOWLEDGEMENTS

We are grateful to Deirde Beecher, the former Information Specialist of Cochrane Multiple Sclerosis and Rare Diseases of the CNS for help in designing and conducting the literature search and to Anna Noel-Storr, Cochrane Dementia and Cognitive Improvement's Information Specialist, for help in updating the literature search. We thank all the study authors who we contacted, who contributed information to this review.

We would like to thank peer reviewers Mara ten Kate and Tim Donovan and consumer reviewer Cathie Hofstetter for their comments and feedback.

REFERENCES

References to studies included in this review

Carmichael 2007 {published data only}

* Carmichael OT, Kuller LH, Lopez OL, Thompson PM, Dutton RA, Lu A, et al. Cerebral ventricular changes associated with transitions between normal cognitive function, mild cognitive impairment, and dementia. *Alzheimer Disease and Associated Disorders* 2007;**21**(1):14–24. [DOI: [10.1097/WAD.0b013e318032d2b1](https://doi.org/10.1097/WAD.0b013e318032d2b1).]

Carmichael OT, Kuller LH, Lopez OL, Thompson PM, Dutton RA, Lu A, et al. Ventricular volume and dementia progression in the Cardiovascular Health Study. *Neurobiology of Aging* 2007;**28**(3):389–97. [PUBMED: 16504345]

Yue NC, Arnold AM, Longstreth WT, Elster AD, Jungreis CA, O'Leary DH, et al. Sulcal, ventricular, and white matter changes at MR imaging in the aging brain: data from the Cardiovascular Health Study. *Radiology* 1997;**202**:33–9.

Caroli 2007 {published data only}

Caroli A, Testa C, Geroldi C, Nobili F, Barnden LR, Guerra UP, et al. Cerebral perfusion correlates of conversion to Alzheimer's disease in amnesic mild cognitive impairment. *Journal of Neurology* 2007;**254**:1698–707. [DOI: [10.1007/s00415-007-0631-7](https://doi.org/10.1007/s00415-007-0631-7)]

Clerx 2013a {published data only}

Clerx L, Van Der Pol L, Rueckert D, De Jong R, Van Schijndel R, Verhey FRJ, et al. Comparison of measurements of medial temporal lobe atrophy in the prediction of Alzheimer's disease in subjects with MCI. *Alzheimer's & Dementia* 2011;**7** Suppl 4:133.

Clerx L, Van Rossum I, Burns L, Knol D, Scheltens P, Verhey FRJ, et al. MRI traits identify several loci influencing degeneration of the hippocampus lobe atrophy in the prediction of Alzheimer's disease in subjects with MCI. *Alzheimer's & Dementia* 2012;**8** Suppl 4:738–9.

* Clerx L, Van Rossum IA, Burns L, Knol DL, Scheltens P, Verhey F, et al. Measurements of medial temporal lobe atrophy for prediction of Alzheimer's disease in subjects with mild cognitive impairment. *Neurobiology of Aging* 2013;**34**(8):2003–13.

deToledo-Morell 2004 {published data only}

de Toledo-Morrell L, Goncharova I, Dickerson B, Wilson RS, Bennett DA. From healthy aging to early Alzheimer's disease: in vivo detection of entorhinal cortex atrophy. *Annals of the New York Academy of Sciences* 2000;**911**:240–53.

* deToledo-Morrell L, Stoub TR, Bulgakova M, Wilson RS, Bennett DA, Leurgans S, et al. MRI-derived entorhinal volume is a good predictor of conversion from MCI to AD. *Neurobiology of Aging* 2004;**25**:1197–203.

Devanand 2007 {published data only}

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**:828–36.

Eckerstrom 2008 {published data only}

Eckerström C, Olsson E, Borga M, Ekholm S, Ribbelin S, Rolstad S, et al. Small baseline volume of left hippocampus is associated with subsequent conversion of MCI into dementia: the Göteborg MCI study. *Journal of the Neurological Sciences* 2008;**272**:48–59.

Eckerstrom 2013 {published data only}

Eckerström C, Olsson E, Bjerke M, Malmgren H, Edman A, Wallin A, et al. A combination of neuropsychological, neuroimaging, and cerebrospinal fluid markers predicts conversion from mild cognitive impairment to dementia. *Journal of Alzheimer's Disease* 2013;**36**(3):421–31.

Erten-Lyons 2006 {published data only}

Erten-Lyons D, Howieson D, Moore MM, Quinn J, Sexton G, Silbert L, et al. Brain volume loss in MCI predicts dementia. *Neurology* 2006;**66**:233–5.

Frolich 2017 {published data only}

Frolich L, Peters O, Lewczuk P, Gruber O, Teipel SJ, Gertz HJ, et al. Incremental value of biomarker combinations to predict progression of mild cognitive impairment to Alzheimer's dementia. *Alzheimer's Research & Therapy* 2017;**9**(1):84.

Galton 2005 {published data only}

* Galton CJ, Erzinçlioglu S, Sahakian BJ, Antoun N, Hodges JR. A comparison of the Addenbrooke's Cognitive Examination (ACE), conventional neuropsychological assessment, and simple MRI-based medial temporal lobe evaluation in the early diagnosis of Alzheimer's disease. *Cognitive and Behavioral Neurology* 2005;**18**:144–50.

Galton CJ, Gomez-Anson B, Antoun N, Scheltens P, Patterson K, Graves M, et al. Temporal lobe rating scale: application to Alzheimer's disease and frontotemporal dementia. *Journal of Neurology, Neurosurgery, and Psychiatry* 2001;**70**:165–73.

Gaser 2013 {published data only}

Gaser C, Franke K, Klöppel S, Koutsouleris N, Sauer H. BrainAGE in mild cognitive impaired patients: predicting the conversion to Alzheimer's disease. *PLoS One* 2013;**8**(6):e67346.

Herukka 2008 {published data only}

* Herukka SK, Pennanen C, Soininen H, Pirttila T. CSF Aβ42, tau and phosphorylated tau correlate with medial temporal lobe atrophy. *Journal of Alzheimer's Disease* 2008;**14**(1):51–7.

Pennanen C, Kivipelto M, Tuomainen S, Hartikainen P, Hanninen T, Laakso MP, et al. Hippocampus and entorhinal cortex in mild cognitive impairment and early AD. *Neurobiology of Aging* 2004;**25**:303–10.

Jack 2000 {published data only}

Jack CR Jr, Petersen RC, Xu Y, O'Brien PC, Smith GE, Ivnik RJ, et al. Rates of hippocampal atrophy correlate with change in clinical status in aging and AD. *Neurology* 2000;**55**(4):484-9.

Jang 2018 {published data only}

Jang JW, Park JH, Kim S, Park YH, Pyun JM, Lim JS, et al. Alzheimer's Disease Neuroimaging Initiative. A 'comprehensive visual rating scale' for predicting progression to dementia in patients with mild cognitive impairment. *PLoS One* 2018;**13**(8):e0201852.

Khan 2015 {published data only}

Khan W, Westman E, Jones N, Wahlund LO, Mecocci P, Vellas B, et al. AddNeuroMed consortium for the Alzheimer's Disease Neuroimaging Initiative. Automated hippocampal subfield measures as predictors of conversion from mild cognitive impairment to Alzheimer's disease in two independent cohorts. *Brain Topography* 2015;**28**(5):746-59.

Ledig 2018 {published data only}

Ledig C, Schuh A, Guerrero R, Heckemann RA, Rueckert D. Structural brain imaging in Alzheimer's disease and mild cognitive impairment: biomarker analysis and shared morphometry database. *Scientific Reports* 2018;**8**(1):11258.

Liu 2010 {published data only}

Liu Y, Paajanen T, Zhang Y, Westman E, Wahlund LO, Simmons A, et al. AddNeuroMed Consortium. Analysis of regional MRI volumes and thickness as predictors of conversion from mild cognitive impairment to Alzheimer's disease. *Neurobiology of Aging* 2010;**31**(8):1375-85.

Monge Argilés 2014 {published data only}

Monge Argilés JA, Blanco Cantó MA, Leiva Salinas C, Flors L, Muñoz Ruiz C, Sánchez Payá J, et al. A comparison of early diagnostic utility of Alzheimer disease biomarkers in brain magnetic resonance and cerebrospinal fluid. *Neurologia* 2014;**29**(7):397-401.

Nesteruk 2016 {published data only}

Nesteruk M, Nesteruk T, Styczyńska M, Barczak A, Mandecka M, Walecki J, et al. Predicting the conversion of mild cognitive impairment to Alzheimer's disease based on the volumetric measurements of the selected brain structures in magnetic resonance imaging. *Neurologia i Neurochirurgia Polska* 2015;**49**(6):349-53.

* Nesteruk M, Nesteruk T, Styczyńska M, Mandecka M, Barczak A, Barcikowska M. Combined use of biochemical and volumetric biomarkers to assess the risk of conversion of mild cognitive impairment to Alzheimer's disease. *Folia Neuropathologica* 2016;**54**(4):369-74.

Ong 2015 {published data only}

Ong KT, Villemagne VL, Bahar-Fuchs A, Lamb F, Langdon N, Catafau AM, et al. Aβ imaging with 18F-florbetaben in prodromal Alzheimer's disease: a prospective outcome study. *Journal of Neurology, Neurosurgery, and Psychiatry* 2015;**86**(4):431-6.

Pereira 2014 {published data only}

Pereira JB, Cavallin L, Spulber G, Aguilar C, Mecocci P, Vellas B, et al. AddNeuroMed consortium and for the Alzheimer's Disease Neuroimaging Initiative. Influence of age, disease onset and ApoE4 on visual medial temporal lobe atrophy cut-offs. *Journal of Internal Medicine* 2014;**275**(3):317-30.

Platero 2019 {published data only}

Platero C, López ME, Carmen Tobar MD, Yus M, Maestu F. Discriminating Alzheimer's disease progression using a new hippocampal marker from T1-weighted MRI: the local surface roughness. *Human Brain Mapping* 2019;**40**(5):1666-76.

Prestia 2013 {published data only}

Prestia A, Caroli A, Van Der Flier WM, Ossenkoppele R, Van Berckel B, Barkhof F, et al. Prediction of dementia in MCI patients based on core diagnostic markers for Alzheimer disease. *Neurology* 2013;**80**(11):1048-56.

Prestia 2013 (ADNI) {published data only}

Prestia A, Caroli A, Herholz K, Reiman E, Chen K, Jagust WJ, et al. Diagnostic accuracy of markers for prodromal Alzheimer's disease in independent clinical series. *Alzheimer's and Dementia* 2013;**9**(6):677-86.

Prieto del Val 2016 {published data only}

Prieto Del Val L, Cantero JL, Atienza M. Atrophy of amygdala and abnormal memory-related alpha oscillations over posterior cingulate predict conversion to Alzheimer's disease. *Scientific Reports* 2016;**6**:31859.

Rhodus-Meester 2016 {published data only}

Rhodus-Meester HF, Koikkalainen J, Mattila J, Teunissen CE, Barkhof F, Lemstra AW, et al. Integrating biomarkers for underlying Alzheimer's disease in mild cognitive impairment in daily practice: comparison of a clinical decision support system with individual biomarkers. *Journal of Alzheimer's Disease* 2016;**50**(1):261-70.

VanderFlier 2005 {published data only}

Van der Flier WM, Van der Vlies AE, Weverling-Rijnsburger AW, De Boer NL, Admiraal-Behloul F, Bollen EL, et al. MRI measures and progression of cognitive decline in non demented elderly attending a memory clinic. *International Journal of Geriatric Psychiatry* 2005;**20**(11):1060-6. [PubMed: 16250078]

Visser 1999 {published data only}

Visser PJ, Scheltens P, Verhey FR, Schmand B, Launer LJ, Jolles J, et al. Medial temporal lobe atrophy and memory dysfunction as predictors for dementia in subjects with mild cognitive impairment. *Journal of Neurology* 1999;**246**(6):477-85. [PubMed: 10431775]

Visser 2002 {published data only}

Visser PJ, Verhey FR, Hofman PA, Scheltens P, Jolles J. Medial temporal lobe atrophy predicts Alzheimer's disease in patients with minor cognitive impairment. *Journal of Neurology, Neurosurgery, and Psychiatry* 2002;**72**(4):491-7. [PubMed: 11909909]

Wang 2006 {published data only}

Wang L, Miller JP, Gado MH, McKeel DW, Rothermich M, Miller MI, et al. Abnormalities of hippocampal surface structure in very mild dementia of the Alzheimer type. *NeuroImage* 2006;**30**(1):52-60. [PUBMED: 16243546]

Westman 2011 {published data only}

Westman E, Cavallin L, Muehlboeck JS, Zhang Y, Mecocci P, Vellas B, et al. Sensitivity and specificity of medial temporal lobe visual ratings and multivariate regional MRI classification in Alzheimer's disease. *PLoS One* 2011;**6**(7):e22506.

Wolz 2011 {published data only}

Wolz R, Julkunen V, Koikkalainen J, Niskanen E, Zhang DP, Rueckert D, et al. Multi-method analysis of MRI images in early diagnostics of Alzheimer's disease. *PLoS One* 2011;**6**(10):e25446.

Wood 2016 {published data only}

Wood RA, Moodley KK, Lever C, Minati L, Chan D. Allocentric spatial memory testing predicts conversion from mild cognitive impairment to dementia: an initial proof-of-concept study. *Frontiers in Neurology* 2016;**7**:215.

References to studies excluded from this review
Aguilar 2014 {published data only}

Aguilar C, Muehlboeck JS, Mecocci P, Vellas B, Tsolaki M, Kloszewska I, et al. AddNeuroMed Consortium. Application of a MRI based index to longitudinal atrophy change in Alzheimer disease, mild cognitive impairment and healthy older individuals in the AddNeuroMed cohort. *Frontiers in Aging Neuroscience* 2014;**6**:145.

Aksu 2011 {published data only}

Aksu Y, Miller DJ, Kesidis G, Bigler DC, Yang QX. An MRI-derived definition of MCI-to-AD conversion for long-term, automatic prognosis of MCI patients. *PLoS One* 2011;**6**(10):e25074.

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. [PUBMED: 16682538]

Apostolova 2014 {published data only}

Apostolova LG, Hwang KS, Kohannim O, Avila D, Elashoff D, Jack CR Jr, et al. Alzheimer's Disease Neuroimaging Initiative. ApoE4 effects on automated diagnostic classifiers for mild cognitive impairment and Alzheimer's disease. *NeuroImage: Clinical* 2014;**4**:461-72.

Archer 2010 {published data only}

Archer HA, Kennedy J, Barnes J, Pepple T, Boyes R, Randlesome K, et al. Memory complaints and increased rates of brain atrophy: risk factors for mild cognitive impairment and Alzheimer's disease. *International Journal of Geriatric Psychiatry* 2010;**25**(11):1119-26. [PUBMED: 20084620]

Ardekani 2017 {published data only}

Ardekani BA, Bermudez E, Mubeen AM, Bachman AH, Alzheimer's Disease Neuroimaging Initiative. Prediction of incipient Alzheimer's disease dementia in patients with mild cognitive impairment. *Journal of Alzheimer's Disease* 2017;**55**(1):269-81.

Bakkour 2009 {published data only}

Bakkour A, Morris JC, Dickerson BC. The cortical signature of prodromal AD: regional thinning predicts mild AD dementia. *Neurology* 2009;**72**(12):1048-55. [PUBMED: 19109536]

Barnes 2014 {published data only}

Barnes DE, Cener IS, Yaffe K, Ritchie CS, Lee SJ, Alzheimer's Disease Neuroimaging Initiative. A point-based tool to predict conversion from mild cognitive impairment to probable Alzheimer's disease. *Alzheimer's & Dementia* 2014;**10**(6):646-55.

Beheshti 2016 {published data only}

Beheshti I, Demirel H, Farokhian F, Yang C, Matsuda H, Alzheimer's Disease Neuroimaging Initiative. Structural MRI-based detection of Alzheimer's disease using feature ranking and classification error. *Computer Methods and Programs in Biomedicine* 2016;**137**:177-93.

Beheshti 2017 {published data only}

Beheshti I, Demirel H, Matsuda H, Alzheimer's Disease Neuroimaging Initiative. Classification of Alzheimer's disease and prediction of mild cognitive impairment-to-Alzheimer's conversion from structural magnetic resonance imaging using feature ranking and a genetic algorithm. *Computers in Biology and Medicine* 2017;**83**:109-19.

Bell-McGinty 2005 {published data only}

Bell-McGinty S, Lopez OL, Meltzer CC, Scanlon JM, Whyte EM, Dekosky ST, et al. Differential cortical atrophy in subgroups of mild cognitive impairment. *Archives of Neurology* 2005;**62**(9):1393-7. [PUBMED: 16157746]

Bernard 2014 {published data only}

Bernard C, Helmer C, Dilharreguy B, Amieva H, Auriacombe S, Dartigues JF, et al. Time course of brain volume changes in the preclinical phase of Alzheimer's disease. *Alzheimer's & Dementia* 2014;**10**(2):143-51.

Blasko 2008 {published data only}

Blasko I, Jellinger K, Kemmler G, Krampla W, Jungwirth S, Wichart I, et al. Conversion from cognitive health to mild cognitive impairment and Alzheimer's disease: prediction by plasma amyloid beta 42, medial temporal lobe atrophy and homocysteine. *Neurobiology of Aging* 2008;**29**(1):1-11. [PUBMED: 17055615]

Bombois 2008 {published data only}

Bombois S, Debette S, Bruandet A, Delbeuck X, Delmaire C, Leys D, et al. Vascular subcortical hyperintensities predict conversion to vascular and mixed dementia in MCI patients. *Stroke* 2008;**39**(7):2046-51. [PUBMED: 18436882]

Borgio 2012 {published data only}

Borgio JG, Baldaçara L, Moraes Wdos S, Lacerda AL, Montañó MB, Jackowski AP, et al. Hippocampal volume and CDR-SB can predict conversion to dementia in MCI patients. *Arquivos de Neuro-Psiquiatria* 2012;**70**(11):839-42.

Boutet 2012 {published data only}

Boutet C, Chupin M, Colliot O, Sarazin M, Mutlu G, Drier A, et al. Alzheimer's Disease Neuroimaging Initiative. Is radiological evaluation as good as computer-based volumetry to assess hippocampal atrophy in Alzheimer's disease?. *Neuroradiology* 2012;**54**(12):1321-30.

Bouwman 2007 {published data only}

Bouwman FH, Schoonenboom SN, Van der Flier WM, Van Elk EJ, Kok A, Barkhof F, et al. CSF biomarkers and medial temporal lobe atrophy predict dementia in mild cognitive impairment. *Neurobiology of Aging* 2007;**28**(7):1070-4.

Brickman 2015 {published data only}

Brickman AM, Zahodne LB, Guzman VA, Narkhede A, Meier IB, Griffith EY, et al. Reconsidering harbingers of dementia: progression of parietal lobe white matter hyperintensities predicts Alzheimer's disease incidence. *Neurobiology of Aging* 2015;**36**(1):27-32.

Bron 2014 {published data only}

Bron EE, Steketee RM, Houston GC, Oliver RA, Achterberg HC, Loog M, et al. for the Alzheimer's Disease Neuroimaging Initiative. Diagnostic classification of arterial spin labeling and structural MRI in presenile early stage dementia. *Human Brain Mapping* 2014;**35**(9):4916-31.

Bron 2015 {published data only}

Bron EE, Smits M, Niessen WJ, Klein S. Feature selection based on the SVM weight vector for classification of dementia. *IEEE Journal of Biomedical and Health Informatics* 2015;**19**(5):1617-26.

Brück 2013 {published data only}

Brück 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.

Brueggen 2015 {published data only}

Brueggen K, Dyrba M, Barkhof F, Hausner L, Filippi M, Nestor PJ, et al. Basal forebrain and hippocampus as predictors of conversion to Alzheimer's disease in patients with mild cognitive impairment - a multicenter DTI and volumetry study. *Journal of Alzheimer's Disease* 2015;**48**(1):197-204.

Brys 2009 {published data only}

Brys M, Glodzik L, Mosconi L, Switalski R, De Santi S, Pirraglia E, et al. Magnetic resonance imaging improves cerebrospinal fluid biomarkers in the early detection of Alzheimer's disease. *Journal of Alzheimer's Disease* 2009;**16**(2):351-62. [PUBMED: 19221425]

Buckley 2017 {published data only}

Buckley CJ, Inglis F, Cherubini A, Zanette M, Farrar G, Brooks DJ, et al. Performance of [¹⁸F] flutemetamol amyloid scanning in a phase III amnesic mild cognitive impairment study: additional influence of other biomarkers in estimating risk of conversion to probable Alzheimer's disease. *Alzheimer's & Dementia* 2017;**13**(7):P221.

Buckner 2005 {published data only}

Buckner RL, Snyder AZ, Shannon BJ, LaRossa G, Sachs R, Fotenos AF, et al. Molecular, structural, and functional characterization of Alzheimer's disease: evidence for a relationship between default activity, amyloid, and memory. *Journal of Neuroscience* 2005;**25**(34):7709-17. [PUBMED: 16120771]

Callahan 2015 {published data only}

Callahan BL, Ramirez J, Berezuk C, Duchesne S, Black SE, for the Alzheimer's Disease Neuroimaging Initiative. Predicting Alzheimer's disease development: a comparison of cognitive criteria and associated neuroimaging biomarkers. *Alzheimer's Research & Therapy* 2015;**7**(1):68.

Cardenas 2003 {published data only}

Cardenas VA, Du AT, Hardin D, Ezekiel F, Weber P, Jagust WJ, et al. Comparison of methods for measuring longitudinal brain change in cognitive impairment and dementia. *Neurobiology of Aging* 2003;**24**(4):537-44. [PUBMED: 12714110]

Carmichael 2013 {published data only}

Carmichael O, McLaren DG, Tommet D, Mungas D, Jones RN, for the Alzheimer's Disease Neuroimaging Initiative. Coevolution of brain structures in amnesic mild cognitive impairment. *NeuroImage* 2013;**66**:449-56.

Caroli 2015 {published data only}

Caroli A, Prestia A, Wade S, Chen K, Ayutyanont N, Landau SM, et al. for the Alzheimer's Disease Neuroimaging Initiative. Alzheimer disease biomarkers as outcome measures for clinical trials in MCI. *Alzheimer Disease and Associated Disorders* 2015;**29**(2):101-9.

Casanova 2013 {published data only}

Casanova R, Hsu FC, Sink KM, Rapp SR, Williamson JD, Resnick SM, et al. Alzheimer's Disease Neuroimaging Initiative. Alzheimer's disease risk assessment using large-scale machine learning methods. *PloS One* 2013;**8**(11):e77949.

Cespedes 2017 {published data only}

Cespedes MI, Fripp J, McGree JM, Drovandi CC, Mengersen K, Doecke JD. Comparisons of neurodegeneration over time between healthy ageing and Alzheimer's disease cohorts via Bayesian inference. *BMJ Open* 2017;**7**(2):e012174.

Cevik 2017 {published data only}

Cevik A, Weber GW, Eyuboglu BM, Oguz KK. Voxel-MARS: a method for early detection of Alzheimer's disease by classification of structural brain MRI. *Annals of Operations Research* 2017;**258**(1):31-57.

Chan 2016 {published data only}

Chan D, Gallaher LM, Moodley K, Minati L, Burgess N, Hartley T. The 4 Mountains test: a short test of spatial memory with high sensitivity for the diagnosis of pre-dementia Alzheimer's disease. *Journal of Visualized Experiments : JoVE* 2016;**116**:54454.

Chao 2005 {published data only}

Chao LL, Schuff N, Kramer JH, Du AT, Capizzano AA, O'Neill J, et al. Reduced medial temporal lobe N-acetylaspartate in cognitively impaired but non demented patients. *Neurology* 2005;**64**(2):282-9. [PUBMED: 15668426]

Cheng 2012 {published data only}

Cheng B, Zhang D, Shen D. Domain transfer learning for MCI conversion prediction. *Medical Image Computing and Computer-assisted Intervention : MICCAI ... International Conference on Medical Image Computing and Computer-Assisted Intervention* 2012;**15**(Pt 1):82-90.

Cheng 2015a {published data only}

Cheng B, Liu M, Zhang D, Munsell BC, Shen D. Domain transfer learning for MCI conversion prediction. *IEEE Transactions on Bio-medical Engineering* 2015;**62**(7):1805-17.

Cheng 2015b {published data only}

Cheng B, Liu M, Suk HI, Shen D, Zhang D, Alzheimer's Disease Neuroimaging Initiative. Multimodal manifold-regularized transfer learning for MCI conversion prediction. *Brain Imaging and Behavior* 2015;**9**(4):913-26.

Chertkow 2012 {published data only}

Chertkow H, Bergman H, Bock C, Wolfson C, McKelvey R, Whitehead V. Amnesic mild cognitive impairment in a memory clinic: longitudinal course and predictors of progression (abstract). *Alzheimer's & Dementia* 2012;**8** Suppl 4:482.

Chetelat 2005 {published data only}

Chetelat G, Landeau B, Eustache F, Mezenge F, Viader F, de la Sayette V, et al. Using voxel-based morphometry to map the structural changes associated with rapid conversion in MCI: a longitudinal MRI study. *NeuroImage* 2005;**27**(4):934-46. [PUBMED: 15979341]

Chincarini 2011 {published data only}

Chincarini A, Bosco P, Calvini P, Gemme G, Esposito M, Olivieri C, et al. Alzheimer's Disease Neuroimaging Initiative. Local MRI analysis approach in the diagnosis of early and prodromal Alzheimer's disease. *NeuroImage* 2011;**58**(2):469-80.

Chincarini 2014 {published data only}

Chincarini A, Bosco P, Gemme G, Esposito M, Rei L, Squarcia S, et al. Alzheimer's Disease Neuroimaging Initiative. Automatic temporal lobe atrophy assessment in prodromal AD: data from the DESCRIPA study. *Alzheimer's & Dementia* 2014;**10**(4):456-67.

Chincarini 2016 {published data only}

Chincarini A, Sensi F, Rei L, Gemme G, Squarcia S, Longo R, et al. Alzheimer's Disease Neuroimaging Initiative. Integrating longitudinal information in hippocampal volume

measurements for the early detection of Alzheimer's disease. *NeuroImage* 2016;**125**:834-47.

Cho 2012 {published data only}

Cho Y, Seong JK, Jeong Y, Shin SY, Alzheimer's Disease Neuroimaging Initiative. Individual subject classification for Alzheimer's disease based on incremental learning using a spatial frequency representation of cortical thickness data. *NeuroImage* 2012;**59**(3):2217-30.

Chow 2015 {published data only}

Chow N, Hwang KS, Hartz S, Green AE, Somme JH, Thompson PM, et al. Alzheimer's Disease Neuroimaging Initiative. Comparing 3T and 1.5T MRI for mapping hippocampal atrophy in the Alzheimer's Disease Neuroimaging Initiative. *American Journal of Neuroradiology* 2015;**36**(4):653-60.

Chu 2012 {published data only}

Chu C, Hsu AL, Chou KH, Bandettini P, Lin C, Alzheimer's Disease Neuroimaging Initiative. Does feature selection improve classification accuracy? Impact of sample size and feature selection on classification using anatomical magnetic resonance images. *NeuroImage* 2012;**60**(1):59-70.

Chung 2016 {published data only}

Chung JK, Plitman E, Nakajima S, Chakravarty MM, Caravaggio F, Takeuchi H, et al. Depressive symptoms and small hippocampal volume accelerate the progression to dementia from mild cognitive impairment. *Journal of Alzheimer's Disease* 2016;**49**(3):743-54.

Chupin 2009 {published data only}

Chupin M, Gérardin E, Cuingnet R, Boutet C, Lemieux L, Lehericy S, et al. Fully automatic hippocampus segmentation and classification in Alzheimer's disease and mild cognitive impairment applied on data from ADNI. *Hippocampus* 2009;**19**(6):579-87.

Citak-Er 2017 {published data only}

Citak-Er F, Goularas D, Ormeci B. A novel convolutional neural network model based on voxel-based morphometry of imaging data in predicting the prognosis of patients with mild cognitive impairment. *Journal of Neurological Sciences* 2017;**34**(1):52-69.

Clerx 2013b {published data only}

Clerx L, Jacobs HI, Burgmans S, Gronenschild EH, Uylings HB, Echávarri C, et al. Sensitivity of different MRI-techniques to assess grey matter atrophy patterns in Alzheimer's disease is region-specific. *Current Alzheimer Research* 2013;**10**(9):940-51.

Clerx 2014 {published data only}

Clerx L, Dierckx E, Van Der Pol L, Van Rossum I, Verhey FR, Aalten P, et al. Added value of MRI biomarkers to neuropsychological test performance for prediction of AD in subjects with MCI. *Alzheimer's & Dementia* 2014;**10**(4):265.

Convit 2000 {published data only}

Convit A, de Asis J, de Leon MJ, Tarshish CY, De Santi S, Rusinek H. Atrophy of the medial occipitotemporal, inferior, and middle temporal gyri in non-demented elderly predict decline

- to Alzheimer's disease. *Neurobiology of Aging* 2000;**21**(1):19-26. [PUBMED: 10794844]
- Cover 2016** {published data only}
 Cover KS, Van Schijndel RA, Versteeg A, Leung KK, Redolfi A, Manset D, et al. The measurement of hippocampal atrophy rates with MRI for a 3-year study appears to be at least 3 times more sensitive than a 1-year study based on back-to-back reproducibility. *Alzheimer's & Dementia* 2016;**12** Suppl 7:99-100.
- Cui 2011** {published data only}
 Cui Y, Liu B, Luo S, Zhen X, Fan M, Liu T, et al. Alzheimer's Disease Neuroimaging Initiative. Identification of conversion from mild cognitive impairment to Alzheimer's disease using multivariate predictors. *PloS One* 2011;**6**(7):E21896.
- Cuingnet 2011** {published data only}
 Cuingnet R, Gerardin E, Tessieras J, Auzias G, Lehericy S, Habert MO, et al. Alzheimer's Disease Neuroimaging Initiative. Automatic classification of patients with Alzheimer's disease from structural MRI: a comparison of ten methods using the ADNI database. *NeuroImage* 2011;**56**(2):766-81.
- Da 2013** {published data only}
 Da X, Toledo JB, Zee J, Wolk DA, Xie SX, Ou Y, et al. Alzheimer's Disease Neuroimaging Initiative. Integration and relative value of biomarkers for prediction of MCI to AD progression: spatial patterns of brain atrophy, cognitive scores, APOE genotype and CSF biomarkers. *NeuroImage: Clinical* 2013;**4**:164-73.
- Damian 2013** {published data only}
 Damian M, Hausner L, Jekel K, Richter M, Froelich L, Almkvist O, et al. Single-domain amnesic mild cognitive impairment identified by cluster analysis predicts Alzheimer's disease in the European prospective DESCRIPA study. *Dementia and Geriatric Cognitive Disorders* 2013;**36**(1-2):1-19.
- Davatzikos 2011** {published data only}
 Davatzikos C, Bhatt P, Shaw LM, Batmanghelich KN, Trojanowski JQ. Prediction of MCI to AD conversion, via MRI, CSF biomarkers, and pattern classification. *Neurobiology of Aging* 2011;**32**(12):2322. e19-27.
- De Leon 1997** {published data only}
 De Leon MJ, George AE, Golomb J, Tarshish C, Convit A, Kluger A, et al. Frequency of hippocampal formation atrophy in normal aging and Alzheimer's disease. *Neurobiology of Aging* 1997;**18**(1):1-11.
- de Leon 2006** {published data only}
 de Leon MJ, DeSanti S, Zinkowski R, Mehta PD, Pratico D, Segal S, et al. Longitudinal CSF and MRI biomarkers improve the diagnosis of mild cognitive impairment. *Neurobiology of Aging* 2006;**27**(3):394-401.
- de Leon 2007** {published data only}
 de Leon MJ, Mosconi L, Li J, De Santi S, Yao Y, Tsui WH, et al. Longitudinal CSF isoprostane and MRI atrophy in the progression to AD. *Journal of Neurology* 2007;**254**(12):1666-75.
- Den Heijer 2006** {published data only}
 Den Heijer T, Geerlings MI, Hoebeek FE, Hofman A, Koudstaal PJ, Breteler MM. Use of hippocampal and amygdala volumes on magnetic resonance imaging to predict dementia in cognitively intact elderly people. *Archives of General Psychiatry* 2006;**63**(1):57-62.
- Desikan 2008** {published data only}
 Desikan RS, Fischl B, Cabral HJ, Kemper TL, Guttman CR, Blacker D, et al. MRI measures of temporoparietal regions show differential rates of atrophy during prodromal AD. *Neurology* 2008;**71**(11):819-25.
- 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.
- Devanand 2012** {published data only}
 Devanand DP, Bansal R, Liu J, Hao X, Pradhaban G, Peterson BS. MRI hippocampal and entorhinal cortex mapping in predicting conversion to Alzheimer's disease. *NeuroImage* 2012;**60**(3):1622-9.
- Dickerson 2001** {published data only}
 Dickerson BC, Goncharova I, Sullivan MP, Forchetti C, Wilson RS, Bennett DA, et al. MRI-derived entorhinal and hippocampal atrophy in incipient and very mild Alzheimer's disease. *Neurobiology of Aging* 2001;**22**(5):747-54.
- Dickerson 2004** {published data only}
 Dickerson BC, Salat DH, Bates JF, Atiya M, Killiany RJ, Greve DN, et al. Medial temporal lobe function and structure in mild cognitive impairment. *Annals of Neurology* 2004;**56**(1):27-35.
- Dickerson 2013** {published data only}
 Dickerson BC, Wolk DA, Alzheimer's Disease Neuroimaging Initiative. Biomarker-based prediction of progression in MCI: comparison of AD signature and hippocampal volume with spinal fluid amyloid- β and tau. *Frontiers in Aging Neuroscience* 2013;**5**:55.
- Douaud 2013** {published data only}
 Douaud G, Menke RA, Gass A, Monsch AU, Rao A, Whitcher B, et al. Brain microstructure reveals early abnormalities more than two years prior to clinical progression from mild cognitive impairment to Alzheimer's disease. *Journal of Neuroscience* 2013;**33**(5):2147-55.
- Doyle 2014** {published data only}
 Doyle OM, Westman E, Marquand AF, Mecocci P, Vellas B, Tsolaki M, et al. Predicting progression of Alzheimer's disease using ordinal regression. *PloS One* 2014;**9**(8):e105542.
- Duara 2008** {published data only}
 Duara R, Loewenstein DA, Potter E, Appel J, Greig MT, Urs R, et al. Medial temporal lobe atrophy on MRI scans and the diagnosis of Alzheimer disease. *Neurology* 2008;**71**(24):1986-92.

Duchesne 2009 {published data only}

Duchesne S, Caroli A, Geroldi C, Collins DL, Frisoni GB. Relating one-year cognitive change in mild cognitive impairment to baseline MRI features. *NeuroImage* 2009;**47**(4):1363-70.

Duchesne 2014 {published data only}

Duchesne S, Valdivia F, Mouiha A, Robitaille N. High-dimensional medial lobe morphometry: an automated MRI biomarker for the new AD diagnostic criteria. *International Journal of Alzheimer's Disease* 2014;**2014**:278096.

Duchesne 2015 {published data only}

Duchesne S, Valdivia F, Mouiha A, Robitaille N, Alzheimer's Disease Neuroimaging Initiative. Single time point high-dimensional morphometry in Alzheimer's disease: group statistics on longitudinally acquired data. *Neurobiology of Aging* 2015;**36**(Supplement 1):11-22.

Dukart 2016 {published data only}

Dukart J, Sambataro F, Bertolino A. Accurate prediction of conversion to Alzheimer's disease using imaging, genetic, and neuropsychological biomarkers. *Journal of Alzheimer's Disease* 2016;**49**(4):1143-59.

Dyrba 2015 {published data only}

Dyrba M, Barkhof F, Fellgiebel A, Filippi M, Hausner L, Hauenstein K, et al. EDSO study group. Predicting prodromal Alzheimer's disease in subjects with mild cognitive impairment using machine learning classification of multimodal multicenter diffusion-tensor and magnetic resonance imaging data. *Journal of Neuroimaging* 2015;**25**(5):738-47.

Eckerström 2015 {published data only}

Eckerström C, Olsson E, Klasson N, Berge J, Nordlund A, Bjerke M, et al. Multimodal prediction of dementia with up to 10 years follow up: the Gothenburg MCI study. *Journal of Alzheimer's Disease* 2015;**44**(1):205-14.

Egli 2015 {published data only}

Egli SC, Hirni DI, Taylor KI, Berres M, Regeniter A, Gass A, et al. Varying strength of cognitive markers and biomarkers to predict conversion and cognitive decline in an early-stage-enriched mild cognitive impairment sample. *Journal of Alzheimer's Disease* 2015;**44**(2):625-33.

El Fakhri 2003 {published data only}

El Fakhri G, Kijewski MF, Johnson KA, Syrkin G, Killiany RJ, Becker JA, et al. MRI-guided SPECT perfusion measures and volumetric MRI in prodromal Alzheimer disease. *Archives of Neurology* 2003;**60**(8):1066-72.

El Fakhri 2004 {published data only}

El Fakhri G, Kijewski MF, Albert MS, Johnson KA, Moore SC. Quantitative SPECT leads to improved performance in discrimination tasks related to prodromal Alzheimer's disease. *Journal of Nuclear Medicine* 2004;**45**(12):2026-31.

Ellis 2014 {published data only}

Ellis KA, Szoek C, Bush AI, Darby D, Graham PL, Lautenschlager NT, et al. AIBL Research Group. Rates of diagnostic transition and cognitive change at 18-month

follow-up among 1112 participants in the Australian imaging, biomarkers and lifestyle flagship study of ageing (AIBL). *International Psychogeriatrics* 2014;**26**(4):543-54.

Eskildsen 2013 {published data only}

Eskildsen SF, Coupé P, García-Lorenzo D, Fonov V, Pruessner JC, Collins DL, Alzheimer's Disease Neuroimaging Initiative. Prediction of Alzheimer's disease in subjects with mild cognitive impairment from the ADNI cohort using patterns of cortical thinning. *NeuroImage* 2013;**65**:511-21.

Eskildsen 2015 {published data only}

Eskildsen SF, Coupé P, Fonov VS, Pruessner JC, Collins DL, Alzheimer's Disease Neuroimaging Initiative. Structural imaging biomarkers of Alzheimer's disease: predicting disease progression. *Neurobiology of Aging* 2015;**36**(Supplement 1):23-31.

Evans 2010 {published data only}

Evans MC, Barnes J, Nielsen C, Kim LG, Clegg SL, Blair M, et al. Volume changes in Alzheimer's disease and mild cognitive impairment: cognitive associations. *European Radiology* 2010;**20**(3):674-82.

Ewers 2012 {published data only}

Ewers M, Walsh C, Trojanowski JQ, Shaw LM, Petersen RC, Jack CR Jr, et al. North American Alzheimer's Disease Neuroimaging Initiative (ADNI). Prediction of conversion from mild cognitive impairment to Alzheimer's disease dementia based upon biomarkers and neuropsychological test performance. *Neurobiology of Aging* 2012;**33**(7):1203-14.

Fan 2008 {published data only}

Fan Y, Batmanghelich N, Clark CM, Davatzikos C. Spatial patterns of brain atrophy in MCI patients, identified via high-dimensional pattern classification, predict subsequent cognitive decline. *NeuroImage* 2008;**39**(4):1731-43.

Fellgiebel 2006 {published data only}

Fellgiebel A, Dellani PR, Greverus D, Scheurich A, Stoeter P, Muller MJ. Predicting conversion to dementia in mild cognitive impairment by volumetric and diffusivity measurements of the hippocampus. *Psychiatry Research* 2006;**146**(3):283-7.

Fjell 2010 {published data only}

Fjell AM, Walhovd KB, Fennema-Notestine C, McEvoy LK, Hagler DJ, Holland D, et al. CSF biomarkers in prediction of cerebral and clinical change in mild cognitive impairment and Alzheimer's disease. *Journal of Neuroscience* 2010;**30**(6):2088-101.

Fleisher 2008 {published data only}

Fleisher AS, Sun S, Taylor C, Ward CP, Gamst AC, Petersen RC, et al. Volumetric MRI vs clinical predictors of Alzheimer disease in mild cognitive impairment. *Neurology* 2008;**70**(3):191-9.

Fouquet 2009 {published data only}

Fouquet M, Desgranges B, Landeau B, Duchesnay E, Mezenge F, de la Sayette V, et al. Longitudinal brain metabolic changes from amnesic mild cognitive impairment to Alzheimer's disease. *Brain* 2009;**132**(Pt 8):2058-67.

Franko 2013 {published data only}

Frankó E, Joly O, Alzheimer's Disease Neuroimaging Initiative. Evaluating Alzheimer's disease progression using rate of regional hippocampal atrophy. *PLoS One* 2013;**8**(8):e71345.

Gao 2018 {published data only}

Gao N, Tao LX, Huang J, Zhang F, Li X, O'Sullivan F, et al. Contourlet-based hippocampal magnetic resonance imaging texture features for multivariant classification and prediction of Alzheimer's disease. *Metabolic Brain Disease* 2018;**33**(6):1899-909.

Gavidia 2017 {published data only}

Gavidia-Bovadilla G, Kanaan-Izquierdo S, Mataró-Serrat M, Perera-Lluna A, Alzheimer's Disease Neuroimaging Initiative. Early prediction of Alzheimer's disease using null longitudinal model-based classifiers. *PLoS One* 2017;**12**(1):e0168011.

Gavrilova 2008 {published data only}

Gavrilova SI, Fedorova YB, Roshchina IF, Korovaitseva GI. Prognosis of mild cognitive impairment syndrome: data from a two-year clinical follow-up study. *Neuroscience and Behavioral Physiology* 2008;**38**(2):129-34.

Geroldi 2006 {published data only}

Geroldi C, Rossi R, Calvagna C, Testa C, Bresciani L, Binetti G, et al. Medial temporal atrophy but not memory deficit predicts progression to dementia in patients with mild cognitive impairment. *Journal of Neurology, Neurosurgery, and Psychiatry* 2006;**77**(11):1219-22.

Gomar 2011 {published data only}

Gomar JJ, Bobes-Bascaran MT, Conejero-Goldberg C, Davies P, Goldberg TE, Alzheimer's Disease Neuroimaging Initiative. Utility of combinations of biomarkers, cognitive markers, and risk factors to predict conversion from mild cognitive impairment to Alzheimer disease in patients in the Alzheimer's disease neuroimaging initiative. *Archives of General Psychiatry* 2011;**68**(9):961-9.

Gómez-Sancho 2018 {published data only}

Gómez-Sancho M, Tohka J, Gómez-Verdejo V. Alzheimer's Disease Neuroimaging Initiative. Comparison of feature representations in MRI-based MCI-to-AD conversion prediction. *Magnetic Resonance Imaging* 2018;**50**:84-95.

Goryawala 2015 {published data only}

Goryawala M, Zhou Q, Barker W, Loewenstein DA, Duara R, Adjouadi M. Inclusion of neuropsychological scores in atrophy models improves diagnostic classification of Alzheimer's disease and mild cognitive impairment. *Computational Intelligence and Neuroscience* 2015;**2015**:865265.

Grothe 2013 {published data only}

Grothe M, Heinsen H, Teipel S. Longitudinal measures of cholinergic forebrain atrophy in the transition from healthy aging to Alzheimer's disease. *Neurobiology of Aging* 2013;**34**(4):1210-20.

Grundman 2002 {published data only}

Grundman M, Sencakova D, Jack CR Jr, Petersen RC, Kim HT, Schultz A, et al. Brain MRI hippocampal volume and prediction of clinical status in a mild cognitive impairment trial. *Journal of Molecular Neuroscience* 2002;**19**(1-2):23-7.

Guo 2017 {published data only}

Guo S, Lai C, Wu C, Cen G, Alzheimer's Disease Neuroimaging Initiative. Corrigendum: conversion discriminative analysis on mild cognitive impairment using multiple cortical features from MR images. *Frontiers in Aging Neuroscience* 2017;**9**:293.

Guzman 2013 {published data only}

Guzman VA, Carmichael OT, Schwarz C, Tosto G, Zimmerman ME, Brickman AM, Alzheimer's Disease Neuroimaging Initiative. White matter hyperintensities and amyloid are independently associated with entorhinal cortex volume among individuals with mild cognitive impairment. *Alzheimer's & Dementia* 2013;**9** Suppl 5:124-31.

Hall 2015 {published data only}

Hall A, Muñoz-Ruiz M, Mattila J, Koikkalainen J, Tsolaki M, Mecocci P, et al. Alzheimer Disease Neuroimaging Initiative, AddNeuroMed consortium, DESCRIPA, Kuopio L-MCI. Generalizability of the disease state index prediction model for identifying patients progressing from mild cognitive impairment to Alzheimer's disease. *Journal of Alzheimer's Disease* 2015;**44**(1):79-92.

Hall 2015b {published data only}

Hall A, Mattila J, Koikkalainen J, Lötjonen J, Wolz R, Scheltens P, et al. Predicting progression from cognitive impairment to Alzheimer's disease with the Disease State Index. *Current Alzheimer Research* 2015;**12**(1):69-79.

Hamalainen 2007 {published data only}

Hamalainen A, Tervo S, Grau-Olivares M, Niskanen E, Pennanen C, Huuskonen J, et al. Voxel-based morphometry to detect brain atrophy in progressive mild cognitive impairment. *NeuroImage* 2007;**37**(4):1122-31.

Hamalainen 2008 {published data only}

Hamalainen A, Grau-Olivares M, Tervo S, Niskanen E, Pennanen C, Huuskonen J, et al. Apolipoprotein E epsilon 4 allele is associated with increased atrophy in progressive mild cognitive impairment: a voxel-based morphometric study. *Neuro-degenerative Diseases* 2008;**5**(3-4):186-9.

Heister 2011 {published data only}

Heister D, Brewer JB, Magda S, Blennow K, McEvoy LK, Alzheimer's Disease Neuroimaging Initiative. Predicting MCI outcome with clinically available MRI and CSF biomarkers. *Neurology* 2011;**77**(17):1619-28.

Henneman 2009 {published data only}

Henneman WJ, Vrenken H, Barnes J, Sluimer IC, Verwey NA, Blankenstein MA, et al. Baseline CSF p-tau levels independently predict progression of hippocampal atrophy in Alzheimer disease. *Neurology* 2009;**73**(12):935-40.

Henry-Feugeas 2008 {published data only}

Henry-Feugeas MC, Onen F, Claeys ES. Classifying late-onset dementia with MRI: is arteriosclerotic brain degeneration the most common cause of Alzheimer's syndrome?. *Clinical Interventions in Aging* 2008;**3**(1):187-99.

Hensel 2005 {published data only}

Hensel A, Wolf H, Busse A, Arendt T, Gertz HJ. Association between global brain volume and the rate of cognitive change in elderly humans without dementia. *Dementia and Geriatric Cognitive Disorders* 2005;**19**(4):213-21.

Hinrichs 2011 {published data only}

Hinrichs C, Singh V, Xu G, Johnson SC, Alzheimers Disease Neuroimaging Initiative. Predictive markers for AD in a multi-modality framework: an analysis of MCI progression in the ADNI population. *NeuroImage* 2011;**55**(2):574-89.

Hu 2016 {published data only}

Hu K, Wang Y, Chen K, Hou L, Zhang X. Multi-scale features extraction from baseline structure MRI for MCI patient classification and AD early diagnosis. *Neurocomputing* 2016;**175**(Part A):132-45.

Huang 2017 {published data only}

Huang M, Yang W, Feng Q, Chen W, Alzheimer's Disease Neuroimaging Initiative. Longitudinal measurement and hierarchical classification framework for the prediction of Alzheimer's disease. *Scientific Reports* 2017;**7**:39880.

Inui 2017 {published data only}

Inui Y, Ito K, Kato T, SEAD-J Study Group. Longer-term investigation of the value of 18F-FDG-PET and magnetic resonance imaging for predicting the conversion of mild cognitive impairment to Alzheimer's disease: a multicenter study. *Journal of Alzheimer's Disease* 2017;**60**(3):877-87.

Jack 2004 {published data only}

Jack CR Jr, Shiung MM, Gunter JL, O'Brien PC, Weigand SD, Knopman DS, et al. Comparison of different MRI brain atrophy rate measures with clinical disease progression in AD. *Neurology* 2004;**62**(4):591-600.

Jack 2005 {published data only}

Jack CR Jr, Shiung MM, Weigand SD, O'Brien PC, Gunter JL, Boeve BF, et al. Brain atrophy rates predict subsequent clinical conversion in normal elderly and amnesic MCI. *Neurology* 2005;**65**(8):1227-31.

Jack 2008a {published data only}

Jack CR Jr, Weigand SD, Shiung MM, Przybelski SA, O'Brien PC, Gunter JL, et al. Atrophy rates accelerate in amnesic mild cognitive impairment. *Neurology* 2008;**70**(19 Pt 2):1740-52.

Jack 2009 {published data only}

Jack CR Jr, Lowe VJ, Weigand SD, Wiste HJ, Senjem ML, Knopman DS, et al. Serial PIB and MRI in normal, mild cognitive impairment and Alzheimer's disease: implications for sequence of pathological events in Alzheimer's disease. *Brain* 2009;**132**(Pt 5):1355-65.

Jacobs 2011 {published data only}

Jacobs HI, Van Boxtel MP, Van der Elst W, Burgmans S, Smeets F, Gronenschild EH, et al. Increasing the diagnostic accuracy of medial temporal lobe atrophy in Alzheimer's disease. *Journal of Alzheimer's Disease* 2011;**25**(3):477-90.

Jie 2013 {published data only}

Jie B, Zhang D, Cheng B, Shen D. Manifold regularized multi-task feature selection for multi-modality classification in Alzheimer's disease. *Medical Image Computing and Computer-assisted Intervention : MICCAI ... International Conference on Medical Image Computing and Computer-Assisted Intervention* 2013;**16**(Pt 1):275-83.

Kalin 2017 {published data only}

Kälin AM, Park MT, Chakravarty MM, Lerch JP, Michels L, Schroeder C, et al. Subcortical shape changes, hippocampal atrophy and cortical thinning in future Alzheimer's disease patients. *Frontiers in Aging Neuroscience* 2017;**9**:38.

Kaneko 2005 {published data only}

Kaneko KT, Momose M, Kadoya M. Neuroimaging tools to rate cognitive impairment. *Psychogeriatrics* 2005;**5**(3):89.

Kantarci 2005 {published data only}

Kantarci K, Petersen RC, Boeve BF, Knopman DS, Weigand SD, O'Brien PC, et al. DWI predicts future progression to Alzheimer disease in amnesic mild cognitive impairment. *Neurology* 2005;**64**(5):902-4. [PUBMED: 15753434]

Kantarci 2009 {published data only}

Kantarci K, Weigand SD, Przybelski SA, Shiung MM, Whitwell JL, Negash S, et al. Risk of dementia in MCI: combined effect of cerebrovascular disease, volumetric MRI, and 1H MRS. *Neurology* 2009;**72**(17):1519-25. [PUBMED: 19398707]

Karas 2008 {published data only}

Karas G, Sluimer J, Goekoop R, Van der Flier W, Rombouts SA, Vrenken H, et al. Amnesic mild cognitive impairment: structural MR imaging findings predictive of conversion to Alzheimer disease. *AJNR. American Journal of Neuroradiology* 2008;**29**(5):944-9. [PUBMED: 18296551]

Kaye 1997 {published data only}

Kaye JA, Swihart T, Howieson D, Dame A, Moore MM, Karnos T, et al. Volume loss of the hippocampus and temporal lobe in healthy elderly persons destined to develop dementia. *Neurology* 1997;**48**(5):1297-304. [PUBMED: 9153461]

Kaye 2005 {published data only}

Kaye JA, Moore MM, Dame A, Quinn J, Camicioli R, Howieson D, et al. Asynchronous regional brain volume losses in presymptomatic to moderate AD. *Journal of Alzheimer's Disease* 2005;**8**(1):51-6. [PUBMED: 16155349]

Khan 2015b {published data only}

Khan W, Aguilar C, Kiddle SJ, Doyle O, Thambisetty M, Muehlboeck S, et al. Alzheimer's Disease Neuroimaging Initiative. A subset of cerebrospinal fluid proteins from a multi-analyte panel associated with brain atrophy, disease

- classification and prediction in Alzheimer's disease. *PLoS One* 2015;**10**(8):e0134368.
- Killiany 2000** {published data only}
 Killiany RJ, Gomez-Isla T, Moss M, Kikinis R, Sandor T, Jolesz F, et al. Use of structural magnetic resonance imaging to predict who will get Alzheimer's disease. *Annals of Neurology* 2000;**47**(4):430-9. [PUBMED: 10762153]
- Kim 2017** {published data only}
 Kim HR, Park YH, Jang JW, Park SY, Wang MJ, Baek MJ, et al. Visual rating of posterior atrophy as a marker of progression to dementia in mild cognitive impairment patients. *Journal of Alzheimer's Disease* 2017;**55**(1):137-46.
- Kloppel 2015** {published data only}
 Klöppel S, Peter J, Ludl A, Pilatus A, Maier S, Mader I, et al. Alzheimer's Disease Neuroimaging Initiative. Applying automated MR-based diagnostic methods to the memory clinic: a prospective study. *Journal of Alzheimer's Disease* 2015;**47**(4):939-54.
- Kong 2014** {published data only}
 Kong D, Giovanello KS, Wang Y, Lee E, Ibrahim JG, Lin W, et al. Role of imaging and genetic data for predicting time to conversion to Alzheimer disease in patients with mild cognitive impairment. *Annals of Neurology* 2014;**76** Suppl 18:S94.
- 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. [PUBMED: 15249617]
- Korolev 2016** {published data only}
 Korolev IO, Symonds LL, Bozoki AC, Alzheimer's Disease Neuroimaging Initiative. Predicting progression from mild cognitive impairment to Alzheimer's dementia using clinical, MRI, and plasma biomarkers via probabilistic pattern classification. *PLoS One* 2016;**11**(2):e0138866.
- Kovacevic 2009** {published data only}
 Kovacevic S, Rafii MS, Brewer JB. High-throughput, fully automated volumetry for prediction of MMSE and CDR decline in mild cognitive impairment. *Alzheimer Disease and Associated Disorders* 2009;**23**(2):139-45. [PUBMED: 19474571]
- Krashenyi 2016** {published data only}
 Krashenyi I, Ramírez J, Popov A, Górriz JM, The Alzheimer's Disease Neuroimaging Initiative. Fuzzy computer-aided Alzheimer's disease diagnosis based on MRI data. *Current Alzheimer Research* 2016;**13**(5):545-56.
- Laforce 2010** {published data only}
 Laforce R Jr, Buteau JP, Paquet N, Verret L, Houde M, Bouchard RW. The value of PET in mild cognitive impairment, typical and atypical/unclear dementias: a retrospective memory clinic study. *American Journal of Alzheimer's Disease and Other Dementias* 2010;**25**(4):324-32. [PUBMED: 20539026]
- Lan 2017** {published data only}
 Lan MJ, Ogden RT, Kumar D, Stern Y, Parsey RV, Pelton GH, et al. Utility of molecular and structural brain imaging to predict progression from mild cognitive impairment to dementia. *Journal of Alzheimer's Disease* 2017;**60**(3):939-47.
- Landau 2010** {published data only}
 Landau SM, Harvey D, Madison CM, Reiman EM, Foster NL, Aisen PS, et al. Alzheimer's Disease Neuroimaging Initiative. Comparing predictors of conversion and decline in mild cognitive impairment. *Neurology* 2010;**75**(3):230-8.
- Lebedev, 2014** {published data only}
 Lebedev AV, Westman E, Van Westen GJ, Kramberger MG, Lundervold A, Aarsland D, et al. Alzheimer's Disease Neuroimaging Initiative and the AddNeuroMed consortium. Random Forest ensembles for detection and prediction of Alzheimer's disease with a good between-cohort robustness. *NeuroImage: Clinical* 2014;**6**:115-25.
- Lee 2015** {published data only}
 Lee E, Zhu H, Kong D, Wang Y, Giovanello KS, Ibrahim JG. BFLCRM: a bayesian functional linear Cox regression model for predicting time to conversion to Alzheimer's disease. *Annals of Applied Statistics* 2015;**9**(4):2153-78.
- Lehman 2013** {published data only}
 Lehmann M, Koedam EL, Barnes J, Bartlett JW, Barkhof F, Wattjes MP, et al. Alzheimer's Disease Neuroimaging Initiative. Visual ratings of atrophy in MCI: prediction of conversion and relationship with CSF biomarkers. *Neurobiology of Aging* 2013;**34**(1):73-82.
- Leung 2010** {published data only}
 Leung KK, Barnes J, Ridgway GR, Bartlett JW, Clarkson MJ, Macdonald K, et al. Automated cross-sectional and longitudinal hippocampal volume measurement in mild cognitive impairment and Alzheimer's disease. *NeuroImage* 2010;**51**(4):1345-59. [PUBMED: 20230901]
- Leung 2013** {published data only}
 Leung KK, Bartlett JW, Barnes J, Manning EN, Ourselin S, Fox NC, Alzheimer's Disease Neuroimaging Initiative. Cerebral atrophy in mild cognitive impairment and Alzheimer disease: rates and acceleration. *Neurology* 2013;**80**(7):648-54.
- Li 2012** {published data only}
 Li X, Jiao J, Shimizu S, Jibiki I, Watanabe K, Kubota T. Correlations between atrophy of the entorhinal cortex and cognitive function in patients with Alzheimer's disease and mild cognitive impairment. *Psychiatry and Clinical Neurosciences* 2012;**66**(7):587-93.
- Li 2014a** {published data only}
 Li H, Liu Y, Gong P, Zhang C, Ye J, Alzheimer's Disease Neuroimaging Initiative. Hierarchical interactions model for predicting mild cognitive impairment (MCI) to Alzheimer's disease (AD) conversion. *PLoS One* 2014;**9**(1):e82450.

Li 2014b {published data only}

Li M, Oishi K, He X, Qin Y, Gao F, Mori S, Alzheimer's Disease Neuroimaging Initiative. An efficient approach for differentiating Alzheimer's disease from normal elderly based on multicenter MRI using gray-level invariant features. *PLoS One* 2014;**9**(8):e105563.

Lillemark 2014 {published data only}

Lillemark L, Sørensen L, Pai A, Dam EB, Nielsen M, Alzheimer's Disease Neuroimaging Initiative. Brain region's relative proximity as marker for Alzheimer's disease based on structural MRI. *BMC Medical Imaging* 2014;**14**:21.

Lin 2018 {published data only}

Lin W, Tong T, Gao Q, Guo D, Du X, Yang Y, et al. Alzheimer's Disease Neuroimaging Initiative. Convolutional neural networks-based MRI image analysis for the Alzheimer's disease prediction from mild cognitive impairment. *Frontiers in Neuroscience* 2018;**12**:777.

Lindemer 2015 {published data only}

Lindemer ER, Salat DH, Smith EE, Nguyen K, Fischl B, Greve DN, Alzheimer's Disease Neuroimaging Initiative. White matter signal abnormality quality differentiates mild cognitive impairment that converts to Alzheimer's disease from non converters. *Neurobiology of Aging* 2015;**36**(9):2447-57.

Liu 2013 {published data only}

Liu Y, Mattila J, Ruiz MÁ, Paaanen T, Koikkalainen J, Van Gils M, et al. Alzheimer's Disease Neuroimaging Initiative. Predicting AD conversion: comparison between prodromal AD guidelines and computer assisted PredictAD tool. *PLoS One* 2013;**8**(2):e55246.

Liu 2014a {published data only}

Liu F, Wee CY, Chen H, Shen D. Inter-modality relationship constrained multi-modality multi-task feature selection for Alzheimer's disease and mild cognitive impairment identification. *NeuroImage* 2014;**84**:466-75.

Liu 2014b {published data only}

Liu M, Zhang D, Shen D, Alzheimer's Disease Neuroimaging Initiative. Identifying informative imaging biomarkers via tree structured sparse learning for AD diagnosis. *Neuroinformatics* 2014;**12**(3):381-94.

Llano 2011 {published data only}

Llano DA, Laforet G, Devanarayan V, Alzheimer's Disease Neuroimaging Initiative. Derivation of a new ADAS-cog composite using tree-based multivariate analysis: prediction of conversion from mild cognitive impairment to Alzheimer disease. *Alzheimer Disease and Associated Disorders* 2011;**25**(1):73-84.

Long 2016 {published data only}

Long Z, Jing B, Yan H, Dong J, Liu H, Mo X, et al. A support vector machine-based method to identify mild cognitive impairment with multi-level characteristics of magnetic resonance imaging. *Neuroscience* 2016;**331**:169-76.

Lopez 2016 {published data only}

López ME, Turrero A, Cuesta P, López-Sanz D, Bruña R, Marcos A, et al. Searching for primary predictors of conversion from mild cognitive impairment to Alzheimer's disease: a multivariate follow-up study. *Journal of Alzheimer's Disease* 2016;**52**(1):133-43.

Luo 2016 {published data only}

Luo Y, Cao Z, Liu Y, Wu L, Shan H, Liu Y, et al. T2 signal intensity and volume abnormalities of hippocampal subregions in patients with amnesic mild cognitive impairment by magnetic resonance imaging. *International Journal of Neuroscience* 2016;**126**(10):904-11.

Ma 2016 {published data only}

Ma X, Li Z, Jing B, Liu H, Li D, Li H, Alzheimer's Disease Neuroimaging Initiative. Identify the atrophy of Alzheimer's disease, mild cognitive impairment and normal aging using morphometric MRI analysis. *Frontiers in Aging Neuroscience* 2016;**8**:243.

MacDonald 2013 {published data only}

Macdonald KE, Bartlett JW, Leung KK, Ourselin S, Barnes J, ADNI investigators. The value of hippocampal and temporal horn volumes and rates of change in predicting future conversion to AD. *Alzheimer Disease and Associated Disorders* 2013;**27**(2):168-73.

Mah 2015 {published data only}

Mah L, Binns MA, Steffens DC, Alzheimer's Disease Neuroimaging Initiative. Anxiety symptoms in amnesic mild cognitive impairment are associated with medial temporal atrophy and predict conversion to Alzheimer disease. *American Journal of Geriatric Psychiatry* 2015;**23**(5):466-76.

Mangialasche 2013 {published data only}

Mangialasche F, Westman E, Kivipelto M, Muehlboeck JS, Cecchetti R, Baglioni M, et al. AddNeuroMed consortium. Classification and prediction of clinical diagnosis of Alzheimer's disease based on MRI and plasma measures of α - γ -tocotrienols and γ -tocopherol. *Journal of Internal Medicine* 2013;**273**(6):602-21.

Manning 2014 {published data only}

Manning EN, Barnes J, Cash DM, Bartlett JW, Leung KK, Ourselin S, et al. Alzheimer's Disease Neuroimaging Initiative. APOE ϵ 4 is associated with disproportionate progressive hippocampal atrophy in AD. *PLoS One* 2014;**9**(5):e97608.

Martínez-Torteya 2015 {published data only}

Martínez-Torteya A, Treviño V, Tamez-Peña JG. Improved diagnostic multimodal biomarkers for Alzheimer's disease and mild cognitive impairment. *BioMed Research International* 2015;**2015**:961314.

Maruyama 2004 {published data only}

Maruyama M, Matsui T, Tanji H, Nemoto M, Tomita N, Ootsuki M, et al. Cerebrospinal fluid tau protein and periventricular white matter lesions in patients with mild cognitive impairment: implications for 2 major pathways. *Archives of Neurology* 2004;**61**(5):716-20. [PUBMED: 15148149]

Mascalchi 2016 {published data only}

Mascalchi M, Bessi V, Ginestroni A, Toschi N, Ciulli S, Padiglioni S, et al. Low MT ratio in hippocampus of amnesic MCI patients who will progress to AD. *Journal of Neurology* 2016;**263**(5):1024-6.

Massaro 2004 {published data only}

Massaro JM, D'Agostino RB Sr, Sullivan LM, Beiser A, DeCarli C, Au R, et al. Managing and analysing data from a large-scale study on Framingham offspring relating brain structure to cognitive function. *Statistics in Medicine* 2004;**23**(2):351-67. [PUBMED: 14716734]

McEvoy 2009 {published data only}

McEvoy LK, Fennema-Notestine C, Roddey JC, Hagler DJ Jr, Holland D, Karow DS, et al. Alzheimer's Disease Neuroimaging Initiative. Alzheimer disease: quantitative structural neuroimaging for detection and prediction of clinical and structural changes in mild cognitive impairment. *Radiology* 2009;**251**(1):195-205.

McEvoy 2011 {published data only}

McEvoy LK, Blennow K, Brewer J, Dale A, Heister D. Predicting progression to Alzheimer's disease in MCI using combined structural atrophy and CSF biomarkers. *Alzheimer's & Dementia* 2011;**7** Suppl 4:3-4.

Meguro 2016 {published data only}

Meguro K, Akanuma K, Meguro M, Yamaguchi S, Ishii H, Tashiro M. Prevalence and prognosis of prodromal Alzheimer's disease as assessed by magnetic resonance imaging and 18F-fluorodeoxyglucose-positron emission tomography in a community: reanalysis from the Osaka-Tajiri Project. *Psychogeriatrics* 2016;**16**(2):116-20.

Meyer 2005a {published data only}

Meyer JS, Huang J, Chowdhury M. MRI abnormalities associated with mild cognitive impairments of vascular (VMCI) versus neurodegenerative (NMCI) types prodromal for vascular and Alzheimer's dementias. *Current Alzheimer Research* 2005;**2**(5):579-85. [PUBMED: 16375661]

Meyer 2005b {published data only}

Meyer JS, Quach M, Thornby J, Chowdhury M, Huang J. MRI identifies MCI subtypes: vascular versus neurodegenerative. *Journal of the Neurological Sciences* 2005;**229-230**:121-9. [PUBMED: 15760630]

Meyer 2007 {published data only}

Meyer JS, Huang J, Chowdhury MH. MRI confirms mild cognitive impairments prodromal for Alzheimer's, vascular and Parkinson-Lewy body dementias. *Journal of the Neurological Sciences* 2007;**257**(1-2):97-104. [PUBMED: 17316690]

Miller 2008 {published data only}

Miller SL, Fenstermacher E, Bates J, Blacker D, Sperling RA, Dickerson BC. Hippocampal activation in adults with mild cognitive impairment predicts subsequent cognitive decline. *Journal of Neurology, Neurosurgery, and Psychiatry* 2008;**79**(6):630-5. [PUBMED: 17846109]

Minhas 2017 {published data only}

Minhas S, Khanum A, Riaz F, Alvi A, Khan SA. A nonparametric approach for mild cognitive impairment to AD conversion prediction: results on longitudinal data. *IEEE Journal of Biomedical and Health Informatics* 2017;**21**(5):1403-10.

Moradi 2015 {published data only}

Moradi E, Pepe A, Gaser C, Huttunen H, Tohka J, Alzheimer's Disease Neuroimaging Initiative. Machine learning framework for early MRI-based Alzheimer's conversion prediction in MCI subjects. *NeuroImage* 2015;**104**:398-412.

Moradi 2016 {published data only}

Moradi E, Hallikainen I, Hänninen T, Tohka J, Alzheimer's Disease Neuroimaging Initiative. Rey's auditory verbal learning test scores can be predicted from whole brain MRI in Alzheimer's disease. *NeuroImage: Clinical* 2016;**13**:415-27.

Moretti 2015 {published data only}

Moretti DV. Understanding early dementia: EEG, MRI, SPECT and memory evaluation. *Translational Neuroscience* 2015;**6**(1):32-46.

Morra 2009 {published data only}

Morra JH, Tu Z, Apostolova LG, Green AE, Avedissian C, Madsen SK, et al. Alzheimer's Disease Neuroimaging Initiative. Automated mapping of hippocampal atrophy in 1-year repeat MRI data from 490 subjects with Alzheimer's disease, mild cognitive impairment, and elderly controls. *NeuroImage* 2009;**45** Suppl 1:S3-15.

Mubeen 2017 {published data only}

Mubeen AM, Asaei A, Bachman AH, Sidtis JJ, Ardekani BA, Alzheimer's Disease Neuroimaging Initiative. A six-month longitudinal evaluation significantly improves accuracy of predicting incipient Alzheimer's disease in mild cognitive impairment. *Journal of Neuroradiology. Journal de Neuroradiologie* 2017;**44**(6):381-7.

Mungas 2002 {published data only}

Mungas D, Reed BR, Jagust WJ, DeCarli C, Mack WJ, Kramer JH, et al. Volumetric MRI predicts rate of cognitive decline related to AD and cerebrovascular disease. *Neurology* 2002;**59**(6):867-73. [PUBMED: 12297568]

Mungas 2005 {published data only}

Mungas D, Harvey D, Reed BR, Jagust WJ, DeCarli C, Beckett L, et al. Longitudinal volumetric MRI change and rate of cognitive decline. *Neurology* 2005;**65**(4):565-71. [PUBMED: 16116117]

Nesteruk 2015 {published data only}

Nesteruk M, Nesteruk T, Styczyńska M, Barczak A, Mandecka M, Walecki J, et al. Predicting the conversion of mild cognitive impairment to Alzheimer's disease based on the volumetric measurements of the selected brain structures in magnetic resonance imaging. *Neurologia i Neurochirurgia Polska* 2015;**49**(6):349-53.

Nordlund 2005 {published data only}

Nordlund A, Rolstad S, Hellström P, Sjögren M, Hansen S, Wallin A. The Goteborg MCI study: mild cognitive impairment is

- a heterogeneous condition. *Journal of Neurology, Neurosurgery, and Psychiatry* 2005;**76**(11):1485-90.
- Ota 2015** {published data only}
 Ota K, Oishi N, Ito K, Fukuyama H, SEAD-J Study Group, Alzheimer's Disease Neuroimaging Initiative. Effects of imaging modalities, brain atlases and feature selection on prediction of Alzheimer's disease. *Journal of Neuroscience Methods* 2015;**256**:168-83.
- Ota 2016** {published data only}
 Ota K, Oishi N, Ito K, Fukuyama H, SEAD-J Study Group, Alzheimer's Disease Neuroimaging Initiative. Prediction of Alzheimer's disease in amnesic mild cognitive impairment subtypes: stratification based on imaging biomarkers. *Journal of Alzheimer's Disease* 2016;**52**(4):1385-401.
- Overdorp 2014** {published data only}
 Overdorp EJ, Kessels RP, Claassen JA, Oosterman JM. Cognitive impairments associated with medial temporal atrophy and white matter hyperintensities: an MRI study in memory clinic patients. *Frontiers in Aging Neuroscience* 2014;**6**:98.
- Park 2013** {published data only}
 Park H, Yang JJ, Seo J, Lee JM, Alzheimer's Disease Neuroimaging Initiative. Dimensionality reduced cortical features and their use in predicting longitudinal changes in Alzheimer's disease. *Neuroscience Letters* 2013;**550**:17-22.
- Park 2015** {published data only}
 Park MH, Han C. Is there an MCI reversion to cognitively normal? Analysis of Alzheimer's disease biomarkers profiles. *International Psychogeriatrics* 2015;**27**(3):429-37.
- Peng 2015** {published data only}
 Peng GP, Feng Z, He FP, Chen ZQ, Liu XY, Liu P, et al. Correlation of hippocampal volume and cognitive performances in patients with either mild cognitive impairment or Alzheimer's disease. *CNS Neuroscience & Therapeutics* 2015;**21**(1):15-22.
- Perani 2015** {published data only}
 Perani D, Cerami C, Caminiti SP, Santangelo R, Coppi E, Ferrari L, et al. Cross-validation of biomarkers for the early differential diagnosis and prognosis of dementia in a clinical setting. *European Journal of Nuclear Medicine and Molecular Imaging* 2016;**43**(3):499-508.
- Persson 2017** {published data only}
 Persson K, Barca ML, Eldholm RS, Cavallin L, Šaltytė Benth J, Selbæk G, et al. Visual evaluation of medial temporal lobe atrophy as a clinical marker of conversion from mild cognitive impairment to dementia and for predicting progression in patients with mild cognitive impairment and mild Alzheimer's disease. *Dementia and Geriatric Cognitive Disorders* 2017;**44**(1-2):12-24.
- Peters 2014** {published data only}
 Peters F, Villeneuve S, Belleville S. Predicting progression to dementia in elderly subjects with mild cognitive impairment using both cognitive and neuroimaging predictors. *Journal of Alzheimer's Disease* 2014;**38**(2):307-18.
- Petersen 2010** {published data only}
 Petersen RC, Aisen PS, Beckett LA, Donohue MC, Gamst AC, Harvey DJ, et al. Alzheimer's Disease Neuroimaging Initiative (ADNI): clinical characterization. *Neurology* 2010;**74**(3):201-9.
- Prasad 2011** {published data only}
 Prasad K, Wiryasaputra L, Ng A, Kandiah N. White matter disease independently predicts progression from mild cognitive impairment to Alzheimer's disease in a clinic cohort. *Dementia and Geriatric Cognitive Disorders* 2011;**31**(6):431-4.
- Prestia 2015** {published data only}
 Prestia A, Caroli A, Wade SK, Van der Flier WM, Ossenkoppele R, Van Berckel B, et al. Prediction of AD dementia by biomarkers following the NIA-AA and IWG diagnostic criteria in MCI patients from three European memory clinics. *Alzheimer's & Dementia* 2015;**11**(10):1191-201.
- Prins 2013** {published data only}
 Prins ND, Van der Flier WM, Brashear HR, Knol DL, Van de Pol LA, Barkhof F, et al. Predictors of progression from mild cognitive impairment to dementia in the placebo-arm of a clinical trial population. *Journal of Alzheimer's Disease* 2013;**36**(1):79-85.
- Qiu 2014** {published data only}
 Qiu Y, Li L, Zhou TY, Lu W, Alzheimer's Disease Neuroimaging Initiative. Alzheimer's disease progression model based on integrated biomarkers and clinical measures. *Acta Pharmacologica Sinica* 2014;**35**(9):1111-20.
- Querbes 2009** {published data only}
 Querbes O, Aubry F, Pariente J, Lotterie JA, Demonet JF, Duret V, et al. Early diagnosis of Alzheimer's disease using cortical thickness: impact of cognitive reserve. *Brain* 2009;**132**(Pt 8):2036-47. [PUBMED: 19439419]
- Raamana 2015** {published data only}
 Raamana PR, Weiner MW, Wang L, Beg MF, Alzheimer's Disease Neuroimaging Initiative. Thickness network features for prognostic applications in dementia. *Neurobiology of Aging* 2015;**36** Suppl 1:S91-S102.
- Rana 2017** {published data only}
 Rana AK, Sandu AL, Robertson KL, McNeil CJ, Whalley LJ, Staff RT, et al. A comparison of measurement methods of hippocampal atrophy rate for predicting Alzheimer's dementia in the Aberdeen Birth Cohort of 1936. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring* 2016;**6**:31-9.
- Redolfi 2015** {published data only}
 Redolfi A, Manset D, Barkhof F, Wahlund LO, Glatard T, Mangin JF, et al. Head-to-head comparison of two popular cortical thickness extraction algorithms: a cross-sectional and longitudinal study. *PLoS One* 2015;**10**(3):e0117692.
- Richard 2013** {published data only}
 Richard E, Schmand BA, Eikelenboom P, Van Gool WA, Alzheimer's Disease Neuroimaging Initiative. MRI and cerebrospinal fluid biomarkers for predicting progression to Alzheimer's disease in patients with mild cognitive impairment: a diagnostic accuracy study. *BMJ Open* 2013;**3**(6):e002541.

Risacher 2010 {published data only}

Risacher SL, Shen L, West JD, Kim S, McDonald BC, Beckett LA, et al. Longitudinal MRI atrophy biomarkers: relationship to conversion in the ADNI cohort. *Neurobiology of Aging* 2010;**31**(8):1401-18. [PUBMED: 20620664]

Ritter 2016 {published data only}

Ritter K, Lange C, Weygandt M, Mäurer A, Roberts A, Estrella M, et al. Combination of structural MRI and FDG-PET of the brain improves diagnostic accuracy in newly manifested cognitive impairment in geriatric inpatients. *Journal of Alzheimer's Disease* 2016;**54**(4):1319-31.

Runtti 2014 {published data only}

Runtti H, Mattila J, Van Gils M, Koikkalainen J, Soininen H, Lötjönen J, Alzheimer's Disease Neuroimaging Initiative. Quantitative evaluation of disease progression in a longitudinal mild cognitive impairment cohort. *Journal of Alzheimer's Disease* 2014;**39**(1):49-61.

Salvatore 2018 {published data only}

Salvatore C, Cerasa A, Castiglioni I. MRI characterizes the progressive course of AD and predicts conversion to Alzheimer's dementia 24 months before probable diagnosis. *Frontiers in Aging Neuroscience* 2018;**10**:135.

Sambuchi 2015 {published data only}

Sambuchi N, Muraccioli I, Alescio-Lautier B, Paban V, Sambuc R, Jouve É, et al. Subjective cognitive impairment and Alzheimer's disease: a two year follow up of 51 subjects during two years [Subjective cognitive impairment et maladie d'Alzheimer: étude d'une cohorte de 51 sujets suivis sur deux ans]. *Geriatric et Psychologie Neuropsychiatrie du Vieillessement* 2015;**13**(4):462-71.

Schmitter 2014 {published data only}

Schmitter D, Roche A, Maréchal B, Ribes D, Abdulkadir A, Bach-Cuadra M, et al. Alzheimer's Disease Neuroimaging Initiative. An evaluation of volume-based morphometry for prediction of mild cognitive impairment and Alzheimer's disease. *NeuroImage: Clinical* 2014;**7**:7-17.

Schuff 2009 {published data only}

Schuff N, Woerner N, Boreta L, Kornfield T, Shaw LM, Trojanowski JQ, et al. MRI of hippocampal volume loss in early Alzheimer's disease in relation to ApoE genotype and biomarkers. *Brain* 2009;**132**(Pt 4):1067-77. [PUBMED: 19251758]

Shaffer 2013 {published data only}

Shaffer JL, Petrella JR, Sheldon FC, Choudhury KR, Calhoun VD, Coleman RE, et al. Predicting cognitive decline in subjects at risk for Alzheimer disease by using combined cerebrospinal fluid, MR imaging, and PET biomarkers. *Radiology* 2013;**266**(2):583-91.

Sheng 2017 {published data only}

Sheng C, Xia M, Yu H, Huang Y, Lu Y, Liu F, et al. Abnormal global functional network connectivity and its relationship to medial temporal atrophy in patients with amnesic mild cognitive impairment. *PLoS One* 2017;**12**(6):e0179823.

Sluimer 2009 {published data only}

Sluimer JD, Van der Flier WM, Karas GB, Van Schijndel R, Barnes J, Boyes RG, et al. Accelerating regional atrophy rates in the progression from normal aging to Alzheimer's disease. *European Radiology* 2009;**19**(12):2826-33. [PUBMED: 19618189]

Smith 2008 {published data only}

Smith EE, Egorova S, Blacker D, Killiany RJ, Muzikansky A, Dickerson BC, et al. Magnetic resonance imaging white matter hyperintensities and brain volume in the prediction of mild cognitive impairment and dementia. *Archives of Neurology* 2008;**65**(1):94-100. [PUBMED: 18195145]

Sohn 2015 {published data only}

Sohn BK, Yi D, Seo EH, Choe YM, Kim JW, Kim SG, et al. Comparison of regional gray matter atrophy, white matter alteration, and glucose metabolism as a predictor of the conversion to Alzheimer's disease in mild cognitive impairment. *Journal of Korean Medical Science* 2015;**30**(6):779-87.

Song 2013 {published data only}

Song X, Mitnitski A, Zhang N, Chen W, Rockwood K, Alzheimer's Disease Neuroimaging Initiative. Dynamics of brain structure and cognitive function in the Alzheimer's Disease Neuroimaging Initiative. *Journal of Neurology, Neurosurgery, and Psychiatry* 2013;**84**(1):71-8.

Sousa 2015 {published data only}

Sousa A, Gomar JJ, Goldberg TE. Neural and behavioral substrates of disorientation in mild cognitive impairment and Alzheimer's disease. *Alzheimer's and Dementia: Translational Research and Clinical Interventions* 2015;**1**(1):37-45.

Sousa 2016 {published data only}

Sousa A, Gomar JJ, Ragland JD, Conejero-Goldberg C, Buthorn J, Keehlisen L, et al. The relational and item-specific encoding task in mild cognitive impairment and Alzheimer disease. *Dementia and Geriatric Cognitive Disorders* 2016;**42**(5-6):265-77.

Spulber 2010 {published data only}

Spulber G, Niskanen E, MacDonald S, Smilovici O, Chen K, Reiman EM, et al. Whole brain atrophy rate predicts progression from MCI to Alzheimer's disease. *Neurobiology of Aging* 2010;**31**(9):1601-5. [PUBMED: 18829136]

Spulber 2013 {published data only}

Spulber G, Simmons A, Muehlboeck JS, Mecocci P, Vellas B, Tsolaki M, et al. An MRI-based index to measure the severity of Alzheimer's disease-like structural pattern in subjects with mild cognitive impairment. *Journal of Internal Medicine* 2013;**273**(4):396-409.

Staekenborg 2009 {published data only}

Staekenborg SS, Koedam EL, Henneman WJ, Stokman P, Barkhof F, Scheltens P, et al. Progression of mild cognitive impairment to dementia: contribution of cerebrovascular disease compared with medial temporal lobe atrophy. *Stroke: a Journal of Cerebral Circulation* 2009;**40**(4):1269-74. [PUBMED: 19228848]

Stephan 2015 {published data only}

Stephan BC, Tzourio C, Auriacombe S, Amieva H, Dufouil C, Alperovitch A, et al. Usefulness of data from magnetic resonance imaging to improve prediction of dementia: a population based cohort study. *BMJ* 2015;**350**:h2863.

Stonnington 2018 {published data only}

Stonnington CM, Chen Y, Savage CR, Lee W, Bauer RJ 3rd, Sharieff S, et al. Predicting imminent progression to clinically significant memory decline using volumetric MRI and FDG PET. *Journal of Alzheimer's Disease* 2018;**63**(2):603-15.

Stoub 2005 {published data only}

Stoub TR, Bulgakova M, Leurgans S, Bennett DA, Fleischman D, Turner DA, et al. MRI predictors of risk of incident Alzheimer disease: a longitudinal study. *Neurology* 2005;**64**(9):1520-4. [PUBMED: 15883311]

Suk 2014 {published data only}

Suk HI, Lee SW, Shen D, Alzheimer's Disease Neuroimaging Initiative. Hierarchical feature representation and multimodal fusion with deep learning for AD/MCI diagnosis. *NeuroImage* 2014;**101**:569-82.

Suppa 2015a {published data only}

Suppa P, Hampel H, Spies L, Fiebach JB, Dubois B, Buchert R, Alzheimer's Disease Neuroimaging Initiative. Fully automated atlas-based hippocampus volumetry for clinical routine: validation in subjects with mild cognitive impairment from the ADNI cohort. *Journal of Alzheimer's Disease* 2015;**46**(1):199-209.

Suppa 2015b {published data only}

Suppa P, Anker U, Spies L, Bopp I, Ruegger-Frey B, Klaghofer R, et al. Fully automated atlas-based hippocampal volumetry for detection of Alzheimer's disease in a memory clinic setting. *Journal of Alzheimer's Disease* 2015;**44**(1):183-93.

Susanto 2015 {published data only}

Susanto TA, Pua EP, Zhou J, Alzheimer's Disease Neuroimaging Initiative. Cognition, brain atrophy, and cerebrospinal fluid biomarkers changes from preclinical to dementia stage of Alzheimer's disease and the influence of apolipoprotein e. *Journal of Alzheimer's Disease* 2015;**45**(1):253-68.

Sørensen 2016 {published data only}

Sørensen L, Igel C, Liv Hansen N, Osler M, Lauritzen M, Rostrup E, et al. Alzheimer's Disease Neuroimaging Initiative, Australian Imaging Biomarkers and Lifestyle Flagship Study of Ageing. Early detection of Alzheimer's disease using MRI hippocampal texture. *Human Brain Mapping* 2016;**37**(3):1148-61.

Tang 2014 {published data only}

Tang X, Holland D, Dale AM, Younes L, Miller MI, Alzheimer's Disease Neuroimaging Initiative. Shape abnormalities of subcortical and ventricular structures in mild cognitive impairment and Alzheimer's disease: detecting, quantifying, and predicting. *Human Brain Mapping* 2014;**35**(8):370-25.

Tang 2015 {published data only}

Tang X, Holland D, Dale AM, Younes L, Miller MI. Baseline shape diffeomorphometry patterns of subcortical and ventricular structures in predicting conversion of mild cognitive impairment to Alzheimer's disease. *Journal of Alzheimer's Disease* 2015;**44**(2):599-611.

Tapiola 2008 {published data only}

Tapiola T, Pennanen C, Tapiola M, Tervo S, Kivipelto M, Hanninen T, et al. MRI of hippocampus and entorhinal cortex in mild cognitive impairment: a follow-up study. *Neurobiology of Aging* 2008;**29**(1):31-8. [PUBMED: 17097769]

Tarnanas 2014 {published data only}

Tarnanas I, Tsolaki M, Nef T, M Müri R, Mosimann UP. Can a novel computerized cognitive screening test provide additional information for early detection of Alzheimer's disease?. *Alzheimer's & Dementia* 2014;**10**(6):790-8.

Teipel 2015 {published data only}

Teipel SJ, Kurth J, Krause B, Grothe MJ, Alzheimer's Disease Neuroimaging Initiative. The relative importance of imaging markers for the prediction of Alzheimer's disease dementia in mild cognitive impairment - beyond classical regression. *NeuroImage: Clinical* 2015;**8**:583-93.

Ten Kate 2017a {published data only}

Ten Kate M, Barkhof F, Visser PJ, Teunissen CE, Scheltens P, Van der Flier WM, et al. Amyloid-independent atrophy patterns predict time to progression to dementia in mild cognitive impairment. *Alzheimer's Research & Therapy* 2017;**9**(1):73.

Tosun 2010 {published data only}

Tosun D, Schuff N, Truran-Sacrey D, Shaw LM, Trojanowski JQ, Aisen P, et al. Relations between brain tissue loss, CSF biomarkers, and the ApoE genetic profile: a longitudinal MRI study. *Neurobiology of Aging* 2010;**31**(8):1340-54. [PUBMED: 20570401]

Trzepacz 2014 {published data only}

Trzepacz PT, Yu P, Sun J, Schuh K, Case M, Witte MM, et al. Alzheimer's Disease Neuroimaging Initiative. Comparison of neuroimaging modalities for the prediction of conversion from mild cognitive impairment to Alzheimer's dementia. *Neurobiology of Aging* 2014;**35**(1):143-51.

Trzepacz 2016 {published data only}

Trzepacz PT, Hochstetler H, Yu P, Castelluccio P, Witte MM, Dell'Agnello G, et al. Alzheimer's Disease Neuroimaging Initiative. Relationship of hippocampal volume to amyloid burden across diagnostic stages of Alzheimer's disease. *Dementia and Geriatric Cognitive Disorders* 2016;**41**(1-2):68-79.

Van Maurik 2016 {published data only}

Van Maurik IS, Bouwman FH, Teunissen CE, Scheltens P, Barkhof F, Wattjes M, et al. Personalized risk estimates for MCI patients: taking biomarkers into the clinic. *Alzheimer's & Dementia* 2016;**12** Suppl 7:P393.

Vannini 2007 {published data only}

Vannini P, Almkvist O, Dierks T, Lehmann C, Wahlund LO. Reduced neuronal efficacy in progressive mild cognitive impairment: a prospective fMRI study on visuospatial processing. *Psychiatry Research* 2007;**156**(1):43-57. [PUBMED: 17719211]

Van Rossum 2012 {published data only}

Van Rossum IA, Vos SJ, Burns L, Knol DL, Scheltens P, Soininen H, et al. Injury markers predict time to dementia in subjects with MCI and amyloid pathology. *Neurology* 2012;**79**(17):1809-16.

Varon 2011 {published data only}

Varon D, Loewenstein DA, Potter E, Greig MT, Agron J, Shen Q, et al. Minimal atrophy of the entorhinal cortex and hippocampus: progression of cognitive impairment. *Dementia and Geriatric Cognitive Disorders* 2011;**31**(4):276-83.

Varon 2015 {published data only}

Varon D, Barker W, Loewenstein D, Greig M, Bohorquez A, Santos I, et al. Alzheimer's Disease Neuroimaging Initiative. Visual rating and volumetric measurement of medial temporal atrophy in the Alzheimer's Disease Neuroimaging Initiative (ADNI) cohort: baseline diagnosis and the prediction of MCI outcome. *International Journal of Geriatric Psychiatry* 2015;**30**(2):192-200.

Vasta 2016 {published data only}

Vasta R, Augimeri A, Cerasa A, Nigro S, Gramigna V, Nonnis M, et al. for The Alzheimer's Disease Neuroimaging Initiative. Hippocampal subfield atrophies in converted and not-converted mild cognitive impairments patients by a Markov random fields algorithm. *Current Alzheimer Research* 2016;**13**(5):566-74.

Vemuri 2009 {published data only}

Vemuri P, Wiste HJ, Weigand SD, Shaw LM, Trojanowski JQ, Weiner MW, et al. MRI and CSF biomarkers in normal, MCI, and AD subjects: predicting future clinical change. *Neurology* 2009;**73**(4):294-301. [PUBMED: 19636049]

Verma 2018 {published data only}

Verma N, Beretvas SN, Pascual B, Masdeu JC, Markey MK, Alzheimer's Disease Neuroimaging Initiative. A biomarker combining imaging and neuropsychological assessment for tracking early Alzheimer's disease in clinical trials. *Current Alzheimer Research* 2018;**15**(5):429-42.

Villemagne 2013 {published data only}

Villemagne VL, Burnham S, Bourgeat P, Brown B, Ellis KA, Salvado, et al. Australian Imaging Biomarkers and Lifestyle (AIBL) Research Group. Amyloid β deposition, neurodegeneration, and cognitive decline in sporadic Alzheimer's disease: a prospective cohort study. *Lancet Neurology* 2013;**12**(4):357-67.

Vos 2012 {published data only}

Vos S, Van Rossum I, Burns L, Knol D, Scheltens P, Soininen H, et al. Test sequence of CSF and MRI biomarkers for

prediction of AD in subjects with MCI. *Neurobiology of Aging* 2012;**33**(10):2272-81.

Vos 2013 {published data only}

Vos SJ, Van Rossum IA, Verhey F, Knol DL, Soininen H, Wahlund LO, et al. Prediction of Alzheimer disease in subjects with amnesic and non amnesic MCI. *Neurology* 2013;**80**(12):1124-32.

Wahlund 2003 {published data only}

Wahlund LO, Blennow K. Cerebrospinal fluid biomarkers for disease stage and intensity in cognitively impaired patients. *Neuroscience Letters* 2003;**339**(2):99-102. [PUBMED: 12614904]

Walhovd 2010 {published data only}

Walhovd KB, Fjell AM, Brewer J, McEvoy LK, Fennema-Notestine C, Hagler DJ Jr, et al. Combining MR imaging, positron-emission tomography, and CSF biomarkers in the diagnosis and prognosis of Alzheimer disease. *American Journal of Neuroradiology* 2010;**31**(2):347-54. [PUBMED: 20075088]

Wang 2009 {published data only}

Wang H, Golob E, Bert A, Nie K, Chu Y, Dick MB, et al. Alterations in regional brain volume and individual MRI-guided perfusion in normal control, stable mild cognitive impairment, and MCI-AD converter. *Journal of Geriatric Psychiatry and Neurology* 2009;**22**(1):35-45. [PUBMED: 19150973]

Wang 2016 {published data only}

Wang S, Zhang Y, Liu G, Phillips P, Yuan TF. Detection of Alzheimer's disease by three-dimensional displacement field estimation in structural magnetic resonance imaging. *Journal of Alzheimer's Disease* 2016;**50**(1):233-48.

Wee 2013 {published data only}

Wee CY, Yap PT, Shen D. Prediction of Alzheimer's disease and mild cognitive impairment using cortical morphological patterns. *Human Brain Mapping* 2013;**34**(12):3411-25.

Wei 2016 {published data only}

Wei R, Li C, Fogelson N, Li L. Prediction of conversion from mild cognitive impairment to Alzheimer's disease using MRI and structural network features. *Frontiers in Aging Neuroscience* 2016;**8**:76.

Weise 2015 {published data only}

Weise D, Tiepolt S, Awissus C, Hoffmann KT, Lobsien D, Kaiser T, et al. Critical comparison of different biomarkers for Alzheimer's disease in a clinical setting. *Journal of Alzheimer's Disease* 2015;**48**(2):425-32.

Westman 2012 {published data only}

Westman E, Muehlboeck JS, Simmons A. Combining MRI and CSF measures for classification of Alzheimer's disease and prediction of mild cognitive impairment conversion. *NeuroImage* 2012;**62**(1):229-38.

Whitwell 2007 {published data only}

Whitwell JL, Przybelski SA, Weigand SD, Knopman DS, Boeve BF, Petersen RC, et al. 3D maps from multiple MRI illustrate changing atrophy patterns as subjects progress from mild

cognitive impairment to Alzheimer's disease. *Brain* 2007;**130**(Pt 7):1777-86. [PUBMED: 17533169]

Whitwell 2008 {published data only}

Whitwell JL, Shiung MM, Przybelski SA, Weigand SD, Knopman DS, Boeve BF, et al. MRI patterns of atrophy associated with progression to AD in amnesic mild cognitive impairment. *Neurology* 2008;**70**(7):512-20. [PUBMED: 17898323]

Willette 2014 {published data only}

Willette AA, Calhoun VD, Egan JM, Kapogiannis D, Alzheimer's Disease Neuroimaging Initiative. Prognostic classification of mild cognitive impairment and Alzheimer's disease: MRI independent component analysis. *Psychiatry Research* 2014;**224**(2):81-8.

Wolk 2009 {published data only}

Wolk DA, Price JC, Saxton JA, Snitz BE, James JA, Lopez OL, et al. Amyloid imaging in mild cognitive impairment subtypes. *Annals of Neurology* 2009;**65**(5):557-68. [PUBMED: 19475670]

Wolz 2010 {published data only}

Wolz R, Heckemann RA, Aljabar P, Hajnal JV, Hammers A, Lötjönen J, et al. Alzheimer's Disease Neuroimaging Initiative. Measurement of hippocampal atrophy using 4D graph-cut segmentation: application to ADNI. *NeuroImage* 2010;**52**(1):109-18.

Xu 2015 {published data only}

Xu L, Wu X, Chen K, Yao L. Multi-modality sparse representation-based classification for Alzheimer's disease and mild cognitive impairment. *Computer Methods and Programs in Biomedicine* 2015;**122**(2):182-90.

Xu 2016 {published data only}

Xu L, Wu X, Li R, Chen K, Long Z, Zhang J, et al. Alzheimer's Disease Neuroimaging Initiative. Prediction of progressive mild cognitive impairment by multi-modal neuroimaging biomarkers. *Journal of Alzheimer's Disease* 2016;**51**(4):1045-56.

Yamaguchi 2002 {published data only}

Yamaguchi S, Meguro K, Shimada M, Ishizaki J, Yamadori A, Sekita Y. Five-year retrospective changes in hippocampal atrophy and cognitive screening test performances in very mild Alzheimer's disease: the Tajiri Project. *Neuroradiology* 2002;**44**(1):43-8. [PUBMED: 11942499]

Yang 2012 {published data only}

Yang X, Tan MZ, Qiu A. CSF and brain structural imaging markers of the Alzheimer's pathological cascade. *PLoS One* 2012;**7**(12):e47406.

Ye 2012 {published data only}

Ye J, Farnum M, Yang E, Verbeeck R, Lobanov V, Raghavan N, et al. Alzheimer's Disease Neuroimaging Initiative. Sparse learning and stability selection for predicting MCI to AD conversion using baseline ADNI data. *BMC neurology* 2012;**12**:46.

Yi 2016 {published data only}

Yi HA, Möller C, Dieleman N, Bouwman FH, Barkhof F, Scheltens P, et al. Relation between subcortical grey matter

atrophy and conversion from mild cognitive impairment to Alzheimer's disease. *Journal of Neurology, Neurosurgery, and Psychiatry* 2016;**87**(4):425-32.

Young 2013 {published data only}

Young J, Modat M, Cardoso MJ, Mendelson A, Cash D, Ourselin S, Alzheimer's Disease Neuroimaging Initiative. Accurate multimodal probabilistic prediction of conversion to Alzheimer's disease in patients with mild cognitive impairment. *NeuroImage: Clinical* 2013;**2**:735-45.

Youssofzadeh 2017 {published data only}

Youssofzadeh V, McGuinness B, Maguire LP, Wong-Lin K. Multi-kernel learning with Dartel improves combined MRI-PET classification of Alzheimer's disease in AIBL data: group and individual analyses. *Frontiers in Human Neuroscience* 2017;**11**:380.

Yu 2012 {published data only}

Yu P, Dean RA, Hall SD, Qi Y, Sethuraman G, Willis BA, et al. Enriching amnesic mild cognitive impairment populations for clinical trials: optimal combination of biomarkers to predict conversion to dementia. *Journal of Alzheimer's Disease* 2012;**32**(2):373-85.

Yu 2014 {published data only}

Yu G, Liu Y, Thung KH, Shen D. Multi-task linear programming discriminant analysis for the identification of progressive MCI individuals. *PLoS One* 2014;**9**(5):e96458.

Yun 2015 {published data only}

Yun HJ, Kwak K, Lee JM, Alzheimer's Disease Neuroimaging Initiative. Multimodal discrimination of Alzheimer's disease based on regional cortical atrophy and hypometabolism. *PLoS One* 2015;**10**(6):e129250.

Zhang 2012a {published data only}

Zhang D, Shen D, Alzheimer's Disease Neuroimaging Initiative. Multi-modal multi-task learning for joint prediction of multiple regression and classification variables in Alzheimer's disease. *NeuroImage* 2012;**59**(2):895-907.

Zhang 2012b {published data only}

Zhang N, Song X, Zhang Y, Alzheimer's Disease Neuroimaging Initiative. Combining structural brain changes improves the prediction of Alzheimer's disease and mild cognitive impairment. *Dementia and Geriatric Cognitive Disorders* 2012;**33**(5):318-26.

Zheng 2015 {published data only}

Zheng W, Yao Z, Hu B, Gao X, Cai H, Moore P. Novel cortical thickness pattern for accurate detection of Alzheimer's disease. *Journal of Alzheimer's Disease* 2015;**48**(4):995-1008.

Zhou 2014 {published data only}

Zhou Q, Goryawala M, Cabrerizo M, Wang J, Barker W, Loewenstein DA, et al. An optimal decisional space for the classification of Alzheimer's disease and mild cognitive impairment. *IEEE Transactions on Biomedical Engineering* 2014;**61**(8):2245-53.

Zhou 2019 {published data only}

Zhou H, Jiang J, Lu J, Wang M, Zhang H, Zuo C, Alzheimer's Disease Neuroimaging Initiative. Dual-model radiomic biomarkers predict development of mild cognitive impairment progression to Alzheimer's disease. *Frontiers in Neuroscience* 2019;**12**:1045.

Additional references
ADI 2019

Alzheimer's Disease International. World Alzheimer Report 2019: Attitudes to Dementia. London: Alzheimer's Disease International, 2019.

ADNI 2010

Alzheimer's Disease Neuroimaging Initiative (ADNI). ADNI procedures, protocols and grants. www.adni-info.org (accessed 25 April 2018).

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 2000

American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, Text Revision (DSM-IV-TR). 4th Edition. Washington, DC: American Psychiatric Association, 2000.

American Psychiatric Association 2013

American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders DSM-5. Washington, DC: American Psychiatric Association, 2013.

Bobinski 2000

Bobinski M, de Leon MJ, Wegiel J, Desanti S, Convit A, Saint Louis LA, et al. The histological validation of post mortem magnetic resonance imaging determined hippocampal volume in Alzheimer's disease. *Neuroscience* 2000;**95**:721-5.

Bossuyt 2013

Bossuyt P, Davenport C, Deeks J, Hyde C, Leeflang M, Scholten R. Chapter 11: Interpreting results and drawing conclusions. In: Deeks JJ, Bossuyt PM, Gatsonis C (editors), *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 0.9*. The Cochrane Collaboration, 2013. Available from: srdta.cochrane.org. [Available from: <http://srdta.cochrane.org/>.]

Bäckman 2004

Bäckman L, Jones S, Berger AK, Laukka EJ, Small BJ. Multiple cognitive deficits during the transition to Alzheimer's disease. *Journal of Internal Medicine* 2004;**256**(3):195-204.

Carmichael 2005

Carmichael OT, Aizenstein HA, Davis SW, Becker JT, Thompson PM, Meltzer CC, et al. Atlas-based hippocampus segmentation in Alzheimer's disease and mild cognitive impairment. *NeuroImage* 2005;**27**(4):979-90. [PUBMED: 15990339]

De Vet 2008

De Vet HC, Eisinga A, Riphagen II, Aertgeerts B, Pewsner D. Chapter 7: Searching for Studies. In: *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 0.4* (September 2008). The Cochrane Collaboration, 2008: Section 7.2.1.2. Available from srdta.cochrane.org/handbook-dta-reviews.

deToledo-Morrell 1997

deToledo-Morrell L, Sullivan MP, Morrell F, Wilson RS, Bennett DA, Spencer S. Alzheimer's disease: in vivo detection of differential vulnerability of brain regions. *Neurobiology of Aging* 1997;**18**:438-63.

Doust 2005

Doust JA, Pietrzak E, Sanders S, Glasziou PP. Identifying studies for systematic reviews of diagnostic tests was difficult due to the poor sensitivity and precision of methodologic filters and the lack of information in the abstract. *Journal of Clinical Epidemiology* 2005;**58**(5):444-9. [PUBMED: 15845330]

Duvernoy 1988

Duvernoy H. The human hippocampus. An atlas of applied anatomy. München: JF Bergmann Verlag, 1988.

Erkinjuntti 2000

Erkinjuntti T, Inzitari D, Pantoni L, Wallin A, Scheltens P, Rockwood K. Research criteria for subcortical vascular dementia in clinical trials. *Journal of Neural Transmission* 2000;**Suppl**(59):23-30.

European Medicines Agency 2011

European Medicines Agency. Qualification opinion of low hippocampal volume (atrophy) by MRI for use in clinical trials for regulatory purpose - in pre-dementia stage of Alzheimer's disease. www.ema.europa.eu/docs/en_GB/document_library/Regulatory_and_procedural_guideline/2011/12/WC500118737.pdf (accessed 16 March 2019).

Fischl 2004

Fischl B, Salat DH, van der Kouwe AJ, Makris N, Ségonne F, Quinn BT, et al. Sequence-independent segmentation of magnetic resonance images. *NeuroImage* 2004;**23** **Suppl** 1.:S69-84.

Frisoni 2010a [pers comm]

Frisoni GB. Numbers needed to construct 2x2 tables for medial temporal lobe volume. Reference standard [personal communication]. Email to: F Pasquier and S Bombois 1 February 2010.

Frisoni 2010b [pers comm]

Frisoni GB. Numbers needed to construct 2x2 table for total hippocampal volume [personal communication]. Email to: H Malmgren and C Eckerström 9 February 2010.

Frisoni 2010c [pers comm]

Frisoni GB. Numbers needed to construct 2x2 table for total hippocampal volume, brain volume, ventricular volume [personal communication]. Email to: D Erten-Lyons 10 February 2010.

Frisoni 2010d [pers comm]

Frisoni GB. Numbers needed to construct 2x2 tables [personal communication]. Email to: LL Chao 20 March 2010.

Frisoni 2010e [pers comm]

Frisoni GB. Numbers needed to construct 2x2 table for total volumes of hippocampus and whole brain volume. Baseline characteristics of participants [personal communication]. Email to: WM van der Flier 30 March 2010.

Frisoni 2010f [pers comm]

Frisoni GB. Numbers needed to construct 2x2 table for the medial temporal lobe atrophy. Baseline patients' characteristics [personal communication]. Email to: GB Karas 30 March 2010.

Frisoni 2010g [pers comm]

Frisoni GB. Numbers needed to construct 2x2 table for total hippocampal volume and intracranial volume. Baseline patients' characteristics [personal communication]. Email to: CR Jack 23 April 2010.

Frisoni 2010h [pers comm]

Frisoni GB. Numbers needed to construct 2x2 table for lateral ventricular and whole brain volume [personal communication]. Email to: OT Charmichael 30 April 2010.

Frisoni 2010i [pers comm]

Frisoni GB. Numbers needed to construct 2x2 table for total volume of hippocampus [personal communication]. Email to: A Caroli 21 May 2010.

Frisoni 2010j [pers comm]

Frisoni GB. Numbers needed to construct 2x2 table for left, right hippocampal volume, left and right entorhinal volume, intracranial volume [personal communication]. Email to: DB Devanand 3 June 2010.

Frisoni 2010k [pers comm]

Frisoni GB. Numbers needed to construct 2x2 table for hippocampus and entorhinal cortex [personal communication]. Email to: T Tapiola 20 June 2010.

Frisoni 2010l [pers comm]

Frisoni GB. Numbers needed to construct 2x2 tables [personal communication]. Email to: L DeToledo 20 July 2010.

Frisoni 2010m [pers comm]

Frisoni GB. Numbers needed to construct 2x2 table. Baseline patients' characteristics [personal communication]. Email to: A Fellgiebel 20 July 2010.

Frisoni 2010n [pers comm]

Frisoni GB. Numbers needed to construct 2x2 table. Baseline patients' characteristics [personal communication]. Email to: C Geroldi 20 July 2010.

Frisoni 2010o [pers comm]

Frisoni GB. Numbers needed to construct 2x2 table [personal communication]. Email to: ES Korf 20 July 2010.

Frisoni 2010p [pers comm]

Frisoni GB. Numbers needed to construct 2x2 tables for hippocampus, parahippocampal gyrus, lateral temporal lobe and intracranial volumes. Baseline patients' characteristics [personal communication]. Email to: PJ Visser 2 September 2010.

Frisoni 2010q [pers comm]

Frisoni GB. Numbers needed to construct 2x2 table for left, right, total volumes of hippocampus and entorhinal cortex [personal communication]. Email to: SK Herukka 24 September 2010.

Frisoni 2010r [pers comm]

Frisoni GB. Numbers needed to construct 2x2 table for total volumes of hippocampus, amygdala volume and total intracranial volume. Baseline characteristics of participants [personal communication]. Email to: PN Wang 29 October 2010.

Frisoni 2010s [pers comm]

Frisoni GB. Numbers needed to construct 2x2 table for the hippocampal volume [personal communication]. Email to: CR Jack 29 October 2010.

Frisoni 2012 [pers comm]

Frisoni GB. Numbers needed to construct 2x2 table for the medial temporal lobe atrophy. Baseline patients' characteristics [personal communication]. Email to: BC Dickerson 20 June 2012.

Frisoni 2013

Frisoni GB, Bocchetta M, Chételat G, Rabinovici GD, de Leon MJ, Kaye J, et al. Imaging markers for Alzheimer disease: which vs how. *Neurology* 2013;**81**(5):487-500.

Frisoni 2016a [pers comm]

Frisoni GB. Numbers needed to construct 2x2 table. MRI technique. Baseline patients' characteristics [personal communication]. Email to: DE Barnes 19 October 2016.

Frisoni 2016b [pers comm]

Frisoni GB. Numbers needed to construct 2x2 table for the medial temporal lobe atrophy. Baseline patients' characteristics [personal communication]. Email to: R Duara 19 October 2016.

Frisoni 2016c [pers comm]

Frisoni GB. Numbers needed to construct 2x2 table for the medial temporal lobe atrophy. Baseline patients' characteristics [personal communication]. Email to: DB Devanand 19 October 2016.

Frisoni 2016d [pers comm]

Frisoni GB. Numbers needed to construct 2x2 table for hippocampal volume [personal communication]. Email to: A Prestia 20 October 2016.

Frisoni 2016e [pers comm]

Frisoni GB. Numbers needed to construct 2x2 table for the medial temporal lobe atrophy. Baseline patients' characteristics [personal communication]. Email to: SJ Vos 20 October 2016.

Frisoni 2016f [pers comm]

Frisoni GB. Numbers needed to construct 2x2 table [personal communication]. Email to: CR Jack 20 October 2016.

Frisoni 2017a

Frisoni GB, Boccardi M, Barkhof F, Blennow K, Cappa S, Chiotis K, et al. Strategic roadmap for an early diagnosis of Alzheimer's disease based on biomarkers. *Lancet Neurology* 2017;**16**(8):661-76.

Frisoni 2017b [pers comm]

Frisoni GB. Numbers needed to construct 2x2 tables for medial temporal lobe atrophy [personal communication]. Email to: HF Rhodius-Meester 27 February 2017.

Galton 2001

Galton CJ, Gomez-Anson B, Antoun N, Scheltens P, Patterson K, Graves M, et al. Temporal lobe rating scale: application to Alzheimer's disease and frontotemporal dementia. *Journal of Neurology, Neurosurgery, and Psychiatry* 2001;**70**(2):165-73.

Goncharova 2001

Goncharova II, Dickerson BC, Stoub TR, deToledo-Morrell L. MRI of entorhinal cortex: a reliable protocol for volumetric measurement. *Neurobiology of Aging* 2001;**22**:737-45.

GRADEpro GDT [Computer program]

McMaster University (developed by Evidence Prime). GRADEproGDT. Hamilton (ON): McMaster University (developed by Evidence Prime), accessed 29 January 2019.

Graham 1997

Graham JE, Rockwood K, Beattie BL, Eastwood R, Gauthier S, Tuokko H, et al. Prevalence and severity of cognitive impairment with and without dementia in an elderly population. *Lancet* 1997;**349**(9068):1793-6.

Hilden 1996

Hilden J, Glasziou P. Regret graphs, diagnostic uncertainty and Youden's Index. *Statistics in Medicine* 1996;**15**(10):969-86.

Jack 1992

Jack CR Jr, Petersen RC, O'Brien PC, Tangalos EG. MR-based hippocampal volumetry in the diagnosis of Alzheimer's disease. *Neurology* 1992;**42**:183-8. [PUBMED: 1734300]

Jack 2008b

Jack CR Jr, Bernstein MA, Fox NC, Thompson P, Alexander G, Harvey D, et al. The Alzheimer's Disease Neuroimaging Initiative (ADNI): MRI methods. *Journal of Magnetic Resonance Imaging* 2008;**4**:685-91.

Jack 2018

Jack CR Jr, Bennett DA, Blennow K, Carrillo MC, Dunn B, Haeberlein SB, et al. NIA-AA research framework: toward a biological definition of Alzheimer's disease. *Alzheimer's & Dementia* 2018;**14**(4):535-62.

Killiany 2002

Killiany RJ, Hyman BT, Gomez-Isla T, Moss MB, Kikinis R, Jolesz F, et al. MRI measures of entorhinal cortex vs hippocampus in preclinical AD. *Neurology* 2002;**58**:1188-96.

Leandrou 2018

Leandrou S, Petroudi S, Kyriacou PA, Reyes-Aldasoro CC, Pattichis CS. Quantitative MRI brain studies in mild cognitive impairment and Alzheimer's disease: a methodological review. *IEEE reviews in biomedical engineering* 2018;**11**:97-111.

Ledig 2015

Ledig C, Heckemann RA, Hammers A, Lopez JC, Newcombe VF, Makropoulos A, et al. Robust whole-brain segmentation: application to traumatic brain injury. *Medical Image Analysis* 2015;**21**:40-58.

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.

Lehéricy 1994

Lehéricy S, Baulac M, Chiras J, Piérot L, Martin N, Pillon B, et al. Amygdalohippocampal MR volume measurements in the early stages of Alzheimer disease. *American Journal of Neuroradiology* 1994;**15**(5):929-37. [PUBMED: 8059663]

Lijmer 1999

Lijmer JG, Mol BW, Heisterkamp S, Bossel GJ, Prins MH, Van derMeulen JH. Empirical evidence of design-related bias in studies of diagnostic tests. *Journal of the American Medical Association* 1999;**282**:1061-6.

Livingston 2017

Livingston G, Sommerlad A, Orgeta V, Costafreda SG, Huntley J, Ames D, et al. Dementia prevention, intervention, and care. *Lancet* 2017;**390**(10113):2673-734.

Lopez 2003

Lopez OL, Jagust WJ, Becker ST. Prevalence and classification of mild cognitive impairment in the Cardiovascular Health Study Cognition Study: Part 1. *Archives of Neurology* 2003;**60**:1385-9. [PUBMED: 14568808]

Lovestone 2009

Lovestone S, Francis P, Kloszewska I, Mecocci P, Simmons A, Soinen H, et al. AddNeuroMed—the European collaboration for the discovery of novel biomarkers for Alzheimer's disease. *Annals of the New York Academy of Sciences* 2009;**1180**:36-46.

Lötjönen 2011

Lötjönen J, Wolz R, Koikkalainen J, Julkunen V, Thurfjell L, Lundqvist R, et al. Fast and robust extraction of hippocampus

from MR images for diagnostics of Alzheimer's disease. *NeuroImage* 2011;**56**(1):185-96.

Macaskill 2010

Macaskill P, Gatsonis C, Deeks JJ, Harbord RM, Takwoingi Y. Chapter10: Analysing and presenting results. In: Deeks JJ, Bossuyt PM, Gatsonis C (editors), *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 1.0*. The Cochrane Collaboration, 2010. Available from: rdta.cochrane.org.

Martínez 2017

Martínez G, Vernooij RW, Fuentes Padilla P, Zamora J, Bonfill Cosp X, Flicker L. 18F PET with florbetapir for the early diagnosis of Alzheimer's disease dementia and other dementias in people with mild cognitive impairment (MCI). *Cochrane Database of Systematic Reviews* 2017, Issue 11. [DOI: [10.1002/14651858.CD012216.pub2](https://doi.org/10.1002/14651858.CD012216.pub2)]

McCleery 2019

McCleery J, Flicker L, Richard E, Quinn TJ. The National Institute on Aging and Alzheimer's Association research framework: a commentary from the Cochrane Dementia and Cognitive Improvement Group. *Alzheimer's & Dementia* 2019;**15**(1):179-81.

McKhaan 2011

McKhaan 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.

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.

Mitchell 2009

Mitchell AJ, Shiri-Feshki M. Rate of progression of mild cognitive impairment to dementia--meta-analysis of 41 robust inception cohort studies. *Acta Psychiatrica Scandinavica* 2009;**119**(4):252-65.

Morris 1993

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

NICE 2018

National Institute for Health and Clinical Excellence Clinical Practice Guideline NG97. Dementia: assessment, management and support for people living with dementia and their carers. Available from www.nice.org.uk/guidance/ng97 2018.

Noel-Storr 2014

Noel-Storr AH, McCleery JM, Richard E, Ritchie CW, Flicker L, Cullum SJ, et al. Reporting standards for studies of diagnostic test accuracy in dementia: the STARDDem Initiative. *Neurology* 2014;**83**(4):364-73.

Palmer 2008

Palmer K, Bäckman L, Winblad B, Fratiglioni L. Mild cognitive impairment in the general population: occurrence and progression to Alzheimer disease. *American Journal of Geriatric Psychiatry* 2008;**16**:603-11.

Pantel 2000

Pantel J, O'Leary DS, Cretsingher K, Bockholt HJ, Keefe H, Magnotta VA, et al. A new method for the in vivo volumetric measurement of the human hippocampus with high neuroanatomical accuracy. *Hippocampus* 2000;**10**(6):752-8.

Patenaude 2011

Patenaude B, Smith SM, Kennedy DN, Jenkinson M. A Bayesian model of shape and appearance for subcortical brain segmentation. *NeuroImage* 2011;**56**:907-22.

Payton 2018

Payton NM, Kalpouzos G, Rizzuto D, Fratiglioni L, Kivipelto M, Bäckman L, et al. Combining cognitive, genetic, and structural neuroimaging markers to identify individuals with increased dementia risk. *Journal of Alzheimer's disease* 2018;**64**(2):533-42.

Petersen 1999

Petersen RC, Smith GE, Waring SC, Ivnik RJ, Tangalos EG, Kokmen E. Mild cognitive impairment. *Archives of Neurology* 1999;**56**:303-8.

Petersen 2001

Petersen RC, Doody R, Kurz A, Mohs RC, Morris JC, Rabins PV, et al. Current concepts in mild cognitive impairment. *Archives of Neurology* 2001;**58**:1985-92.

Petersen 2004

Petersen RC. Mild cognitive impairment as a diagnostic entity. *Journal of Internal Medicine* 2004;**256**:183-94.

Petersen 2009

Petersen R, Knopman D, Boeve B, Geda YE, Ivnik RJ, Smith GE, et al. Mild cognitive impairment: ten years later. *Archives of Neurology* 2009;**66**:1447-55.

Pruessner 2002

Pruessner JC, Köhler S, Crane J, Pruessner M, Lord C, Byrne A, et al. Volumetry of temporo-polar, perirhinal, entorhinal and parahippocampal cortex from high-resolution MR images: considering the variability of the collateral sulcus. *Cerebral Cortex* 2002;**12**:1342-53.

Reisberg 1982

Reisberg B, Ferris SH, de Leon MJ, Crook T. The Global Deterioration Scale for assessment of primary degenerative dementia. *American Journal of Psychiatry* 1982;**139**:1136-9.

Ritchie 2014

Ritchie C, Smailagic N, Noel-Storr AH, Takwoingi Y, Flicker L, Mason SE, 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. [DOI: [10.1002/14651858.CD008782.pub4](https://doi.org/10.1002/14651858.CD008782.pub4)]

Ritchie 2017

Ritchie C, Smailagic N, Noel-Storr AH, Ukoumunne O, Ladds EC, Martin S. CSF tau and the CSF tau/ABeta ratio for the diagnosis of Alzheimer's disease dementia and other dementias in people with mild cognitive impairment (MCI). *Cochrane Database of Systematic Reviews* 2017, Issue 3. [DOI: [10.1002/14651858.CD010803.pub2](https://doi.org/10.1002/14651858.CD010803.pub2)]

Roberts 2010

Roberts JS, Karlawish JH, Uhlmann WR, Petersen RC, Green RC. Mild cognitive impairment in clinical care: a survey of American Academy of Neurology members. *Neurology* 2010;**75**(5):425-31.

Rutter 2001

Rutter CA, Gatsonis CA. A hierarchical regression approach to meta-analysis of diagnostic test accuracy evaluations. *Statistics in Medicine* 2001;**20**:2865-84.

Scheltens 1992

Scheltens P, Leys D, Barkhof F, Huglo D, Weinstein HC, Vermersch P, et al. Atrophy of medial temporal lobes on MRI in "probable" Alzheimer's disease and normal ageing: diagnostic value and neuropsychological correlates. *Journal of Neurology, Neurosurgery, and Psychiatry* 1992;**55**(10):967-72. [PUBMED: 1431963]

Scheltens 1997

Scheltens P, Pasquier F, Weerts J, Barkhof F, Leys D. Qualitative assessment of cerebral atrophy on MRI: inter- and intra-observer reproducibility in dementia and normal aging. *European Journal of Neurology* 1997;**37**:95-9.

Schunemann 2008

Schunemann HJ, Oxman AD, Brozek J, Glasziou P, Jaeschke R, Vist GE, et al. Grading quality of evidence and strength of recommendations for diagnostic tests and strategies. *BMJ* 2008;**336**:1106-10.

Schünemann 2016

Schünemann HJ, Mustafa R, Brozek J, Santesso N, AlonsoCoello P, Guyatt G, et al. GRADE Guidelines: 16. GRADE evidence to decision frameworks for tests in clinical practice and public health. *Journal of Clinical Epidemiology* 2016;**76**:89-98.

Simmons 2009

Simmons A, Westman E, Muehlboeck S, Mecocci P, Vellas B, Tsolaki M, et al. for the AddNeuroMed Consortium. MRI measures of Alzheimer's disease and the AddNeuroMed study. *Biomarkers in Brain Disease: the New York Academy of Sciences* 2009;**1180**:47-55.

Smailagic 2015

Smailagic N, Vacante M, Hyde C, Martin S, Ukoumunne O, Sachpekidis C. 18F-FDG 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* 2015, Issue 1. [DOI: [10.1002/14651858.CD010632.pub2](https://doi.org/10.1002/14651858.CD010632.pub2)]

Takwoingi 2010

Takwoingi Y, Deeks JJ. MetaDAS: A SAS macro for meta-analysis of diagnostic accuracy studies. User Guide Version 1.3.. Available from: <http://srdta.cochrane.org/> 2010 July.

Teipel 2013

Teipel S, Grothe M, Lista S, Toschi N, Francesco G, Garaci FG, et al. Relevance of magnetic resonance imaging for early detection and diagnosis of Alzheimer disease. *Medical Clinics of North America* 2013;**97**(3):399-424.

Ten Kate 2017b

Ten Kate M, Barkhof F, Boccardi M, Visser PJ, Jack CR Jr, Lovblad KO, et al. Clinical validity of medial temporal atrophy as a biomarker for Alzheimer's disease in the context of a structured 5-phase development framework. *Neurobiology of Aging* 2017;**52**:167-82.

The 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**(4):416-8.

US Food and Drug Administration 2015

US Food, Drug Administration. Biomarker letter of support for baseline low hippocampal volume measured by MRI as an exploratory prognostic biomarker for enrichment in clinical trials for Alzheimer's disease. www.fda.gov/downloads/Drugs/DevelopmentApprovalProcess/UCM439714.pdf (accessed 16 March 2019).

Van de Pol 2007

Van de Pol LA, Barnes J, Scahill RI, Frost C, Lewis EB, Boyes RG, et al. Improved reliability of hippocampal atrophy rate measurement in mild cognitive impairment using fluid registration. *NeuroImage* 2007;**34**(3):1036-41.

Van de Pol 2014

Van de Pol LA, Scheltens P. Medial temporal lobe atrophy scores translated to clinical practice: editorial comment on 'influence of age, disease onset and ApoE4 on visual medial temporal lobe atrophy cut-offs'. *Journal of Internal Medicine* 2014;**275**:331-3.

Van der Flier 2014

Van der Flier WM, Pijnenburg YA, Prins N, Lemstra AW, Bouwman FH, Teunissen CE, et al. Optimizing patient care and research: the Amsterdam dementia cohort. *Journal of Alzheimer's Disease* 2014;**41**:313-27.

Vernooij 2019

Vernooij MW, Pizzini FB, Schmidt R, Smits M, Yousry TA, Bargallo N, et al. Dementia imaging in clinical practice: a European-wide survey of 193 centres and conclusions by the ESNR working group. *Neuroradiology* 2019;**61**(6):633-42.

Visser 2008

Visser PJ, Verhey FR, Boada M, Bullock R, De Deyn PP, Frisoni GB, et al. Development of screening guidelines and clinical criteria for predementia Alzheimer's disease. The DESCRIPA study. *Neuroepidemiology* 2008;**30**(4):254-65.

Whiting 2004

Whiting P, Rutjes AW, Reitsma JB, Glas AS, Bossuyt PM, Kleijnen J. Sources of variation and bias in studies of diagnostic accuracy: a systematic review. *Annals of Internal Medicine* 2004;**140**:189-202.

Whiting 2011

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

Winblad 2004

Winblad B, Palmer K, Kivipelto M, Jelic V, Fratiglioni L, Wahlund L-O, 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**:240-6.

Yaffe 2006

Yaffe K, Petersen RC, Lindquist K, Kramer J, Miller B. Subtype of mild cognitive impairment and progression to dementia

and death. *Dementia and Geriatric Cognitive Disorders* 2006;**22**:312-9.

Youden 1950

Youden WJ. An index for rating diagnostic tests. *Cancer* 1950;**3**:32-5.

Yue 1997

Yue NC, Arnold AM, Longstreth WT Jr, Elster AD, Jungreis CA, O'Leary DH. Sulcal, ventricular, and white matter changes at MR imaging in the aging brain: data from the cardiovascular health study. *Radiology* 1997;**202**(1):33-9. [PUBMED: 8988189]

References to other published versions of this review
Filippini 2012

Filippini G, Casazza G, Bellatorre AG, Lista C, Duca P, Beecher D, et al. The role of MRI in the early diagnosis of Alzheimer's disease or other dementias in persons with mild cognitive impairment (MCI). *Cochrane Database of Systematic Reviews* 2012, Issue 2. [DOI: [10.1002/14651858.CD009628](https://doi.org/10.1002/14651858.CD009628)]

* Indicates the major publication for the study

CHARACTERISTICS OF STUDIES
Characteristics of included studies [ordered by study ID]
Carmichael 2007
Study characteristics

Patient sampling	<p>Primary objectives: to analyse several types of cognitive state transitions—normal to MCI, normal to dementia, MCI to dementia, in an attempt to capture relationships between ventricular structure and highly variable rates of cognitive decline in a population-based sample (the Pittsburgh Cardiovascular Health Study Cognition Study; Lopez 2003)</p> <p>Study population: participants with MCI</p> <p>Selection criteria: inclusion criteria were designed to capture different forms of cognitive impairment, with or without specific conditions that could themselves cause cognitive deficits. Exclusion criteria: not reported in detail; comorbid conditions were permitted for "possible MCI" (cerebrovascular disease, history of head trauma, encephalopathy, infectious disease, developmental disabilities, systemic illnesses that may cause cognitive deficits)</p> <p>Study design: prospective longitudinal</p>
Patient characteristics and setting	<p>Clinical presentation: MCI defined as "impairments in delayed recall of verbal material, nonverbal materials, or both; or impairment in at least 1 cognitive domain (other than memory); or 1 abnormal test (which could be a memory test) in at least 2 domains, without severe impairment of IADLs. The cognitive deficits must represent a decline from a previous level of functioning (Lopez 2003). Participants were classified as probable MCI when there were no psychiatric, neurologic (e.g. cerebrovascular disease, history of head trauma encephalopathy, infectious diseases, developmental disabilities), or systemic illnesses that may cause cognitive deficits, or possible when any comorbid condition was present".</p> <p>Age years mean (SD): MCI who progressed to dementia: 84 ± 4 years; stable MCI: 90 ± 5 years; MCI who reverted to normal: 84</p> <p>Gender (% men): MCI who progressed to dementia: 33%; stable MCI: 56%; reverted to normal: 11%</p> <p>Education (up to/beyond high school): MCI who progressed to dementia: 7/9; stable MCI: 3/9; reverted to normal: 0/1</p>

Carmichael 2007 (Continued)

ApoE4 carriers (%): not stated

Neuropsychological tests: modified mean (SD) MCI stable 91.17 ± 5.41 , MCI converts 92.44 ± 5.60

Clinical stroke excluded: not specified

Co-morbidities: participants with comorbidities were not excluded

Number enrolled: 130

Number available for analysis: 29 (amnesic MCI: 3; multiple cognitive deficit MCI: 26)

Setting: University of Pittsburgh

Country: USA

Period of study: 1997-2005

Language: English

Index tests

Index test: MRI automated method for estimation of lateral ventricular and whole-brain volume

Manufacturer: GE

Tesla strength: 1.5 Signa scanner with high-performance gradients

Assessment methods: automatic atlas-based segmentation ([Carmichael 2005](#)), for estimation of lateral ventricular volume (validated technique against visual and manual method); whole-brain mask was output by the skull-stripping procedure and was used to estimate brain volume in each image

Description of positive case definition by index test as reported: criteria for scoring lateral ventricle-to-brain ratio volume according to [Yue 1997](#) (normal value in healthy participants reported in the manuscript; cut-off value for normality not clearly reported; study authors recommended caution against use of the word "abnormal" for grades above the 95th percentile)

Examiners: trained neuroradiologists

Interobserver variability: not provided

Target condition and reference standard(s)

Target condition: AD

Prevalence of AD in the sample: 12/29 (41% of cases included in the analysis)

Stable MCI or converted to other dementia: 17/29 (59%) (12 stable MCI; 1 MCI reverted to normal; 4 MCI converted to other dementia)

Reference standard: NINCDS-ADRDA ([McKhann 1984](#))

Mean clinical follow-up: 3.2 years

Flow and timing

Withdrawals and losses to follow-up: 101 (78%) of the 130 enrolled participants excluded for the following: 99 participants excluded because they had not MRI follow-up, 2 of the remaining 31 participants were excluded and the reason was not reported.

Uninterpretable MRI results were not reported

Comparative

Key conclusions by the authors

Change rate in lateral ventricle-to-brain ratio was faster in participants who were had dementia or transitioned from MCI to dementia, compared with participants normal at follow-up images and participants who transitioned from normal to MCI or dementia

Conflict of interests

Study authors declared no conflict of interest

Carmichael 2007 (Continued)

Notes **Source of funding:** National Heart, Lung, and Blood Institute contracts N01-HC-85079 through N01-HC-85086, N01-HC-35129, N01 HC-15103, N01 HC-55222, and U01 HL080295; NIH grants NS07391, MH064625, AG05133, DA015900-01, MH01077, EB001561, RR019771, RR021813, AG016570, AG20098, and AG15928; and additional contribution from the National Institute of Neurological Disorders and Stroke

2 x 2 table: data to complete 2 x 2 table provided by the study authors

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		High	Low
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
Did the study provide a clear pre-specified definition of what was considered to be a "positive" result of the index test?	No		
Was the index test performed by a single operator or interpreted by consensus in a joint session?	Unclear		
		High	Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify	Yes		

Carmichael 2007 (Continued)

the target condition?

Were the reference standard results interpreted without knowledge of the results of the index tests? Yes

Low
Low
DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard? Yes

Did all patients receive the same reference standard? Yes

Were all patients included in the analysis? No

High
Caroli 2007
Study characteristics

Patient sampling **Primary objectives:** to use voxel-based analysis to find cerebral perfusion correlates of conversion to dementia in people with amnesic MCI

Study population: participants with amnesic MCI, either single-domain or multi-domain

Selection criteria: exclusion criteria: history of depression or psychosis of juvenile onset, clinical major stroke, alcohol abuse, craniocerebral trauma, or heavy use of psychotropic drugs. Participants were included if they agreed to undergo a SPECT-scan

Study design: prospective longitudinal

Patient characteristics and setting

Clinical presentation: amnesic MCI defined as "complaint of memory or other cognitive disturbances; MMSE: 24-27/30 or ≥ 28 plus low performance (score of 2/6 or higher) on the clock drawing test; sparing of IADL and ADL or functional impairment due to causes other than cognitive impairment"

Age years mean (SD): MCI who progressed to AD (AD): 69 ± 3 years; stable MCI: 71 ± 8 years

Gender (% men): MCI who progressed to AD: 56%; stable MCI: 58%

Education years mean (SD): MCI who progressed to AD: 11.4 ± 5.7 ; stable MCI: 8.6 ± 3.6

ApoE4 carriers (%): MCI who progressed to AD: 56%; stable MCI: 43%

Caroli 2007 (Continued)

Neuropsychological tests: MMSE mean (SD): MCI who progressed to AD: 26.8 ± 1.8; stable MCI: 27.0 ± 2.0

Clinical stroke excluded: history or neurological signs of major stroke was a cause of exclusion

Co-morbidities: participants with comorbidities such as hypertension, diabetes and heart disease were included

Number enrolled: 56

Number available for analysis: 23

Setting: tertiary psychogeriatrics unit – IRCCS S. Giovanni di Dio-FBF Brescia, Italy

Country: Italy

Period of study: April 2002-March 2005

Language: English

Index tests

Index test: MRI manual and visual method for estimation of hippocampal volume and medial temporal lobe

Manufacturer: Philips Gyroscan

Tesla strength: 1.0

Assessment methods: manual segmentation and visual scale (Scheltens 1992). Manual tracings of hippocampal volume was performed using DISPLAY

Description of positive case definition by index test as reported: not specified for the manual method; according to Scheltes for the visual method

Examiners: no information about radiologist

Interobserver variability: not provided

Target condition and reference standard(s)

Target condition: AD, subcortical VD, LBD, and FTD

Prevalence of AD in the sample: 9/23 (39% of cases included in the analysis)

Stable MCI or converted to other dementia: 14/23 (61%) stable

Reference standard: NINCDS-ADRDA (McKhann 1984) for AD

Mean clinical follow-up: 1.6 years

Flow and timing

Withdrawals and losses to follow-up: 33 MCI. Participants were divided into amnesic (N = 28) and non-amnesic (N = 28) and only amnesic MCI underwent a yearly follow-up visit from 1-3 years after enrolment. 4 amnesic MCI participants refused to have any follow-up visit and dropped out and 1 participant who converted to FTD was excluded from the analysis.

Uninterpretable MRI results were not reported.

Comparative

Key conclusions by the authors

In conclusion, our results suggest that parahippocampal and inferior temporal hypoperfusion in amnesic MCI patients could be considered as a correlate of conversion to AD.

Conflict of interests

Not reported

Notes

Source of funding: Fondazione Polizzotto (www.fondazionepolizzotto.it) for an unrestricted educational grant

Caroli 2007 (Continued)

2 x 2 table: data to complete 2 x 2 table provided by the study authors

Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	No		
		High	Low
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
Did the study provide a clear pre-specified definition of what was considered to be a "positive" result of the index test?	No		
Was the index test performed by a single operator or interpreted by consensus in a joint session?	Unclear		
		High	Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
		Low	Low
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		

Caroli 2007 (Continued)

Did all patients receive the same reference standard? Yes

Were all patients included in the analysis? No

High

Clerx 2013a
Study characteristics

Patient sampling	<p>Primary objectives: to compare the diagnostic accuracy of 4 different measures of the medial temporal lobe volume</p> <p>Study population: participants with mild cognitive impairment (MCI) selected from the "Development of Screening Guidelines and Clinical Criteria for Predementia AD (DESCRIPA) study (Visser 2008), and the Alzheimer Center of the VU University Medical centre (VUmc) in Amsterdam</p> <p>Selection criteria: inclusion criteria for both cohorts were age ≥ 54 years, diagnosis of MCI, availability of results for each MRI measure and outcome at follow-up. Patients were included if data of the 4 MRI tests were available; no differences between included and excluded participants were found with respect to age, sex, education, cognitive results. Exclusion criteria: diagnosis of dementia at baseline or any somatic, psychiatric, or neurological disorder (e.g. epilepsy) that might have caused cognitive impairment. Patients excluded from the study: for 54 participants follow-up data were missing, for 21 digital format scan was not available, and for 53 cases not all 4 MRI measurements were available. Data for all 4 medial temporal lobe measurements were available for 328.</p> <p>Study design: prospective cohort study</p>
Patient characteristics and setting	<p>Clinical presentation: MCI defined according to the criteria of Petersen 1999; Petersen 2004 (37% non-amnestic and 63% amnestic MCI)</p> <p>Age years mean (SD): 70.6 \pm 7.6 years</p> <p>Gender (% men): 48.5 % (40.4 DESCRIPA; 55.8 University Medical Center Amsterdam (VUmc)</p> <p>Education years mean (SD): 10.0 \pm 3.8 (8.5 \pm 3.9 DESCRIPA; 11.2 \pm 3.3 VUmc)</p> <p>ApoE4 carriers (%): 51% (45% DESCRIPA; 56% VUmc)</p> <p>Neuropsychological tests: MMSE mean (SD): 27.0 \pm 2.5 (27.2 \pm 2.3 DESCRIPA; 26.6 \pm 2.6 VUmc)</p> <p>Clinical stroke excluded: not specified</p> <p>Co-morbidities: not reported</p> <p>Number enrolled: 328</p> <p>Number available for analysis: 328, 156 from the DESCRIPA cohort and 172 from the VUmc cohort</p> <p>Setting: tertiary university hospitals; 9 of the 20 participating centres in the DESCRIPA study and Alzheimer Center of the VU University Medical centre (VUmc) in Amsterdam</p> <p>Country: Europe</p> <p>Period of study: not reported</p> <p>Language: English</p>

Clerx 2013a (Continued)

Index tests	<p>Index test: MRI visual, manual and automated methods for estimation of hippocampal volume, MRI automated method for estimation of lateral ventricles measure</p> <p>Manufacturer: several (Philips and Siemens)</p> <p>Tesla strength: 1.0 or 1.5 T</p> <p>Assessment methods: regarding the manual method, the hippocampal ROI was constructed by manual delineation of hippocampal borders using the software package developed in house, Show_Images 3.7.0 and using criteria according to Van de Pol 2007. Automated hippocampal volumetry was performed using a multi-atlas segmentation method named "learning embedding for atlas propagation" (LEAP). Measurements of the lateral ventricle was executed with an extension of SIENAX that is a package for single-time-point ('cross-sectional') analysis of brain change. After the tissue segmentation, a registered mask was used to exclude the CSF on the outer side of the brain. The visual rating of MTA was performed using a qualitative scale (Scheltens 1992). Rating was performed on coronal T1-weighted images</p> <p>Description of positive case definition by index test as reported: specified only for the visual method: score ranging from 0 (no atrophy)-4 (severe atrophy). Diagnostic accuracy based on Youden index (Youden 1950), was used to construct 2 x 2 table; results based on a cut point for a sensitivity of 85% were also available</p> <p>Examiners: 3 trained technicians, blinded to the diagnosis, performed the manual segmentation of the hippocampus. Scans from the VU University Medical centre were rated using the visual method by a group of 3 trained raters supervised by a neuroradiologist. Scans from the DESCRIPA study were rated by a single rater from the VU University Medical centre. The blindness of raters was specified only for the manual method</p> <p>Interobserver variability: the inter-rater coefficient of variation was < 8% and the intra-rater was < 5% for the manual method. In the VU University Medical centre cohort, the inter- and intra-rater Cohen kappa values were > 0.80, and in the DESCRIPA cohort the intra-rater weighted Cohen kappa value was 0.68 for the visual method</p>
Target condition and reference standard(s)	<p>Target condition: AD</p> <p>Prevalence of AD in the sample: 91/328 (28% of cases included in the analysis); 19.2% in the DESCRIPA and 35.5% in the VU University Medical centre</p> <p>Stable MCI or converted to other dementia: 237/328 (72.0%; 80.8% in the DESCRIPA and 64.5% in the VU University Medical centre)</p> <p>Stable MCI 225, converted to other dementia 12 (4 participants converted to FTD, 6 converted to LBD, 1 to VD, 1 to another form of dementia).</p> <p>Reference standard: DSM IV and NINCDS-ADRDA criteria (McKhann 1984). According to the DESCRIPA protocol, clinician was blinded to index test results</p> <p>Mean clinical follow-up: 2 years</p>
Flow and timing	<p>Withdrawals and losses to follow-up: none reported</p> <p>Uninterpretable results were not reported</p>
Comparative	
Key conclusions by the authors	<p>Volumetric hippocampal measurements are the best predictors of conversion to AD-type dementia in subjects with MCI after 2-year follow-up and are able to predict annual cognitive decline. Because of the limited rater time, learning embedding for atlas propagation (LEAP) automated hippocampal measurement might be preferred</p>
Conflict of interests	<p>Study authors declared conflicts of interest</p>
Notes	<p>Source of funding: grant 02N-01, FP7/2007-2013</p> <p>2 x 2 table: data extracted from the article</p>

Clerx 2013a (Continued)

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	No		
		High	Low
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
Did the study provide a clear pre-specified definition of what was considered to be a "positive" result of the index test?	No		
Was the index test performed by a single operator or interpreted by consensus in a joint session?	No		
		High	Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		

Clerx 2013a (Continued)

Were the reference standard results interpreted without knowledge of the results of the index tests?

Yes

Low

Low

DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard?

Yes

Did all patients receive the same reference standard?

Yes

Were all patients included in the analysis?

Yes

Low

deToledo-Morell 2004
Study characteristics

Patient sampling

Primary objectives: to compare the entorhinal cortex and hippocampal volumes as predictors of conversion to AD in a group of patients diagnosed with amnesic MCI

Study population: participants with amnesic MCI

Selection criteria: exclusion criteria: < 65 years, neurologic, psychiatric, or systemic conditions that may cause cognitive impairment (e.g. clinical stroke, alcoholism, major depression, a history of temporal lobe seizures)

Study design: prospective longitudinal

Patient characteristics and setting

Clinical presentation: amnesic MCI defined according to [Petersen 2001](#) diagnostic criteria

Age mean (SD): people with MCI who progressed to AD: 83 ± 5 years; stable MCI: 81 ± 8 years

Gender (% men): MCI who progressed to AD 40%; MCI non-converters 47%

Education years mean (SD) MCI who progressed to AD: 18.4 ± 2.1; MCI non-converters: 15.2 ± 3.1

ApoE4 carriers (%): not reported

Neuropsychological tests: MMSE mean (SD): people with MCI who progressed to AD: 26.1 ± 1.4; MCI non-converters: 28.0 ± 1.8

Clinical stroke excluded: yes

deToledo-Morell 2004 (Continued)

Co-morbidities: not reported

Number enrolled: 27

Number available for analysis: 27

Setting: a tertiary university hospital; the Rush Alzheimer's Disease Center

Country: USA

Period of study: not reported

Language: English

Index tests

Index test: MRI manual method for estimation of hippocampal and entorhinal cortex volumes

Manufacturer: GE (Signa scanners)

Tesla strength: 1.5

Assessment methods: Analyze software package (Mayo Clinic Foundation) was used for determining the volumes of ROI. Both the hippocampal and the entorhinal cortex volumes were manually segmented respectively according to [deToledo-Morrell 1997](#) and [Goncharova 2001](#)

Description of positive cases definition by index test as reported: not specified

Examiners: all tracings were carried out by T.R.S. (who was trained to be within 95% of L.deT.-M.) and were checked, slice by slice, by L.deT.-M. Investigators involved in the MRI analyses were blinded to clinical information until all volumetric determinations were completed

Interobserver variability: inter- and intra-rater correlation coefficients, based on a sample of 10 MRI scans, were 0.97 and 0.97, respectively, for the hippocampal, 0.99 and 0.99, respectively, for the entorhinal cortex

Target condition and reference standard(s)

Target condition: AD

Prevalence of AD in the sample: 10/27 (37% of cases enrolled in the study)

Stable MCI or converted to other dementia: 17 (63%)

Reference standard: NINCDS-ADRDA ([McKhann 1984](#))

Mean clinical follow-up: 3 years

Flow and timing

Withdrawals and losses to follow-up: none reported

Uninterpretable MRI results have not been reported

Comparative

Key conclusions by the authors

Results of the study were in agreement with post mortem pathological findings and underscored the early involvement of the entorhinal cortex in AD. Findings demonstrated the potential of sensitive neuroimaging techniques for the development of early anatomical markers of AD and for tracking, longitudinally, MCI at risk of developing AD

Conflict of interests

Not reported

Notes

Source of funding: this research was supported by Grants P01 AG09466, P30 AG10161, and R01 AG17917 from the NIA, NIH

2 x 2 table: data to complete 2 x 2 table provided by the study authors

Methodological quality

deToledo-Morell 2004 (Continued)

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		High	Low
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Did the study provide a clear pre-specified definition of what was considered to be a "positive" result of the index test?	No		
Was the index test performed by a single operator or interpreted by consensus in a joint session?	Yes		
		High	Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
		Low	Low
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		

deToledo-Morell 2004 (Continued)

Were all patients included in the analysis? Yes

Low

Devanand 2007
Study characteristics

Patient sampling **Primary objectives:** to evaluate the utility of MRI hippocampal and entorhinal cortex atrophy in predicting conversion from MCI to AD

Study population: participants with MCI

Selection criteria: the inclusion/exclusion criteria aimed at enrolling a broad sample of cognitively impaired outpatients who presented with memory complaints and were found to have cognitive impairment without dementia based on comprehensive evaluation, but without a specific cause for the cognitive impairment. Exclusion criteria: diagnosis of dementia, schizophrenia, current major affective disorder, alcohol or substance dependence, history of stroke, cortical stroke or infarct 2 cm in diameter based on MRI, cognitive impairment entirely caused by medications, or other major neurologic illness, e.g. Parkinson disease

Study design: prospective longitudinal

Patient characteristics and setting **Clinical presentation:** mild cognitive impairment (MCI); subtypes definition according to the criteria of [Peterson 1999](#)

Age mean (SD): MCI who progressed to AD: 72 ± 7 years; MCI non-converters to AD: 65 ± 10 years

Gender (% men): MCI who progressed to AD: 43.2%; MCI non-converters to AD: 44.4%

Education years mean (SD): MCI who progressed to AD: 14.1 ± 4.5; MCI non-converters to AD: 15.6 ± 4

ApoE4 carriers (%): MCI who progressed to AD: 32%; MCI non-converters to AD: 20%

Neuropsychological tests: employed; MMSE mean (SD): MCI who progressed to AD: 26 ± 2; MCI non-converters to AD: 28 ± 2

Clinical stroke excluded: yes

Co-morbidities: not reported

Number enrolled: 141

Number available for analysis: 139 for estimation of the hippocampal volume, 138 for estimation of the entorhinal cortex volume

Setting: tertiary university hospitals; the Memory Disorders Center at New York State Psychiatric Institute and Columbia-Presbyterian Medical Center. The majority (52%) were physician referred, 25% were self-referred, and 23% were referred by family or friends or other sources

Country: USA

Period of study: not reported

Language: English

Index tests **Index test:** MRI manual method for estimation of hippocampal and entorhinal cortex volumes

Manufacturer: GE

Tesla strength: 1.5

Devanand 2007 (Continued)

Assessment methods: both the hippocampal and the entorhinal cortex volumes were manually segmented respectively according to [Bobinski 2000](#) and [Killiany 2002](#).

Description of positive cases definition by index test as reported: not specified

Examiners: a single trained rater (G.P.) evaluated all scans on a Sun UltraSPARC workstation blind to all clinical information, using a dedicated software package (MIDAS) for image segmentation and coregistration

Interobserver variability: the single MRI rater was trained with expert raters and showed high interrater reliability on 10 scans (sum of left and right volumes): hippocampal volume ICC 0.90, parahippocampal gyrus volume ICC 0.96, and entorhinal cortex volume ICC 0.92

Target condition and reference standard(s)

Target condition: AD and dementia

Prevalence of AD in the sample: 35/139 (25% of cases included in the analysis)

Stable MCI or converted to other dementia: 104 (75%); 102 stable MCI, 1 MCI converted to corticobasal degeneration, 1 MCI converted to amyotrophic lateral sclerosis with frontal lobe deficits

Reference standard: NINCDS-ADRDA ([McKhann 1984](#)) and DSM-IV ([American Psychiatric Association 2000](#))

A consensus diagnosis was made between 2 expert clinical raters who remained blind to data from previous visits

Mean clinical follow-up: 3 years

Flow and timing

Withdrawals and losses to follow-up: 1 participant was excluded for head motion during MRI acquisition. Within 6 months of presentation, 2 participants with MCI were diagnosed with other neurologic disorders (corticobasal degeneration, and amyotrophic lateral sclerosis presenting with frontal lobe deficits) and were excluded
Uninterpretable MRI results have not been reported

Comparative

Key conclusions by the authors

In logistic regression analyses in the 3-year follow-up sample, entorhinal cortex and hippocampal volume each showed moderately strong diagnostic accuracy. The combined effects of hippocampal and entorhinal cortex volumes further improved test accuracy

Conflict of interests

Study authors reported no conflicts of interest

Notes

Source of funding: supported in part by grants AG17761, AG12101, MH55735, MH35636, MH55646, P50 AG08702, and P30 AG08051 from the NIA and the National Institute of Mental Health

2 x 2 table: data to complete 2 x 2 table provided by the study authors

Methodological quality

Item

Authors' judgement

Risk of bias

Applicability concerns

DOMAIN 1: Patient Selection

Was a consecutive or random sample of patients enrolled?

No

Was a case-control design avoided?

Yes

Devanand 2007 (Continued)

Did the study avoid inappropriate exclusions? No

High
Low
DOMAIN 2: Index Test All tests

Were the index test results interpreted without knowledge of the results of the reference standard? Yes

Did the study provide a clear pre-specified definition of what was considered to be a "positive" result of the index test? No

Was the index test performed by a single operator or interpreted by consensus in a joint session? Yes

High
Low
DOMAIN 3: Reference Standard

Is the reference standards likely to correctly classify the target condition? Yes

Were the reference standard results interpreted without knowledge of the results of the index tests? Yes

Low
Low
DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard? Yes

Did all patients receive the same reference standard? Yes

Devanand 2007 (Continued)

Were all patients included in the analysis? No

High

Eckerstrom 2008
Study characteristics

Patient sampling	<p>Primary objectives: to validate the hypothesis that hippocampal atrophy predicts conversion from MCI to dementia, to relate baseline hippocampal volume to different forms of dementia, and to investigate the role of hippocampal side differences and rate of volume loss over time</p> <p>Study population: participants with mild cognitive impairment (MCI)</p> <p>Selection criteria: exclusion criteria: age > 79 or < 49, MMSE score < 19, acute/instable somatic disease, severe psychiatric disorder, substance abuse, pseudodementia, or confusion caused by drugs</p> <p>Study design: prospective longitudinal (the Gothenburg MCI study)</p>
Patient characteristics and setting	<p>Clinical presentations: MCI defined according to the criteria of the International Working Group on Mild Cognitive Impairment (Winblad 2004), GDS:3</p> <p>Age mean (SD): MCI who progressed to AD: 70; MCI who progressed into non-AD: 68.1; stable MCI: 66.6</p> <p>Gender (% men): 43%</p> <p>Education years mean: MCI who progressed to AD: 10.0; MCI who progressed into non-AD 10.9; stable MCI: 12.5</p> <p>ApoE4 carriers (%): not reported</p> <p>Neuropsychological tests: employed; MMSE mean (SD): MCI who progressed to AD: 28.0; MCI who progressed into non-AD 26.0; stable MCI: 28.3.</p> <p>Clinical stroke excluded: yes (Gothenburg study protocol)</p> <p>Co-morbidities: not reported</p> <p>Number enrolled: 42</p> <p>Number available for analysis: 42</p> <p>Setting: memory clinic (single-centre study)</p> <p>Country: Sweden</p> <p>Period: Gothenburg study started in 1999</p> <p>Language: English</p>
Index tests	<p>Index test: MRI manual method for estimation of hippocampal volume</p> <p>Manufacturer: Philips</p> <p>Tesla strength: 0.5 (in Gothenburg MCI study T1-weighted images from the 0.5 T scanner were used for manual volumetry of the hippocampus)</p> <p>Assessment methods: manual segmentation process consisted of two steps:</p>

Eckerstrom 2008 (Continued)

1. pointwise landmark setting was done in the sagittal view of the reformatted coronal image where the demarcation in the original coronal image is indiscernible or difficult to interpret (Pantel 2000);
2. segmentation of the hippocampus in the coronal images was done by continuous pen drawing. The Hipposegm program samples pointwise from the pen drawing

Description of positive cases definition by index test as reported: 2 cut-off criteria were chosen post hoc for left hippocampal volume (2200 and 1800 mm³)

Examiners: 1st rater initially segmented all the scans in 1 session that extended over several weeks. The data from this session were used in the main analysis of the study. The raters were always blinded for group belonging, participant/control ID and other header data

Interobserver variability: performed between 2 raters on 30 MRI scans. Single measure ICC was 0.663, average measure ICC was 0.797

Target condition and reference standard(s)	<p>Target condition: AD was the primary target condition; other dementias were the secondary target conditions</p> <p>Prevalence of AD in the sample: 13/42 (31% of enrolled participants)</p> <p>Stable MCI or converted to other dementia: 29 (69%). In this group, 8 cases converted to non-AD dementia</p> <p>Reference standards: NINCDS-ADRDA (McKhann 1984) for AD diagnosis, Erkinjuntti criteria (Erkinjuntti 2000) for VD diagnosis, Lund and Manchester criteria (The Lund and Manchester Groups 1994) for FTD diagnosis.</p> <p>According to the MCI Gothenburg study protocol, the specialist physician was blinded to psychometric, CSF, and imaging results, except for assessment of white matter change</p> <p>Mean clinical follow-up: 2 years</p>
Flow and timing	<p>Withdrawals and losses to follow-up: none reported</p> <p>Uninterpretable MRI results have not been reported</p>
Comparative	
Key conclusions by the authors	The main findings in this the published article are that hippocampal volume predicts conversion to dementia in MCI patients, and that left hippocampal volume seems to be the best marker for conversion
Conflict of interests	Not reported
Notes	<p>Source of funding: this work was supported by grants from the Swedish Research Council (grants 2002-5462, K2002-21P-14359-01A and 09946)</p> <p>2 x 2 table: data to complete 2 x 2 table provided by the study authors</p>

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		

Eckerstrom 2008 (Continued)

Did the study avoid inappropriate exclusions? No

High
Low
DOMAIN 2: Index Test All tests

Were the index test results interpreted without knowledge of the results of the reference standard? Yes

Did the study provide a clear pre-specified definition of what was considered to be a "positive" result of the index test? No

Was the index test performed by a single operator or interpreted by consensus in a joint session? Yes

High
Low
DOMAIN 3: Reference Standard

Is the reference standards likely to correctly classify the target condition? Yes

Were the reference standard results interpreted without knowledge of the results of the index tests? Yes

Low
Low
DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard? Yes

Did all patients receive the same reference standard? Yes

Were all patients included in the analysis? Yes

Low

Eckerstrom 2013
Study characteristics

Patient sampling	<p>Primary objectives: to study prediction of dementia in MCI using neuropsychological tests, commonly used CSF biomarkers, and hippocampal volume</p> <p>Study population: participants with MCI</p> <p>Selection criteria: exclusion criteria: age > 79 or < 49, MMSE score < 19, acute/instable somatic disease, severe psychiatric disorder, substance abuse, pseudodementia, or confusion caused by drugs</p> <p>Study design: prospective longitudinal (the Gothenburg MCI study)</p>
Patient characteristics and setting	<p>Clinical presentations: MCI defined according to the criteria of the International Working Group on Mild Cognitive Impairment (Winblad 2004), GDS:3</p> <p>Age mean (SD): MCI who progressed to AD: 70 ± 6.5; MCI who progressed to all dementia subtypes: 69.3 ± 6.3; stable MCI: 66.4 ± 6.8</p> <p>Gender (% men): 43%</p> <p>Education years mean (SD): MCI who progressed to AD: 10.0 ± 2.9; MCI who progressed to all dementia subtypes: 10.4 ± 3.4; Stable MCI: 12.3 ± 3.6</p> <p>ApoE4 carriers (%): not reported</p> <p>Neuropsychological tests: employed; MMSE mean (SD): MCI who progressed to AD: 28.0 ± 1.1; MCI who progressed to all dementia subtypes: 27.2 ± 2; stable MCI: 28.3 ± 1.8</p> <p>Clinical stroke excluded: yes (Gothenburg study protocol)</p> <p>Co-morbidities: not reported</p> <p>Number enrolled: 42</p> <p>Number available for analysis: 42 for dementia (AD and other type); 34 for AD</p> <p>Setting: memory clinic (single-centre study)</p> <p>Country: Sweden</p> <p>Period: Gothenburg study started in 1999</p> <p>Language: English</p>
Index tests	<p>Index test: MRI manual method for estimation of left and right hippocampus volumes</p> <p>Manufacturer: Philips NT5</p> <p>Tesla strength: 0.5 T</p> <p>Assessment methods: manual segmentation protocol as in Eckerstrom 2008</p> <p>Description of positive cases definition by index test as reported: not specified</p> <p>Examiners: as in Eckerstrom 2008</p> <p>Interobserver variability: as in Eckerstrom 2008</p>
Target condition and reference standard(s)	<p>Target condition: AD was the primary target condition; other dementias were the secondary target conditions</p> <p>Prevalence of AD in the sample: 13/42 (31% of 42 enrolled participants)</p>

Eckerstrom 2013 (Continued)

Stable MCI or converted to other dementia: 29 (69%). In this group, 8 cases converted to non-AD dementia

Reference standards: NINCDS-ADRDA (Mckhann 1984) for AD diagnosis, Erkinjuntti criteria (Erkinjuntti 2000) for VD diagnosis, Lund and Manchester criteria (The Lund and Manchester Groups 1994) for FTD diagnosis.

According to the MCI Gothenburg study protocol, the specialist physician was blinded to psychometric, CSF, and imaging results, except for assessment of white matter change

Mean clinical follow-up: 2 years

Flow and timing	<p>Withdrawals and losses to follow-up: not specified if the 8 cases that converted to other dementias were included in the analysis (in the non-AD dementia group)</p> <p>Uninterpretable MRI results have not been reported</p>
Comparative	
Key conclusions by the authors	The main findings of the study were that in the diagnosis of dementia, combinations of markers performed better than individual markers, and that neuropsychological tests had overall better diagnostic accuracy than CSF biomarkers and, particularly, hippocampal volume
Conflict of interests	Not reported
Notes	<p>Source of funding: grants from Swedish research council; Demensfonden, Pffannenstills stiftelse, Alzheimerfonden, Swedish Brain Power, Sahlgrenska University Hospital</p> <p>2 x 2 table: data from the published article</p>

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	No		
		High	Low
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Did the study provide a clear pre-specified definition of what was considered	No		

Eckerstrom 2013 (Continued)

to be a "positive" result of the index test?

Was the index test performed by a single operator or interpreted by consensus in a joint session?

Yes

High
Low
DOMAIN 3: Reference Standard

Is the reference standards likely to correctly classify the target condition?

Yes

Were the reference standard results interpreted without knowledge of the results of the index tests?

Yes

Low
Low
DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard?

Yes

Did all patients receive the same reference standard?

Yes

Were all patients included in the analysis?

Unclear

Unclear
Erten-Lyons 2006
Study characteristics

Patient sampling

Primary objectives: examine whether presymptomatic rates of regional and total brain volume loss distinguish MCI participants that subsequently decline to dementia

Study population: participants with MCI

Selection criteria: all Oregon Brain Aging Study participants who had at least 3 annual neuropsychological examinations and ≥ 2 analysed MRI scans for the time of interest. MCI defined clinically as CDR = 0.5 on 2 independent examinations separated by at least 6 months and not meeting diagnostic criteria for dementia. Exclusion criteria: not reported

Study design: longitudinal population study

Patient characteristics and setting

Clinical presentations: MCI defined clinically as CDR = 0.5 on 2 independent examinations separated by at least 6 months and not meeting diagnostic criteria for dementia

Age mean (SD): MCI who progressed to dementia: 88 ± 5 ; stable MCI: 86 ± 7

Erten-Lyons 2006 (Continued)

Gender (% men): MCI who progressed to dementia: 17%; stable MCI: 50%

Education years mean (SD): MCI who progressed to dementia: 13 ± 3.3; stable MCI: 15 ± 3.4

ApoE4 carriers (%): MCI who progressed to dementia: 21%; stable MCI: 17%

Neuropsychological tests: employed; MMSE mean (SD): MCI who progressed to dementia: 27.13 ± 2.05; stable MCI: 27.5 ± 1.87

Clinical stroke excluded: not specified

Co-morbidities: not specified

Number enrolled: 37

Number available for analysis: 37

Setting: memory clinic (single-centre)

Country: USA

Period: not reported

Language: English

Index tests

Index test: MRI semi-automated method for estimation of the hippocampal, ventricular and whole brain volumes

Manufacturer: not reported

Tesla strength: 1.5 T ([Kaye 1997](#))

Assessment methods: semi-automated according to ([Kaye 1997](#)). Analysis of the MRI images was performed with computer-assisted techniques utilising a program called REGION

Description of positive cases definition by index test as reported: not specified

Examiners: not reported

Interobserver variability: among 3 operators who scored the same set of 5 MRI scans, the ICC was 0.98 for the intracranial volume and 0.97 for the temporal lobe atrophy ([Kaye 1997](#))

Target condition and reference standard(s)

Target condition: probable AD

Prevalence of AD in the sample: 22/37 (59.5% of enrolled participants), including 17 probable AD and 5 possible AD

Stable MCI or converted to other dementia: 15 (40.5%), including 14 participants who remained stable and 1 with VD

Reference standards: not reported

Mean clinical follow-up: 7.6 years

Flow and timing

Withdrawals and losses to follow-up: none reported

Uninterpretable MRI results have not been reported

Comparative

Key conclusions by the authors

Results suggest that patients with MCI that decline to dementia have smaller hippocampal volumes at baseline and greater rates of generalised and regional brain atrophy several years before symptom onset compared with those who remain stable

Erten-Lyons 2006 (Continued)

Conflict of interests Study authors declared no conflict of interest

Notes **Source of funding:** supported by the Merit Review Grant, Office of Research and Development, Department of Veterans Affairs, National Institute on Aging, NIH AG08017, MO1 RR000334

2 x 2 table: data to complete 2 x 2 table provided by the study authors

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
		High	Low
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
Did the study provide a clear pre-specified definition of what was considered to be a "positive" result of the index test?	No		
Was the index test performed by a single operator or interpreted by consensus in a joint session?	No		
		High	Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
		Unclear	Unclear

Erten-Lyons 2006 (Continued)

DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard? Yes

Did all patients receive the same reference standard? Unclear

Were all patients included in the analysis? Yes

Unclear

Frolich 2017
Study characteristics

Patient sampling **Primary objectives:** investigate which combination of biomarkers best predict a short-term conversion from MCI to AD dementia

Study population: participants from the Dementia Competence Network (DCN), a German multicenter cohort study

Selection criteria: patients who sought evaluation at the participating memory clinics, aged ≥ 50 years and in whom organic cognitive impairment was suspected underwent a screening assessment.

The diagnoses of MCI were made on the basis of clinical and neuropsychological data. A specific MMSE threshold was not applied for diagnosis. A broader definition of MCI was used, the core features being complaints of cognitive deficit in ADL and objectified decline of cognitive abilities (> 1 SD) in at least 1 of main cognitive domains as evidenced by standardised neuropsychological tests; B-ADL score < 4 ; no major depressive episode. The following exclusion criteria were applied: substance abuse or dependence, insufficient German language skills, multimorbidity, comorbid condition with excess mortality, circumstances that make regular attendance at follow-up visits questionable and lack of an informant. In most centres, a consecutive series of patients was included in the study, nevertheless 115/1071 MCI patients (12%) were included on the basis of biomarkers availability, follow-up length (for at least 12 months) and outcome (MCI stable or progressed to AD only)

Study design: prospective multisite longitudinal observational study on memory clinic patients with MCI or early dementia

Patient characteristics and setting **Clinical presentation:** MCI was broadly defined including amnesic and nonamnesic syndromes according to [Winblad 2004](#)

Age mean (SD): MCI who progressed to AD: 65.4 ± 9.37 years; MCI non-converters to AD: 66.5 ± 8.95 years

Gender (% men): MCI who progressed to AD: 54%; MCI non-converters to AD: 60%

Education years mean (SD): MCI who progressed to AD: 8.75 ± 1.58 years;

MCI non-converters to AD: 9.75 ± 1.95 years

ApoE4 carriers (%): MCI who progressed to AD: 36%; MCI non-converters to AD: 41%

Neuropsychological tests: employed; MMSE mean (SD): MCI who progressed to AD: 28 ± 2.34 ; MCI non-converters to AD: 27.5 ± 1.87

Clinical stroke excluded: not specified

Co-morbidities: not reported

Number enrolled: 115

Number available for analysis: 115

Setting: Dementia Competence Network

Frolich 2017 (Continued)

Country: Germany

Period of study: participants were recruited between May 2003 and November 2007

Language: English

Index tests	<p>Index test: automated method for estimation of hippocampal volume</p> <p>Manufacturer: Siemens, Philips</p> <p>Tesla strength: 1.5 T</p> <p>Assessment methods: hippocampal volume was calculated as the mean value of the left and right hemisphere and was determined from high-resolution structural MRI using Oxford Centre for Functional MRI of the Brain Integrated Registration and Segmentation Tool from the Functional MRI of the Brain Software Library package of tools (Pate-naude 2011).</p> <p>Description of positive cases definition by index test as reported: not specified. Sensitivity, specificity and Youden's index of various predictors were calculated (Youden 1950).</p> <p>Examiners: no details about radiologist</p> <p>Interobserver variability: not reported. Special measures were taken for standardisation of MRI acquisition across centres. Acquisition parameters were provided to all centres as guideline. The phantom test of the American College of radiology MRI Accreditation Program was conducted repeatedly at 11 sites of the Dementia Competence Network. Furthermore a single volunteer was investigated at each of these centres.</p>
Target condition and reference standard(s)	<p>Target condition: AD</p> <p>Prevalence of AD in the sample: 28/115 (24% of cases included in the analysis)</p> <p>Stable MCI or converted to other dementia: 87/115 (76%) stable MCI</p> <p>Reference standard: Dementia Competence Network study used NINCDS-ADRDA (McKhann 1984) for the AD diagnosis</p> <p>Mean clinical follow-up: 2.2 years</p>
Flow and timing	<p>Withdrawals and losses to follow-up: none reported</p> <p>Uninterpretable MRI results have not been reported</p>
Comparative	
Key conclusions by the authors	<p>A combination of two biomarkers of neurodegeneration (e.g., HCV and t-Tau) is not superior over the single parameters in identifying patients with MCI who are most likely to progress to AD dementia, although there is a gradual increase in the statistical measures across increasing biomarker combinations</p>
Conflict of interests	<p>LF has received payment for consultancy, expert testimony, honorarium, or travelling support from AstraZeneca, Eisai, Eli Lilly, GE Healthcare, Janssen-Cilag, Lundbeck, Merz Pharma, Novartis, Pfizer, and Schering-Plough, and Apotex Inc., and has received a research grant from Novartis, Pfizer paid to his institution. OP is on scientific advisory boards for Roche, Kyowa Kirin, Novartis, Lilly, and Piramal. He has received funding for travel or speaker honoraria from GSK, Nutricia, and Merck Serono. He has acted as a consultant for Affiris and Roche. He has received research support from Affiris, Piramal, BMS, Eli Lilly, Pfizer, Servier, TRX Pharmaceuticals, Lundbeck, and Genentech. FJ has received consultation board honoraria and speaker's fees from AC Immune, Lilly, GE Healthcare, Janssen, USB, Schwabe, Eisai, Pfizer, Novartis, and Roche. He has received a research grant paid to his institution from Schwabe. JP has received honoraria from Merz, Janssen-Cilag, and Novartis. MH has received a research grant from Schwabe GmbH, has received speaker's honoraria from Pfizer Inc., Merz Pharmaceuticals, and GlaxoSmithKline, and served on an advisory board for Hoffmann-La Roche. ER is a Merz GmbH collaborator, is on the speakers' bureau of/has received a travel grant from BMS, Lundbeck, Servier, and Otsuka, and has received a research grant from Lilly, BMS, AstraZeneca, and Lundbeck. FH was a consultant to AstraZeneca in the area of depression between June 2011 and November 2012. WM has received payment for educational lectures from Merz. JW is on the advisory board for Eli Lilly and has received consulting fee or honorarium and support for travelling to Board meetings. He received payment for lectures from Novartis. JK has received financial support for conducting clinical trials from various pharmaceutical companies manufacturing anti-dementia drugs. He is mentioned as coinventor on the following patents: Substituted piperidines or pyrrolidine compounds for treating sigma-receptor modulated disorders (WO001996031208A3); Method of differentially diagnosing dementias (WO002008058764A1); Soluble amyloid precursor</p>

Frolich 2017 (Continued)

sor proteins in CSF as biomarkers of AD (EP000002068151A1); Immunoglobulinbound A β and immunoglobulins-binding A β peptides in diagnosis and therapy of AD (WO002007082750A1); and Method of diagnosing acute cerebral ischemia (WO002008058764A1). The remaining study authors declare that they have no competing interests.

Notes **Source of funding:** this study was supported by a grant from the German Federal Ministry of Education and Research (BMBF): Kompetenznetz Demenzen (01GI0420)

2 x 2 table: data from the published article

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		
---	---------	--	--

High

Low

DOMAIN 2: Index Test All tests

Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
---	---------	--	--

Did the study provide a clear pre-specified definition of what was considered to be a "positive" result of the index test?	No		
--	----	--	--

Frolich 2017 (Continued)

Was the index test performed by a single operator or interpreted by consensus in a joint session? Unclear

High
Low
DOMAIN 3: Reference Standard

Is the reference standards likely to correctly classify the target condition? Yes

Were the reference standard results interpreted without knowledge of the results of the index tests? Yes

Low
Low
DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard? Yes

Did all patients receive the same reference standard? Yes

Were all patients included in the analysis? Yes

Low

Galton 2005
Study characteristics

Patient sampling

Primary objectives: examine the contribution of the Addenbrooke's Cognitive Examination (ACE), neuropsychological assessment, and a MRI-based temporal lobe rating scale to the prediction of which patients with questionable dementia will progress to AD

Study population: a subgroup of patients with QD was recruited from the memory clinic in Addenbrooke's Hospital, Cambridge, for a longitudinal project. Included participants had undergone suitable MRI imaging at entry.

Selection criteria: participants unable to undergo MRI for a variety of reasons including claustrophobia or pacemakers were excluded. Consecutive referrals were approached and screened to exclude extrapyramidal signs or hallucinations, VD, current cancer treatment, uncontrolled diabetes, and serious head injury. Patients with cerebrovascular events, epilepsy, and major depression were excluded. All participants were aged between 50 and 80 years at the time of recruitment.

Study design: prospective longitudinal study

Patient characteristics and setting

Clinical presentations: QD met the following criteria:

1. subjective complaint of memory impairment;
2. normal ADL; and
3. nondemented as evident from a MMSE score of ≥ 23

Such definition corresponded to MCI as defined by some groups of workers using the CDR scale but not by other groups who based the diagnosis of MCI on neuropsychological test performance. All included participants had a CDR score of 0.5.

Age mean (SD): people with QD who progressed to AD: 71 ± 9 ; stable QD: 59 ± 8

Gender (% men): QD who progressed to AD: 54.5%; stable QD: 44.4%

Education years mean (SD): not reported

ApoE4 carriers (%): not reported

Neuropsychological tests: employed; MMSE mean (SD) QD who progressed to AD: 24.8 ± 1.5 ; stable QD: 29 ± 0.9

Clinical stroke excluded: yes

Co-morbidities: not reported

Number enrolled: 31

Number available for analysis: 29

Setting: memory clinic in Addenbrooke's Hospital, Cambridge

Country: UK

Period: 1997-1999

Language: English

Index tests

Index test: MRI visual method for estimation of the left and right hippocampus, left and right parahippocampus, right lateral temporal lobe volumes

Manufacturer: GE

Tesla strength: not specified for all images (66% were performed at 1.5 T, 34% were scanned before with different not specified protocol)

Galton 2005 (Continued)

Assessment methods: visual method according to the Temporal Lobe Rating scale (Galton 2001) on coronal T1 images. The method built on the work of Scheltens 1992 but incorporates also temporal pole, parahippocampal region and the inferolateral temporal region

Description of positive cases definition by index test as reported: ratings of each region were scored dichotomising as normal or questionable (0-1) versus abnormal (2-3)

Examiners: the 50 coronal films were blinded and randomised to the examiner/examiners (not specified if > 1). Imagings were assessed in 2 sessions (to maintain concentration)

Interobserver variability: not provided

Target condition and reference standard(s)	<p>Target condition: AD</p> <p>Prevalence of AD in the sample: 11/29 (38% of cases included in the analysis)</p> <p>Stable MCI or converted to other dementia: 20 (69%). In this group 2 participants converted to LBD and were excluded from the analysis</p> <p>Reference standards: NINCDS-ADRDA criteria (McKhann 1984). Onset of dementia was determined on functional grounds after interview and examination of the participant and family members and was independent of the neuropsychological and imaging findings</p> <p>Mean clinical follow-up: 1.6 years</p>
Flow and timing	<p>Withdrawals and losses to follow-up: 2/31 participants (6.5% of enrolled participants) converted to LBD and were excluded from the analysis.</p> <p>Uninterpretable MRI results have not been reported.</p>
Comparative	
Key conclusions by the authors	The medial temporal lobe is more atrophied in those participants at risk for dementia who subsequently convert to presumed AD. The Addenbrooke's Cognitive Examination was the best single predictor of progression to AD
Conflict of interests	Not reported
Notes	<p>Source of funding: supported by an MRC-LINK grant</p> <p>2 x 2 table: data from the published article</p>

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?	No		
		High	Low

Galton 2005 (Continued)

DOMAIN 2: Index Test All tests

Were the index test results interpreted without knowledge of the results of the reference standard? Yes

Did the study provide a clear pre-specified definition of what was considered to be a "positive" result of the index test? Yes

Was the index test performed by a single operator or interpreted by consensus in a joint session? Unclear

Low
Low
DOMAIN 3: Reference Standard

Is the reference standards likely to correctly classify the target condition? Yes

Were the reference standard results interpreted without knowledge of the results of the index tests? Yes

Low
Low
DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard? Yes

Did all patients receive the same reference standard? Yes

Were all patients included in the analysis? No

High

Gaser 2013
Study characteristics

Patient sampling

Primary objectives: explore the potential of applying the BrainAGE approach in early detection of abnormal brain changes in order to predict the conversion from MCI to AD within a time span of 3 years

Study population: data obtained from the ADNI database including all MCI for whom baseline MRI data 1.5 T, at least moderately confident diagnoses, hippocampus volumes calculated by Freesurfer Version 4.3 and test scores in certain cognitive scales were available

Selection criteria: inclusion and exclusion criteria according to the ADNI protocol ([ADNI 2010](#); [Petersen 2010](#)). All participants were aged 55–90 and had no evidence of cerebrovascular disease (Modified Hachinski Ischaemia Score ≤ 4), depression (GDS < 6), no significant neurological disease, no visual or hearing impairment, stable medications, good general health, 6 grades of education or equivalent, a study partner, English or Spanish language fluency, and no medical contraindications to MRI

Study design: prospective longitudinal study (ADNI study)

Patient characteristics and setting

Clinical presentations: MCI defined according to the ADNI protocol and these criteria: memory complaint verified by study partner; abnormal memory function based on education-adjusted cut-off on the Logical Memory II subscale (delayed paragraph recall) from the WMS-R; MMSE score of 24–30 (inclusive); CDR score of 0.5; and cognitive and functional impairment not severe enough to meet criteria for AD or dementia

Age mean (SD): MCI who progressed early to AD: 74 ± 7 ; MCI who progressed late to AD: 75 ± 7 ; stable MCI: 76 ± 6

Gender (% men): MCI who progressed early to AD: 57%; MCI who progressed late to AD: 64%; stable MCI: 79%

Education years mean (SD): MCI who progressed early to AD: 15.4 ± 2.9 ; MCI who progressed late to AD: 16.0 ± 2.9 ; stable MCI: 16.5 ± 2.6

ApoE4 carriers (%): not reported

Neuropsychological tests: employed; MMSE mean (SD): MCI who progressed early to AD: 26.5 ± 1.9 ; MCI who progressed late to AD: 26.8 ± 1.6 ; stable MCI: 27.7 ± 1.8

Clinical stroke excluded: not specified

Co-morbidities: not reported

Number enrolled: 195

Number available for analysis: 195

Setting: ADNI database (multicentre study)

Country: USA and Canada

Period: data downloaded in May 2010

Language: English

Index tests

Index test: MRI automated method for estimation of left and right hippocampal volumes

Manufacturer: GE Healthcare, Philips Medical System, Siemens Medical Solution ([Jack 2008b](#))

Tesla strength: 1.5 Tesla

Assessment methods: automated method using FreeSurfer version 4.3 (data obtained from the ADNI database). For the exact procedures of data collection and up-to-date information, see adni.loni.usc.edu/.

Description of positive cases definition by index test as reported: not specified

Gaser 2013 (Continued)

Examiners: not specified

Interobserver variability: not specified

Target condition and reference standard(s)

Target condition: AD

Prevalence of AD in the sample: 133/195 (68%) of the enrolled participants at 3-year follow-up; 58/195 (30%) at 1-year follow-up

Stable MCI or converted to other dementia: 62 (32%) stable MCI at 3-year follow-up

Reference standard: AD diagnosis was made according to NINCDS-ADRDA criteria

Flow and timing

Withdrawals and losses to follow-up: none reported

Uninterpretable MRI results have not been reported

Comparative

Key conclusions by the authors

With accuracy rates of up to 81%, BrainAGE outperformed all cognitive scales and cerebrospinal fluid biomarkers in predicting conversion of MCI to AD within 3 years of follow-up. Furthermore, the post-test probability was increased to 90% when using baseline BrainAGE scores to predict conversion to AD

Conflict of interests

Study authors declared no competing interests

Notes

Source of funding: BMBF grant 01EV0709, data collection and sharing was funded by ADNI

2 x 2 table: data from the published article

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
------	--------------------	--------------	------------------------

DOMAIN 1: Patient Selection

Was a consecutive or random sample of patients enrolled?	No		
--	----	--	--

Was a case-control design avoided?	Yes		
------------------------------------	-----	--	--

Did the study avoid inappropriate exclusions?	No		
---	----	--	--

High
Low
DOMAIN 2: Index Test All tests

Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
---	---------	--	--

Did the study provide a clear pre-specified definition of what was consid-	No		
--	----	--	--

Gaser 2013 (Continued)

ered to be a "positive" result of the index test?

Was the index test performed by a single operator or interpreted by consensus in a joint session? Unclear

High
Low
DOMAIN 3: Reference Standard

Is the reference standards likely to correctly classify the target condition? Yes

Were the reference standard results interpreted without knowledge of the results of the index tests? Yes

Low
Low
DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard? Yes

Did all patients receive the same reference standard? Yes

Were all patients included in the analysis? Yes

Low
Herukka 2008
Study characteristics

Patient sampling **Primary objectives:** investigate the association between the CSF biomarkers and MTA and the ability of these measures to predict AD in MCI
Study population: MCI according to [Petersen 2001](#)
Selection criteria: the substudy included all participants for whom both lumbar puncture and a volumetric MRI scan had been performed. Exclusion criteria not reported
Study design: prospective longitudinal study

Patient characteristics and setting **Clinical presentations:** MCI had an objective impairment in at least 1 cognitive domain (performance < 1.5 SD below age-adjusted values) and a CDR score of 0.5

Herukka 2008 (Continued)

Age mean (SD): MCI who progressed to AD: 72 ± 5; stable MCI: 71 ± 5

Gender (% men): MCI who progressed to AD: 37.5%; stable MCI: 31%

Education years mean (SD): not reported

ApoE4 carriers (%): not reported

Neuropsychological tests: employed; MMSE mean (SD): not available

Clinical stroke excluded: not specified

Co-morbidities: not reported

Number enrolled: 21

Number available for analysis: 21

Setting: participants examined in Neurological Department of Kuopio University Hospital and participants in an ongoing population-based follow-up study in the University of Kuopio

Country: Finland

Period: not reported

Language: English

Index tests

Index test: MRI manual method for estimation of left, right, and total volumes of hippocampus and entorhinal cortex

Manufacturer: Siemens

Tesla strength: 1.5

Assessment methods: the hippocampi and entorhinal cortex were manually traced using custom-made software for a standard Siemens work console

Description of positive cases definition by index test as reported: not specified

Examiners: single rater, blinded to clinical data

Interobserver variability: ICCs for intra-rater reliability were 0.96 for the hippocampus and 0.95 for the entorhinal cortex measured from 10 participants

Target condition and reference standard(s)

Target condition: AD

Prevalence of AD in the sample: 8/21 (38% of enrolled participants)

Stable MCI or converted to other dementia: 13 (62%) stable MCI

Reference standards: NINCDS-ADRDA criteria ([McKhann 1984](#))

Mean clinical follow-up: MCI who progressed to AD: 3.38 ± 1.85 years; stable MCI: 4.77 ± 1.09 years

Flow and timing

Withdrawals explained and losses to follow-up: none reported
 Uninterpretable MRI results have not been reported

Comparative

Key conclusions by the authors

The results of the study suggest that cerebrospinal fluid biomarkers and MTA measured by volumetric MRI correlate with each other and are associated with impairment in memory performance. These findings provide further evidence that both cerebrospinal fluid biomarkers and MRI of the medial temporal lobe structures are useful in the confirmation of early, perhaps even the preclinical diagnosis of AD

Herukka 2008 (Continued)

Conflict of interests Study authors declared no conflict of interest

Notes **Source of funding:** study was supported by Academy of Finland grant number 201495, Kuopio University Hospital EVO grants 5883, 5772720 and 5772725

2 x 2 table: data to complete 2 x 2 table provided by the study authors

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	No		
		High	Low
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Did the study provide a clear pre-specified definition of what was considered to be a "positive" result of the index test?	No		
Was the index test performed by a single operator or interpreted by consensus in a joint session?	Yes		
		High	Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
		Low	Low

Herukka 2008 (Continued)

DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard? Yes

Did all patients receive the same reference standard? Yes

Were all patients included in the analysis? Yes

Low

Jack 2000
Study characteristics

Patient sampling **Primary objectives:** to test the hypothesis that the annualised rates of hippocampal atrophy differ as a function of both baseline and change in clinical group membership (control, MCI, or AD)

Study population: amnesic MCI, community and referral patients

Selection criteria: criteria for the diagnosis of MCI were:

1. memory complaint documented by the patient and collateral source
2. normal general cognition
3. normal ADL
4. not demented (DSM-III-R) American Psychiatric Association 1987
5. CDR score of 0.5.

Participants must have had 2 Mayo ADPR/ADRC clinical assessments separated by a minimum of 2 and a maximum of 4 years. Participants who had symptoms which were clinically felt to be unrelated to AD were excluded. For example, participants who suffered a stroke or who developed depression before or during the follow-up period were excluded

Study design: prospective longitudinal study

Patient characteristics and setting **Clinical presentations:** amnesic MCI (single domain)

Age mean (SD): MCI who progressed to AD: 77 ± 8; stable MCI: 78 ± 8

Gender (% men): MCI who progressed to AD: 50%; stable MCI: 44%

Education years mean (SD): MCI who progressed to AD: 13.1 ± 3.1; stable MCI: 14.0 ± 3.1

ApoE4 carriers (%): MCI who progressed to AD: 39%; stable MCI: 28%

Neuropsychological tests: fixed cut-off scores on specific psychometric test of general cognitive performance and memory were not employed; MMSE mean (SD): MCI who progressed to AD: 24 ± 3.2; stable MCI: 27 ± 1.9

Clinical stroke excluded: yes

Co-morbidities: cases with depression were excluded

Number enrolled: 43

Number available for analysis: 43

Jack 2000 (Continued)

Setting: Mayo ADPR (population-based) and ADRC (referral)

Country: USA

Period: not reported

Language: English

Index tests

Index test: MRI manual method for estimation of the hippocampal volume

Manufacturer: GE

Tesla strength: 1.5

Assessment methods: manual segmentation according to (Jack 1992). The boundaries of the hippocampi were delineated on each anatomic slice, the number of voxel was calculated automatically with a summing ROI function. The borders of the hippocampi were manually traced with a mouse-driven cursor for each slice sequentially from posterior to anterior. Inplane hippocampal anatomic boundaries were defined to include the CA1 through CA4 sectors of the hippocampus proper, the dentate gyrus, and subiculum

Description of positive cases definition by index test as reported: deviation from normal value assessed with a W score, i.e. the value from a standard normal distribution corresponding to the observed percentile in controls. A cut-off value was not clearly specified.

Examiners: all image processing steps in every participant were performed by the same research associate who was blinded to all clinical information

Interobserver variability: not provided (single rater)

Target condition and reference standard(s)

Target condition: AD

Prevalence of AD in the sample: 18/43 (42% of enrolled participants)

Stable MCI or converted to other dementia: 25 (58%) stable

Reference standards: NINCDS/ADRDA criteria (McKhann 1984). The hippocampal volume data derived from the MRI examination and ApoE genotypes were not known to the ADPR/ADRC consensus committee throughout the study

Mean clinical follow-up (SD): 3 ± 1 years

Flow and timing

Withdrawals explained and losses to follow-up: none reported

Uninterpretable MRI results has not been reported

Comparative

Key conclusions by the authors

The data demonstrate a correlation between the rate of change in hippocampal volume and change in cognitive status; data also indicate that the distinction between stable vs a declining members of a group should be detectable both in early symptomatic patient groups (i.e., MCI) and in presymptomatic participants (i.e., controls)

Conflict of interests

Not reported

Notes

Source of funding: grant support: NIH, NIA, AG11378, AG08031, AG06786, AG16574, NS 29059, The DANA Foundation, The Alzheimer's Association

2 x 2 table: data to complete 2 x 2 table provided by the study authors

Methodological quality

Jack 2000 (Continued)

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	No		
		High	Low
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Did the study provide a clear pre-specified definition of what was considered to be a "positive" result of the index test?	No		
Was the index test performed by a single operator or interpreted by consensus in a joint session?	Yes		
		High	Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
		Low	Low
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between in-	Yes		

Jack 2000 (Continued)
 dex test and reference
 standard?

Did all patients receive
 the same reference
 standard? Yes

Were all patients in-
 cluded in the analysis? Yes

Low

Jang 2018

Study characteristics

Patient sampling **Primary objectives:** investigate the use of CVRS for predicting dementia and elucidate its association with cognitive change in patients with MCI over a 3-year follow-up

Study population: participants from the ADNI study

Selection criteria: MCI who had a baseline MRI scan as well as amyloid PET study and at least ≥ 1 follow-up visits after initial assessment were included. All participants had a MMSE score of ≥ 24 , a global CDR score of 0.5, a CDR memory score of ≥ 0.5 , and a score indicating impairment on the delayed recall of Story A of the WMS-R

Exclusion criteria: according to ADNI protocol

Study design: prospective longitudinal study (participants from ADNI study)

Patient characteristics and setting **Clinical presentation:** MCI according to ADNI protocol. Diagnosis of MCI was made according to the presence of objective memory impairment

Age mean (SD): MCI who progressed to AD: 72.1 ± 7.2 years; MCI non-converters to AD: 71.1 ± 7.5 years

Gender (% men): MCI who progressed to AD: 54%; MCI non-converters to AD: 53%

Education years median (range): 16 (14-18) in both MCI groups

ApoE4 carriers (%): MCI who progressed to AD: 68%; MCI non-converters to AD: 40%

Neuropsychological tests: employed; MMSE (range) MCI who progressed to AD: 28 (26-29); MCI non-converters to AD: 29 (28-30)

Clinical stroke excluded: not specified

Co-morbidities: not reported

Number enrolled: 340

Number available for analysis: 340

Setting: ADNI database (multicentre study)- ADNI-GO/ADNI2 cohort

Country: USA and Canada

Period of study: ADNI was launched in 2003, data used in this study downloaded from the ADNI database on December 2017

Language: English

Index tests **Index test:** MRI visual method to estimate hippocampal atrophy, cortical atrophy, subcortical atrophy, small vessel disease

Manufacturer: those used in ADNI study (GE Healthcare, Philips Medical System, Siemens Medical Solution (Jack 2008b))

Tesla strength: 3 T

Jang 2018 (Continued)

Assessment methods: CVRS includes visual scales of hippocampal atrophy, cortical atrophy, subcortical atrophy (ventricular enlargement) and for staging small vessel disease; the rater used a template-based scoring program on a tablet computer that calculated the total score automatically.

Description of positive cases definition by index test as reported: not specified (score ranging from 0-30, a higher score represents more deficits but a threshold was not reported; allocation scores by the CVRS was 8 points for hippocampal atrophy, 9 points for cortical atrophy, 6 points for ventricular atrophy, and 7 points for small vessel disease)

Examiners: 3 raters blinded to demographic and clinical information

Interobserver variability: inter-rater and intra-rater reliability with 34 randomly selected MRI scans were 0.941 and 0.936

Target condition and reference standard(s)

Target condition: AD
Prevalence of AD in the sample: 69/340 (20% of cases included in the analysis)
Stable MCI or converted to other dementia: 271/340 (80%); 271 stable MCI
Reference standard: AD diagnosis according to NINCDS-ADRDA ([McKhann 1984](#))
Median clinical follow-up: 3 years

Flow and timing

Withdrawals and losses to follow-up: none reported
 Uninterpretable MRI results have not been reported

Comparative

Key conclusions by the authors

Baseline CVRS predicted the progression to dementia in MCI and was independently associated with longitudinal cognitive decline

Conflict of interests

Funding for this work was derived in part from the following commercial sources: Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC; Johnson & Johnson Pharmaceutical Research & Development LLC; Lumosity; Lundbeck; Merck & Co, Inc; Meso Scale Diagnostics, LLC; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics

Notes

Source of funding: data collection and sharing for this project was funded by ADNI (NIH Grant U01 AG024904) and Department of Defense ADNI (award number W81XWH-12-2-0012)

2 x 2 table: data from the published article; in order to avoid duplicate, we only used the results of lateral ventricles for the purposes of this review

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
------	--------------------	--------------	------------------------

DOMAIN 1: Patient Selection

Was a consecutive or random sample of patients enrolled?

No

Was a case-control design avoided?

Yes

Did the study avoid inappropriate exclusions?

No

Jang 2018 (Continued)

	High	Low
DOMAIN 2: Index Test All tests		
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes	
Did the study provide a clear pre-specified definition of what was considered to be a "positive" result of the index test?	No	
Was the index test performed by a single operator or interpreted by consensus in a joint session?	No	
	High	Low
DOMAIN 3: Reference Standard		
Is the reference standards likely to correctly classify the target condition?	Yes	
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes	
	Low	Low
DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	Yes	
	Low	

Khan 2015

Study characteristics

Patient sampling

Primary objectives:

1. investigate patterns of hippocampal subfield volume loss in MCI and AD
2. determine the pattern of subfield volume loss due to age, gender, education, APOE e4 genotype, and neuropsychological test scores
3. compare combined subfield volumes to hippocampal volume alone at discriminating between AD and healthy controls, and predicting future MCI conversion to AD at 12 months

Study population: MCI participants from ADNI ([ADNI 2010](#)) and AddNeuroMed cohorts ([Lovestone 2009](#)). Data from the ADNI study were downloaded from the ADNI at the LONI website. For the AddNeuroMed cohort, patients attended local memory clinics and received a diagnosis of MCI.

Selection criteria:

1. Participants had MMSE scores between 24 and 30
2. Subjective memory complaint with preserved ADL
3. CDR score of 0.5
4. GDS score ≤ 5
5. absence of dementia in accordance with NINCDS-ADRDA criteria.

Exclusion criteria according to the corresponding study protocols: for ADNI as in [Gaser 2013](#); for AddNeuroMed, other neurological or psychiatric disease, significant unstable systemic illness or organ failure, and alcohol or substance misuse

Study design: prospective longitudinal study (data from ADNI study and AddNeuroMed study)

Patient characteristics and setting

Clinical presentations: MCI with subjective memory complaint

Age mean (SD): MCI who progressed to AD: 74 ± 7 ; stable MCI: 75 ± 7

Gender (% men): MCI who progressed to AD: 60%; stable MCI: 61%

Education years mean (SD): MCI who progressed to AD: 14.0 ± 4.1 ; stable MCI: 14.3 ± 4.4

ApoE4 carriers (%): MCI who progressed to AD: 63%; stable MCI: 48%

Neuropsychological tests: employed; MMSE mean (SD): MCI who progressed to AD: 26.5 ± 1.8 ; stable MCI: 27.1 ± 1.7

Clinical stroke excluded: not specified

Co-morbidities: not reported

Number enrolled: 447

Number available for analysis: 447

Setting: ADNI cohort and AddNeuroMed cohort

Country: USA and Canada (ADNI); Finland, Italy, Greece, UK, Poland, France (AddNeuroMed)

Period: not specified

Language: English

Index tests

Index test: MRI automated method for estimation of hippocampal volume and subvolumes

Manufacturer: several. Standardised MRI data acquisition techniques were in place for AddNeuroMed and ADNI to ensure homogeneity across data acquisition sites. A comprehensive quality control procedure was carried out on all MRIs according to the AddNeuroMed quality control framework

Khan 2015 (Continued)

Tesla strength: 1.5

Assessment methods: image analyses were carried out using the Freesurfer image analysis pipeline (version 5.1.0). Automated segmentation of the hippocampus was performed to define anatomical subfields. Hippocampal subfields were analysed using a supervised multivariate data analysis method included in the software package SIMCA

Description of positive cases definition by index test as reported: the MRI images of MCI participants were classified as "more similar to healthy" (negative cases) and "more similar to AD" (positive cases)

Examiners: no details about radiologist; imaging interpretation was reserved to an automatic classifier

Interobserver variability: not evaluable

Target condition and reference standard(s)

Target condition: AD

Prevalence of AD in the sample: 90/447 (20% enrolled participants)

Stable MCI or converted to other dementia: 357 (80%) stable MCI

Reference standards: not reported in the published article. Referring to the study protocols, NINCDS-ADR-DA criteria were applied for both cohorts (McKhann 1984).

Mean clinical follow-up: 1 year

Flow and timing

Withdrawals explained and losses to follow-up: none reported

Uninterpretable MRI results have not been reported

Comparative

Key conclusions by the authors

Multiple subfield volumes were atrophied in AD and MCI and were related to age, gender, education, APOE e4 genotype, and neuropsychological test scores. For predicting MCI conversion to AD combined subfield volumes and presubiculum volume were more accurate than total hippocampal volume

Conflict of interests

Information not available

Notes

Source of funding: ADNI (Grant U01 AG024904) and Department of Defense ADNI (award number W81XWH-12-2-0012); InnoMed (an integrated project funded by the European Union of the sixth framework program priority FP6-2004-LIFESCIHEALTH-5, Life Sciences, Genomics and Biotechnology for Health)

2 x 2 table: data from the published article; we only used data regarding the total hippocampal volume for the review

Methodological quality
Item
Authors' judgement
Risk of bias
Applicability concerns
DOMAIN 1: Patient Selection

Was a consecutive or random sample of patients enrolled?

No

Was a case-control design avoided?

Yes

Did the study avoid inappropriate exclusions?

No

Khan 2015 (Continued)

	High	Low
DOMAIN 2: Index Test All tests		
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear	
Did the study provide a clear pre-specified definition of what was considered to be a "positive" result of the index test?	Yes	
Was the index test performed by a single operator or interpreted by consensus in a joint session?	Unclear	
	Unclear	Low
DOMAIN 3: Reference Standard		
Is the reference standards likely to correctly classify the target condition?	Yes	
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes	
	Low	Low
DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	Yes	
	Low	

Ledig 2018
Study characteristics

Patient sampling	<p>Primary objectives: assess biomarkers by determining their power to predict diagnostic classification and by comparing atrophy rates to published meta-studies</p> <p>Study population: participants from the ADNI study including MCI with 2-year follow-up</p> <p>Selection criteria: the stable MCI group was represented by “late MCI” at baseline that remained stable for at least 2 years and until the most recent diagnosis that was available; the progressive MCI group was represented by MCI at baseline that converted within 2 years to probable AD. Exclusion criteria: participants who reverted at any time point from a more severe to a less severe disease stage (N = 68), participants with baseline diagnosis of early MCI (N = 277), participants who were diagnosed as MCI at baseline but converted to probable AD > 2 years later (N = 54).</p> <p>Study design: prospective longitudinal study (participants from ADNI study)</p>
Patient characteristics and setting	<p>Clinical presentation: MCI according to ADNI protocol</p> <p>Age median (min; max): MCI who progressed to AD: 74.3 (48.1; 88.3) years; MCI non-converters to AD: 74.4 (55.9; 91.4) years</p> <p>Gender (% men): MCI who progressed to AD: 59%; MCI non-converters to AD: 59%</p> <p>Education years mean (SD): not reported</p> <p>ApoEers (%): MCI who progressed to AD: 68%; MCI non-converters to AD: 44%</p> <p>Neuropsychological tests: employed; MMSE median (min; max) MCI who progressed to AD: 26 (23;30); MCI non-converters to AD: 28 (24; 30)</p> <p>Clinical stroke excluded: not specified</p> <p>Co-morbidities: not reported</p> <p>Number enrolled: 343 (399 excluded a priori)</p> <p>Number available for analysis: 343</p> <p>Setting: ADNI database (multicentre study)</p> <p>Country: USA and Canada</p> <p>Period of study: not reported</p> <p>Language: English</p>
Index tests	<p>Index test: MRI automated method for estimation of brain volumes of hippocampal, entorhinal cortex, amygdala, middle temporal gyrus, ventricles, cortical grey matter, white matter, deep grey matter, brain tissue</p> <p>Manufacturer: those used in ADNI study (GE Healthcare, Philips Medical System, Siemens Medical Solution (Jack 2008b))</p> <p>Tesla strength: 1.5-3 T</p> <p>Assessment methods: brain extraction (pinfram) and MRI automatic segmentation method for robust segmentation of whole brain MRIs into 138 distinct anatomical structures using Multi-Atlas Label Propagation with Expectation–Maximisation based refinement (MALPEM) (Ledig 2015)</p> <p>Description of positive cases definition by index test as reported: not specified (a threshold was not reported)</p> <p>Examiners: not specified</p> <p>Interobserver variability: not specified in the paper; test-retest reliability for the method was estimated in Ledig 2015: using MALPEM the average ICC was 0.97 for non-cortical and 0.94 for cortical regions</p>
Target condition and reference standard(s)	<p>Target condition: AD</p> <p>Prevalence of AD in the sample: 177/343 (52% of cases included in the analysis)</p> <p>Stable MCI or converted to other dementia: 166/343 (48%) stable MCI</p> <p>Reference standard: AD diagnosis according to NINCDS-ADRDA (McKhann 1984)</p> <p>Mean clinical follow-up: 2 years</p>
Flow and timing	<p>Withdrawals and losses to follow-up: none reported</p>

Ledig 2018 (Continued)

Uninterpretable MRI results have not been reported

Comparative

Key conclusions by the authors The identified biomarkers hold great potential for deeper analysis, and the validated methodology can readily be applied to other imaging cohorts.

Conflict of interests CL, AS and RG conducted this research while being employees of Imperial College London, UK (CL, AS, RG) and IXICO plc, UK (CL, RG). DR is a co-founder and scientific advisor of IXICO plc, UK, a provider of medical image analysis services. CL is currently employed by Imagen Technologies, Inc, NY, USA. This does not alter their adherence to Scientific Reports policies on sharing data and materials.

Notes **Source of funding:** European Union's Seventh Framework Programme under grant agreement no. 611005. RG was funded by an Innovative UK (101685) grant. Data collection and sharing for this project was funded by the ADNI and Department of Defense ADNI.

2 x 2 table: data from the published article; we considered volume estimation data for these regions: hippocampal, entorhinal cortex, amygdala, middle temporal gyrus, lateral ventricles, cortical grey matter, whole brain

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
------	--------------------	--------------	------------------------

DOMAIN 1: Patient Selection

Was a consecutive or random sample of patients enrolled?	No		
--	----	--	--

Was a case-control design avoided?	Yes		
------------------------------------	-----	--	--

Did the study avoid inappropriate exclusions?	No		
---	----	--	--

High
Low
DOMAIN 2: Index Test All tests

Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
---	---------	--	--

Did the study provide a clear pre-specified definition of what was considered to be a "positive" result of the index test?	No		
--	----	--	--

Was the index test performed by a single operator or interpreted	Unclear		
--	---------	--	--

Ledig 2018 (Continued)
 by consensus in a joint session?

	High	Low
DOMAIN 3: Reference Standard		
Is the reference standards likely to correctly classify the target condition?	Yes	
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes	
	Low	Low
DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	Yes	
	Low	

Liu 2010

Study characteristics

Patient sampling	<p>Primary objectives: determine predictors of conversion to AD from MCI with automated MRI regional cortical volume and thickness measures</p> <p>Study population: amnesic MCI</p> <p>Selection criteria: data were collected from 6 medical centres across Europe. All the MCI participants had successfully undergone MRI and cognitive tests evaluated at baseline, and clinical evaluation and cognitive tests were repeated 1 year later. None of the MCI and AD participants had other neurological or psychiatric disease, significant unstable systemic illness or organ failure, and alcohol or substance misuse</p> <p>Study design: prospective longitudinal study (data from AddNeuroMed study; Lovestone 2009)</p>
Patient characteristics and setting	<p>Clinical presentations: amnesic MCI according to Petersen 1999 and Petersen 2001 criteria:</p> <ol style="list-style-type: none"> 1. memory complaint by patient, family, or physician 2. normal ADL

Liu 2010 (Continued)

3. MMSE score range between 24 and 30
4. GDS score ≤ 5
5. participant aged ≥ 65 years
6. CDR memory score of 0.5 or 1
7. absence of dementia according to NINCDS/ADRDA criteria ([McKhann 1984](#))

Age mean (SD): MCI who progressed to AD: 72 ± 6 ; stable MCI: 74 ± 6

Gender (% men): MCI who progressed to AD: 57%; stable MCI: 44%

Education years mean (SD): MCI who progressed to AD 9 ± 4 ; stable MCI: 9 ± 4

ApoE ϵ 4 carriers (%): not stated

Neuropsychological tests: MMSE mean (SD): MCI who progressed to AD: 27 ± 2 ; stable MCI: 27 ± 2

Clinical stroke excluded: not specified

Co-morbidities: not reported

Number enrolled: 100

Number available for analysis: 100

Setting: AddNeuroMed cohort ([Lovestone 2009](#))

Country: Finland, Italy, Greece, UK, Poland, France

Period: not reported

Language: English

Index tests

Index test: MRI automated method for estimation of hippocampal volume

Manufacturer: 6 different MRI systems (4 Signa, GE, Waukesha, WI; 1 Avanto, Siemen, Erlangen, Germany; and 1 Edge 1.5T, Picker, Cleveland, OH)

Tesla strength: 1.5

Assessment methods: a highly automated structural MRI image processing pipeline, which was developed by Fischl and his colleagues was utilised for data analysis ([Fischl 2004](#)). It produced both regional cortical thickness measures from 34 areas and regional volume measures from 24 areas, including the hippocampus

Description of positive cases definition by index test as reported: not specified

Examiners: details not reported

Interobserver variability: not provided

Target condition and reference standard(s)

Target condition: AD

Prevalence of AD in the sample: 21/100 (21% of enrolled participants)

Stable MCI or converted to other dementia: 79 (79%) stable MCI

Reference standards: NINCDS/ADRDA criteria ([McKhann 1984](#)).

Mean clinical follow-up: 1 year

Flow and timing

Withdrawals explained and losses to follow-up: none reported

Uninterpretable MRI results have not been reported

Liu 2010 (Continued)

Comparative

Key conclusions by the authors The cortical volumes achieved higher diagnostic accuracy than did cognitive tests or cortical thickness. Combining the volumes, thickness, and cognitive tests did not improve the test accuracy. The volume of amygdala and caudate were independent variables in predicting conversion from MCI to dementia due to AD

Conflict of interests The authors declared no conflicts of interest

Notes **Source of funding:** the study was funded by the European Union, AddNeuroMed/Innovative Medicines LSHB-CT-2005-518170; Yawu Liu was funded by Health Research Council of the Academy of Finland, grant 121038, and EVO grants 577209 and 5772720 from Kuopio University Hospital

2 x 2 table: data from the published article; only hippocampal volume measure was considered for the review purpose

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
------	--------------------	--------------	------------------------

DOMAIN 1: Patient Selection

Was a consecutive or random sample of patients enrolled?	No		
--	----	--	--

Was a case-control design avoided?	Yes		
------------------------------------	-----	--	--

Did the study avoid inappropriate exclusions?	Yes		
---	-----	--	--

High
Low
DOMAIN 2: Index Test All tests

Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
---	---------	--	--

Did the study provide a clear pre-specified definition of what was considered to be a "positive" result of the index test?	No		
--	----	--	--

Was the index test performed by a single operator or interpreted by consensus in a joint session?	Unclear		
---	---------	--	--

High
Low
DOMAIN 3: Reference Standard

Liu 2010 (Continued)

Is the reference standards likely to correctly classify the target condition? Yes

Were the reference standard results interpreted without knowledge of the results of the index tests? Yes

Low

Low

DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard? Yes

Did all patients receive the same reference standard? Yes

Were all patients included in the analysis? Yes

Low

Monge Argilés 2014
Study characteristics

Patient sampling	<p>Primary objectives: compare the early diagnostic utility of AD biomarkers in the CSF with those in brain MRI in conditions found in our clinical practice, and to ascertain the diagnostic accuracy of both techniques used together</p> <p>Study population: pure amnesic MCI or multidomain MCI</p> <p>Selection criteria: patients with amnesic MCI, > 55 years, MMSE score of 20-27 and score of < 78 on the informant questionnaire. Participants signed an informed consent form to be included in the study and to undergo lumbar puncture. Exclusion criteria: dementia or any other neurological, psychiatric or systemic condition that could lead to cognitive impairment; anticoagulant treatment; absence of informed consent; GDS score > 5</p> <p>Study design: prospective longitudinal study</p>
Patient characteristics and setting	<p>Clinical presentations: pure amnesic MCI or multidomain MCI according to Petersen criteria of 2006 (Yaffe 2006)</p> <p>Age mean (SD): MCI who progressed to AD: 73 ± 7; stable MCI: 73 ± 7</p> <p>Gender (% men): MCI who progressed to AD: 33%; stable MCI: 47%</p> <p>Education years mean: MCI who progressed to AD: 6; stable MCI: 4.3</p> <p>ApoEε4 carriers (%): not stated</p> <p>Neuropsychological tests: employed; MMSE mean (SD): MCI who progressed to AD: 23 ± 1.2; stable MCI: 24 ± 2.4</p>

Monge Argilés 2014 (Continued)

Clinical stroke excluded: not specified

Co-morbidities: not reported

Number enrolled: 30

Number available for analysis: 30

Setting: Alicante Hospital Universitario

Country: Spain

Period: 2008-2009

Language: English

Index tests	<p>Index test: MRI visual method for estimation of MTA</p> <p>Manufacturer: GE</p> <p>Tesla strength: 1.5 Tesla</p> <p>Assessment methods: 2 radiologists visually quantified MTA according to the method described by Korf 2004. According to Scheltens 1997, the MTA scale ranges from 0 (no atrophy)-4 (severe atrophy) and takes into account the width of the choroid fissure, the height of the hippocampus, and the width of the temporal horn. The MTA scale was applied to the right and left medial temporal lobe</p> <p>Description of positive cases definition by index test as reported: the summed score of left and right temporal lobes was used as well as the dichotomised summed score: no atrophy (score 0-2) and atrophy (score ≥ 3)</p> <p>Examiners: the radiologists were skilled and blinded to clinical data</p> <p>Interobserver variability: ICCs between the 2 radiologists were 0.80 and 0.85</p>
Target condition and reference standard(s)	<p>Target condition: AD</p> <p>Prevalence of AD in the sample: 15/30 (50% of enrolled participants)</p> <p>Stable MCI or converted to other dementia: 15 (50%) stable MCI</p> <p>Reference standards: NINCDS-ADRDA criteria (McKhann 1984).</p> <p>Mean clinical follow-up: 2 years</p>
Flow and timing	<p>Withdrawals explained and losses to follow-up: none reported</p> <p>Uninterpretable MRI results have not been reported</p>
Comparative	
Key conclusions by the authors	<p>Diagnostic accuracy of biomarkers in cerebrospinal fluid is higher than that of biomarkers in MRI. Combined use of both techniques is highly accurate for either early diagnosis or exclusion of AD in patients with MCI</p>
Conflict of interests	<p>Sutdy authors declare having no conflict of interest</p>
Notes	<p>Source of funding: the study was partially funded by Novartis Espana and Grunenthal Espana</p> <p>2 x 2 table: data from the published article</p>

Methodological quality

Monge Argilés 2014 (Continued)

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	No		
		High	Low
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Did the study provide a clear pre-specified definition of what was considered to be a "positive" result of the index test?	Yes		
Was the index test performed by a single operator or interpreted by consensus in a joint session?	Unclear		
		Low	Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
		Low	Low
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		

Monge Argilés 2014 (Continued)

Were all patients included in the analysis? Yes

Low

Nesteruk 2016
Study characteristics

Patient sampling	<p>Primary objectives: evaluate the usefulness of several biomarkers in predicting the conversion of MCI to AD: β-amyloid and tau proteins in CSF and the volumetric evaluation of brain structures including the hippocampus in MRI</p> <p>Study population: MCI diagnosed in the Alzheimer's Department</p> <p>Selection criteria: MCI according to Winblad 2004: not normal, not demented; self- and/or informant-report and impairment on objective cognitive tasks; evidence of decline over time on objective cognitive tasks and/or preserved basic ADL/minimal impairment in complex instrumental functions. The concept of MCI is comprehensive of heterogeneity of clinical presentation (amnestic/nonamnestic/single domain/multiple domains) and different aetiologies. Exclusion criteria not specified</p> <p>Study design: prospective longitudinal study</p>
Patient characteristics and setting	<p>Clinical presentations: definition of MCI according to Petersen 2004; all participants received a CDR score of 0.5</p> <p>Age mean (SD): MCI who progressed to AD 70 ± 10; stable MCI: 61 ± 9</p> <p>Gender (% men): 45%</p> <p>Education years mean: MCI who progressed to AD: 13.33 ± 3.43; stable MCI: 14.13 ± 2.74</p> <p>ApoE4 carriers (%): not reported</p> <p>Neuropsychological tests: employed; MMSE mean (SD): MCI who progressed to AD: 27.22 ± 1.56; stable MCI: 27.58 ± 1.79</p> <p>Clinical stroke excluded: not specified</p> <p>Co-morbidities: not reported; multiple aetiologies were considered in the definition of MCI</p> <p>Number enrolled: 40</p> <p>Number available for analysis: 40</p> <p>Setting: Warsaw memory clinic</p> <p>Country: Poland</p> <p>Period: not reported</p> <p>Language: English</p>
Index tests	<p>Index test: MRI automated method for estimation of volumetric measures of the hippocampus and entorhinal cortex, posterior cingulate gyrus, parahippocampal gyrus, superior, medial, inferior temporal gyri</p> <p>Manufacturer: Toshiba</p> <p>Tesla strength: 1.5</p>

Nesteruk 2016 (Continued)

Assessment methods: volume estimation was performed using Freesurfer software (version not specified).

Discriminant analysis was conducted in MCI separately for all volumetric measurements, for CSF biomarkers, and for volumetric and CSF biomarkers together. Classification rate was available for single volumetric regions. No references to a validation of the classification method

Description of positive cases definition by index test as reported: positive cases were defined presumably after the discriminant analysis

Examiners: no details. It is not clear if Freesurfer operator conducted also the discriminant analysis.

Interobserver variability: not provided

Target condition and reference standard(s)

Target condition: AD

Prevalence of AD in the sample: 9/40 (22.5% of enrolled participants)

Stable MCI or converted to other dementia: 31 (77.5%) stable MCI

Reference standards: NIA-AA criteria (McKhaan 2011). It was not specified if only the core clinical criteria were used for the diagnosis of AD or if MRI results were also used to support the diagnosis

Mean clinical follow-up: 0.8 ± 0.5 (conversion was evaluated within 2 years)

Flow and timing

Withdrawals explained and losses to follow-up: none reported

Uninterpretable MRI results have not been reported

Comparative

Key conclusions by the authors

Biomarkers seem to be important parameters in predicting the conversion of MCI to AD, in particular when biochemical biomarkers are used together with volumetric ones

Conflict of interests

Study authors reported no conflicts of interest

Notes

Source of funding: not reported

2 x 2 table: data from the published article

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
------	--------------------	--------------	------------------------

DOMAIN 1: Patient Selection

Was a consecutive or random sample of patients enrolled?

No

Was a case-control design avoided?

Yes

Did the study avoid inappropriate exclusions?

Unclear

High

Low

DOMAIN 2: Index Test All tests

Nesteruk 2016 (Continued)

Were the index test results interpreted without knowledge of the results of the reference standard? Unclear

Did the study provide a clear pre-specified definition of what was considered to be a "positive" result of the index test? No

Was the index test performed by a single operator or interpreted by consensus in a joint session? Unclear

High
Low
DOMAIN 3: Reference Standard

Is the reference standards likely to correctly classify the target condition? Yes

Were the reference standard results interpreted without knowledge of the results of the index tests? Unclear

Unclear
Unclear
DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard? Yes

Did all patients receive the same reference standard? Yes

Were all patients included in the analysis? Yes

Low
Ong 2015
Study characteristics

Patient sampling **Primary objectives:** assess the clinical utility of β -amyloid imaging with 18F-florbetaben in MCI by evaluating its prognostic accuracy for progression to AD, comparing semiquantitative with visual scan assessment, and exploring the relationships among β -amyloid imaging, hippocampal volume (HV) and memory over time

Study population: participants with MCI referred from local memory clinics

Ong 2015 (Continued)

Selection criteria: entry criteria were MCI with presentation of progressive cognitive decline and at least 1 neuropsychological test score falling 1.5 SDs below published means. Exclusion criteria not reported

Study design: prospective longitudinal study

Patient characteristics and setting

Clinical presentations: MCI according to [Petersen 2004](#) criteria

Age mean (SD): MCI with normal hippocampal volume: 71 ± 6; MCI with hippocampal atrophy: 75 ± 7

Gender (% men): not stated

Education years mean (SD): MCI with normal hippocampal volume: 12.4 ± 3.2; MCI with hippocampal atrophy: 15.3 ± 3.5

ApoE4 carriers (%): not stated

Neuropsychological tests: employed; MMSE mean (SD): MCI with normal hippocampal volume: 27.6 ± 1.8; MCI with hippocampal atrophy: 26.9 ± 1.7

Clinical stroke excluded: not specified

Co-morbidities: not reported

Number enrolled: 45

Number available for analysis: 45

Setting: local memory clinics

Country: Australia

Period: May 2008-December 2009

Language: English

Index tests

Index test: MRI automated method for estimation of hippocampal volume

Manufacturer: not specified

Tesla strength: not specified

Assessment methods: hippocampal volume was derived from T1 MPRAGE MRI sequence using Neuroquant software

Description of positive cases definition by index test as reported: cut-off for hippocampal atrophy was determined by double ROC analysis on the hippocampal volumes measured by Neuroquant of 23 AD participants and 143 healthy controls from the "Australian Imaging, Biomarkers and Lifestyle (AIBL)" study

Examiners: blindness was specified only for the amyloid PET rater

Interobserver variability: not provided

Target condition and reference standard(s)

Target condition: AD

Diagnostic classification at 2 years was performed by a neurologist blind to PET and quantitative MRI results

Prevalence of AD in the sample: 20/45 (44% of enrolled participants)

Stable MCI or converted to other dementia: 25 (56%) MCI not converted to AD (21 stable MCI and 4 MCI converted to other dementia: 1 to progressive supranuclear palsy, 2 to frontotemporal lobar degeneration, 1 to LBD)

Ong 2015 (Continued)

Reference standards: NINCDS-ADRDA criteria (McKhann 1984) for AD

Mean clinical follow-up: 2 years

Flow and timing	Withdrawals explained and losses to follow-up: none reported Uninterpretable MRI results have not been reported
Comparative	
Key conclusions by the authors	Amyloid imaging facilitates accurate detection of prodromal AD. As neurodegeneration progresses, and in contrast with the early stages of the disease, hippocampal atrophy and not amyloid status, seems to drive memory decline
Conflict of interests	Study authors declare their competing interests
Notes	Source of funding: study was initiated by Professor CC Rowe, sponsored by Bayer Healthcare AG, Berlin, Germany, and funded in part by NHMRC grant 509166 2 x 2 table: data from the published article

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
		High	Low
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
Did the study provide a clear pre-specified definition of what was considered to be a "positive" result of the index test?	Yes		
Was the index test performed by a single operator or interpreted by consensus in a joint session?	Unclear		
		Unclear	Low

Ong 2015 (Continued)

DOMAIN 3: Reference Standard

Is the reference standards likely to correctly classify the target condition? Yes

Were the reference standard results interpreted without knowledge of the results of the index tests? Yes

Low

Low

DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard? Yes

Did all patients receive the same reference standard? Yes

Were all patients included in the analysis? Yes

Low

Pereira 2014
Study characteristics

Patient sampling	<p>Primary objectives: assess the influence of age, disease onset and ApoE4 on visual MTA cut-offs</p> <p>Study population: participants with MCI from 2 large independent cohorts: ADNI and AddNeuroMed. Participants were recruited through local hospital and memory clinics.</p> <p>Selection criteria: participants with MCI and clinical follow-up at 1 year were included. Inclusion criteria for both cohorts: MMSE score 24-30, memory complaint reported by the patient, family member or physician, normal ADL, CDR memory score of 0.5 or 1 (total CDR = 0.5), memory loss measured by the WMS Logical Memory II for the ADNI cohort only, GDS score of ≤ 5, age ≥ 65 years, stable medication and good general health. Exclusion criteria: meeting the DSM-IV and NINCDS-ADRDA criteria for AD, significant neurological or psychiatric illness other than AD and significant unstable systemic illness or organ failure</p> <p>Study design: prospective longitudinal study. Participants from ADNI (ADNI 2010) and AddNeuroMed studies (Lovestone 2009)</p>
Patient characteristics and setting	<p>Clinical presentations: amnesic MCI</p> <p>Age mean (SD): MCI who progressed to AD: 74 ± 6.5; stable MCI: 75 ± 7</p> <p>Gender (% men): MCI who progressed to AD: 59%; stable MCI: 61%</p> <p>Education years mean (SD): MCI who progressed to AD: 13.7 ± 4.2; stable MCI: 14.0 ± 4.6</p> <p>ApoE4 carriers (%): MCI who progressed to AD: 62.1%; stable MCI: 46%</p> <p>Neuropsychological tests: employed; MMSE mean (SD): MCI who progressed to AD: 26.5 ± 1.8; stable MCI 27.1 ± 1.7</p>

Pereira 2014 (Continued)

Clinical stroke excluded: not specified

Co-morbidities: not reported

Number enrolled: 480

Number available for analysis: 480

Setting: ADNI cohort and AddNeuroMed cohort

Country: USA and Canada (ADNI); Finland, Italy, Greece, UK, Poland, France (AddNeuroMed)

Period: not reported

Language: English

Index tests

Index test: MRI visual method for estimation of MTA

Manufacturer: standardised MRI data acquisition techniques were in place for AddNeuroMed and ADNI to ensure homogeneity across data acquisition sites. A comprehensive quality control procedure was carried out on all MRI images according to the AddNeuroMed quality control framework

Tesla strength: not specified (information from study protocols: 1.5-3 T for ADNI, 1.5 T for AddNeuroMed)

Assessment methods: for each participant, MTA was rated on a single MRI slice posterior to the amygdala and the mammillary bodies, positioned in such a way that the hippocampus, cerebral peduncles and pons were all visible. MTA score attributed according to [Scheltens 1992](#). The right and left sides of the medial temporal lobe are rated separately

Description of positive cases definition by index test as reported: 2 different and independent cut-off values: the age-dependent cut-off (an MTA score of ≥ 2 was considered abnormal for participants < 75 years, whereas a score of ≥ 3 was considered abnormal for participants > 75 years) and the averaged left and right cut-off (the average of the MTA scores of both hemispheres with a resulting score ≥ 1.5 was considered abnormal). We used the averaged cut-off for this review.

Examiners: single MTA rater (LC) was blind to gender, age and diagnosis. High intrarater reliability (weighted kappa 0.93 and 0.94)

Interobserver variability: a highly significant correlation was found between the MTA score and manual delineation by hippocampal volume by another experienced radiologist

Target condition and reference standard(s)

Target condition: AD

Prevalence of AD in the sample: 95/480 (20% of enrolled participants)

Stable MCI or converted to other dementia: 385 (80%) stable MCI

Reference standards: not specified in the published article. NINCDS-ADRDA criteria ([McKhann 1984](#)) were used according to the study protocols (as reported in [Gaser 2013](#) and [Liu 2010](#))

Mean clinical follow-up: 1 year

Flow and timing

Withdrawals explained and losses to follow-up: none reported

Uninterpretable MRI results have not been reported

Comparative

Key conclusions by the authors

Clinical, demographic and genetic variables can influence the classification of MTA cut-off scores, leading to misdiagnosis in some cases. These variables, in addition to the differential sensitivity and specificity of each cut-off, should be carefully considered when performing visual MTA assessment.

Conflict of interests

Study authors report no conflict of interest

Pereira 2014 (Continued)

Notes

Source of funding: study was supported by InnoMed (Innovative Medicines in Europe), an Integrated Project funded by the European Union of the Sixth Framework programme priority FP6-2004-LIFESCIHEALTH-5. Data collection and sharing for this project was funded by the ADNI NIH Grant U01 AG024904. This research was also supported by NIH Grants P30 AG010129 and K01 AG030514

2 x 2 table: data from the published article

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		High	Low
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Did the study provide a clear pre-specified definition of what was considered to be a "positive" result of the index test?	Yes		
Was the index test performed by a single operator or interpreted by consensus in a joint session?	Yes		
		Low	Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results in-	Yes		

Pereira 2014 (Continued)

terpreted without knowledge of the results of the index tests?

Low

Low

DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard? Yes

Did all patients receive the same reference standard? Yes

Were all patients included in the analysis? Yes

Low

Platero 2019
Study characteristics

Patient sampling **Primary objectives:** discriminate AD progression using a new hippocampal marker from T1-wighted MRI: the local surface roughness

Study population: participants were recruited from the Hospital Universitario San Carlos (Madrid), the Centre for Prevention of Cognitive Impairment (Madrid) and the Senior Center of the district of Chamartin (Madrid)

Selection criteria: clinical diagnosis of MCI included the following features:

1. self- or informant-reported cognitive complaints
2. objective evidence of impairment in ≥ 1 cognitive domains
3. preserved independence in functional abilities
4. not demented.

Exclusion criteria not reported. All cases of MCI categorised as "MCI due to AD" with an intermediate level of likelihood ([Albert 2011](#))

Study design: prospective study (no details)

Patient characteristics and setting **Clinical presentation:** amnesic MCI. In addition to meeting the clinical criteria, MCI participants showed signs of loss of the hippocampal volume compared with controls; they were definable as "MCI due to AD" with an intermediate likelihood according to [Albert 2011](#). Hippocampal volumes were not used to establish any of the different diagnoses that were explored, since clinical and cognitive performance were used for this purpose

Age mean (SD): MCI who progressed to AD: 75.6 ± 4.9 years; MCI non-converters to AD: 73.2 ± 5.2 years

Gender (% men): MCI who progressed to AD: 39%; MCI non-converters to AD: 36%

Education years mean (SD): MCI who progressed to AD: 7.94 ± 4.07 years; MCI non-converters to AD: 8.77 ± 4.36 years

Platero 2019 (Continued)

ApoE4 carriers (%): not reported
Neuropsychological tests: employed; MMSE mean (SD): MCI who progressed to AD: 25.8 ± 3.1; MCI non-converters to AD: 27.0 ± 2.3

Clinical stroke excluded: not reported

Co-morbidities: not reported

Number enrolled: 137

Number available for analysis: 97

Setting: Hospital Universitario San Carlos (Madrid), the Centre for Prevention of Cognitive Impairment (Madrid) and the Senior Center of the district of Chamartin (Madrid)

Country: Spain

Period of study: not specified

Language: English

Index tests

Index test: automated method for estimation of hippocampal volume; hippocampal surface roughness and local surface roughness were also evaluated.

Manufacturer: GE

Tesla strength: 1.5 T

Assessment methods: the markers were extracted from the automated hippocampal segmentation. For more robustness in terms of hippocampal segmentation errors, both left and right volumes were averaged and normalised with ICV.

Description of positive cases definition by index test as reported: not specified

Examiners: no details about radiologist

Interobserver variability: not reported

Target condition and reference standard(s)

Target condition: AD

Prevalence of AD in the sample: 36/97 (37% of cases included in the analysis)

Stable MCI or converted to other dementia: 61/97 (63%); 61 stable MCI

Reference standard: NINCDS-ADRDA criteria

Mean clinical follow-up: 3 years

Flow and timing

Withdrawals and losses to follow-up: 40 lost (missing follow-up)

Uninterpretable MRI results have not been reported

Comparative

Key conclusions by the authors

The LSR [local surface roughness] marker show better prediction of conversion to AD than normalised hippocampal volume. The results suggest the relevance of considering the LSR as a new hippocampal marker for the AD continuum.

Conflict of interests

No details

Notes

Source of funding: Spanish Ministry of Economy and Competitiveness, Grant/Award Numbers: IJCI-2016-30662, PSI2012-38375-C03-01, PSI2009-14415-C03-01

2 x 2 table: data from the published article

Methodological quality
Item
Authors' judgement
Risk of bias
Applicability concerns

Platero 2019 (Continued)

DOMAIN 1: Patient Selection

Was a consecutive or random sample of patients enrolled? No

Was a case-control design avoided? Yes

Did the study avoid inappropriate exclusions? No

High
Low
DOMAIN 2: Index Test All tests

Were the index test results interpreted without knowledge of the results of the reference standard? Unclear

Did the study provide a clear pre-specified definition of what was considered to be a "positive" result of the index test? No

Was the index test performed by a single operator or interpreted by consensus in a joint session? Unclear

High
Low
DOMAIN 3: Reference Standard

Is the reference standards likely to correctly classify the target condition? Yes

Were the reference standard results interpreted without knowledge of the results of the index tests? Yes

Low
Low
DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard? Yes

Did all patients receive the same reference standard? Yes

Platero 2019 (Continued)

Were all patients included in the analysis? No

High

Prestia 2013

Study characteristics

Patient sampling	<p>Primary objectives: provide clinical evidence that brain amyloidosis biomarkers turn abnormal earlier than other biomarkers in patients with MCI converting to AD</p> <p>Study population: participants with MCI coming to observation at 3 independent memory clinics: Translational Outpatient Memory Clinic (TOMC), VU University Medical Center (VUMC), Karolinska University Hospital Huddinge (KUHH)</p> <p>Selection criteria: all patients with diagnosis of MCI at baseline, with available baseline MRI, FDG-PET, CSF sampling and clinically followed to ascertain incident AD dementia. Patients included in the study were MCI at baseline and remained stable or developed AD during follow-up. Exclusion criteria: not reported</p> <p>Study design: prospective longitudinal study (participants coming from 3 cohorts)</p>
Patient characteristics and setting	<p>Clinical presentations: in all 3 memory clinics diagnosis of MCI at baseline was made according to Peterson 1999 criteria.</p> <p>Age mean (SD): MCI who progressed to AD: 68 ± 9; stable MCI: 65 ± 9</p> <p>Gender (% men): MCI who progressed to AD: 38%; stable MCI: 48%</p> <p>Education years mean (SD): not reported</p> <p>ApoE4 carriers (%): MCI who progressed to AD: 58%; stable MCI: 51%</p> <p>Neuropsychological tests: employed; MMSE mean (SD): MCI who progressed to AD: 26.7 ± 1.6; stable MCI 27.5 ± 1.8</p> <p>Clinical stroke excluded: yes for the TOMC cohort; not specified for VUMC and KUHH cohorts</p> <p>Co-morbidities: not reported</p> <p>Number enrolled: 73 (31 from TOMC, 25 VUMC, 17 KUHH)</p> <p>Number available for analysis: 73 (31 from TOMC, 25 VUMC, 17 KUHH)</p> <p>Setting: 3 independent memory clinics: TOMC in Brescia, VUMC in Amsterdam, KUHH in Stockholm</p> <p>Country: Italy, the Netherlands, Sweden</p> <p>Period: not reported</p> <p>Language: English</p>
Index tests	<p>Index test: MRI automated method for estimation of hippocampal volume</p> <p>Manufacturer: Philips for TOMC cohort (Caroli 2007); Siemens for VUMC and KUHH cohorts</p> <p>Tesla strength: 1.0 for TOMC; 1.5 for VUMC; 3 for KUHH</p> <p>Assessment methods: left and right hippocampal volumes were automatically computed using Freesurfer.</p>

Prestia 2013 (Continued)

The smallest between left and right volumes was retained for analyses. As the processing procedures do not account for age, pertinent age-corrected scores, hereafter called W-scores, were computed and retained for statistical analyses.

Description of positive cases definition by index test as reported: hippocampal volume abnormality ($W < -2.90$) was defined as W-score below the 5th percentile of its distribution in a group of 143 cognitively healthy elderly people taken from the ADNI dataset

Examiners: no details

Interobserver variability: not provided

Target condition and reference standard(s)	<p>Target condition: AD</p> <p>Prevalence of AD in the sample: 29/73 (40% of enrolled participants); 18 from TOMC, 6 VUMC, 5 KUHH</p> <p>Stable MCI or converted to other dementia: 44 (60%) stable MCI (13 from TOMC, 19 VUMC, 12 KUHH)</p> <p>Reference standards: NINCDS-ADRDA criteria (McKhann 1984)</p> <p>Mean clinical follow-up: 1.9 ± 1.3 years for MCI who progressed to AD; 2.7 ± 1.4 years for stable MCI</p>
Flow and timing	<p>Withdrawals explained and losses to follow-up: none</p> <p>Uninterpretable MRI results have not been reported</p>
Comparative	
Key conclusions by the authors	The core biomarker pattern is in line with the current pathophysiologic model of AD. Fully normal and fully abnormal pattern is associated with exceptional and universal development of dementia. Cases not in line might be due to atypical neurobiology or inaccurate thresholds for biomarker (ab)normality
Conflict of interests	The study authors report no disclosures relevant to the manuscript. Invitation to go to Neurology.org for full disclosures
Notes	<p>Source of funding: Swedish Research Council (project 05817), Strategic Research Program in Neuroscience at Karolinska Institutet, Swedish Brain Power, Sottoprogetto finalizzato strategico 2006, Programma Strategico 2006, Convenzione 71; Programma Strategico 2007, Convenzione PS39, Ricerca Corrente Italian Ministry of Health</p> <p>2 x 2 table: data to complete 2 x 2 table provided by the study authors</p>

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	No		
		High	Low

Prestia 2013 (Continued)

DOMAIN 2: Index Test All tests

Were the index test results interpreted without knowledge of the results of the reference standard? Unclear

Did the study provide a clear pre-specified definition of what was considered to be a "positive" result of the index test? Yes

Was the index test performed by a single operator or interpreted by consensus in a joint session? Unclear

Unclear

Low

DOMAIN 3: Reference Standard

Is the reference standards likely to correctly classify the target condition? Yes

Were the reference standard results interpreted without knowledge of the results of the index tests? Yes

Low

Low

DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard? Yes

Did all patients receive the same reference standard? Yes

Were all patients included in the analysis? Yes

Low

Prestia 2013 (ADNI)
Study characteristics

Patient sampling **Primary objectives:** capitalise on data from different clinical series to compare sensitivity and specificity of individual biomarkers for predicting MCI progression to AD

Study population: MCI patients come from two independent data sets: ADNI and Translational Outpatient Memory Clinic (TOMC)

Selection criteria: patients included in the study were all MCI with prodromal AD taken from ADNI and TOMC databases, with available baseline MRI, FDG-PET and CSF sampling. MCI who converted to non-AD dementia were excluded from the study. Stable MCI coming from the same databases were included. Exclusion criteria according to ADNI protocol for the ADNI cohort, as in [Gaser 2013](#). For the TOMC cohort all patients were self-referred or referred by general practitioners or specialists, complaining of memory or other cognitive disturbances unaccountable by focal cerebral, physical, psychiatric, or metabolic diseases. Patients who converted to non-AD dementia were excluded from the study.

Study design: prospective longitudinal study

Patient characteristics and setting

Clinical presentations: MCI according to [Petersen 1999](#)

Age mean (SD): MCI who progressed to AD in ADNI: 75 ± 8; stable MCI in ADNI: 75 ± 8; MCI who progressed to AD in TOMC: 71 ± 8; stable MCI in TOMC: 72 ± 8

Gender (% men): MCI who progressed to AD in ADNI: 58%; stable MCI in ADNI: 61%; MCI who progressed to AD in TOMC: 33%; stable MCI in TOMC: 50%

Education years mean (SD): not reported

ApoE4 carriers (%): not reported

Neuropsychological tests: employed; MMSE mean (SD): MCI who progressed to AD in ADNI: 28 ± 2; stable MCI in ADNI: 27 ± 2; MCI who progressed to AD in TOMC: 26 ± 2; stable MCI in TOMC: 26 ± 2

Clinical stroke excluded: yes in the TOMC cohort, not specified in ADNI cohort

Co-morbidities: not reported

Number enrolled: 57 MCI from ADNI, 36 MCI from TOMC

Number available for analysis: 57 MCI from ADNI, 36 MCI from TOMC (manual method), 32 MCI from TOMC (automated method)

Setting: ADNI and TOMC databases

Country: USA and Canada for ADNI; Italy for TOMC

Period: not reported

Language: English

Index tests

Index test: MRI manual, automated, semiautomated methods for estimation of hippocampal volume

Manufacturer: several for the ADNI cohort; Philips for TOMC ([Caroli 2007](#))

Tesla strength: high-resolution MRI (Tesla strength not specified in the published article, anyway 1.5-3 T for the ADNI cohort according to the study protocol, 1.0 T for the TOMC cohort as reported in ([Caroli 2007](#)))

Assessment methods: Freesurfer software version 4.5.0 was used in ADNI and TOMC cohort for automated volume estimation, DISPLAY was used in the TOMC cohort for manual volume estimation, Medtronic Surgical Navigation Technologies tool was used in the ADNI cohort for semiautomated volume estimation. Manual tracing was performed according the protocol of Pruessner 2000. The semiautomated method was applied basing on fluid imaging transformation according to Christensen 1997

Prestia 2013 (ADNI) *(Continued)*

Description of positive cases definition by index test as reported: hippocampal volume cutoffs were computed specifically based on hippocampal volume performance in correctly identifying a group of 287 cognitively healthy elders taken from a reference normative database (manual segmentation) or a group of 66 ADNI cognitively healthy elders (both semi-automated and automated procedures).

Examiners: a single rater performed the manual method. Regarding the semiautomated method, boundaries were checked by qualified reviewers and in case of failure, edited manually

Interobserver variability: not reported

Target condition and reference standard(s)	<p>Target condition: AD</p> <p>Prevalence of AD in the sample: 24/57 (42%) of enrolled participants in the ADNI cohort; 18/36 (50%) of enrolled participants in the TOMC cohort</p> <p>Stable MCI or converted to other dementia: 33 (58%) stable MCI in the ADNI cohort 18 (50%) stable MCI in the TOMC cohort.</p> <p>Reference standards: NINCDS-ADRDA criteria (McKhann 1984). Baseline biomarker results were at the clinician's disposal; anyway progression to AD was ascertained basing on clinical criteria.</p> <p>Mean clinical follow-up: 3 ± 1 years for the ADNI cohort, 2.2 ± 1 years for the TOMC cohort</p>
Flow and timing	<p>Withdrawals explained and losses to follow-up: Missing data for 3 stable MCI and 1 converted MCI patient because of failed the automated processing.</p> <p>Uninterpretable MRI results have not been reported</p>
Comparative	
Key conclusions by the authors	Current findings suggest that β -amyloid concentrations in CSF and hippocampal volumes may be used in combination to best identify podromal AD.
Conflict of interests	information not available
Notes	<p>Source of funding: ADNI National Institutes of Health Grant U01 AG024904, Progetto Finalizzato Strategico 2006 and 2007, Ricerca Corrente Italian Ministry of Health, grant from the Associazione Fatebenefratelli per la Ricerca</p> <p>Information from authors: for the selection of patients the same criteria were used for MCI converters and non converters</p> <p>2 x 2 table: data from the published article; in order to avoid duplicate, only the results of the TOMC cohort were used for the review purpose</p>

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		

Prestia 2013 (ADNI) *(Continued)*

Did the study avoid inappropriate exclusions? No

High
Low
DOMAIN 2: Index Test All tests

Were the index test results interpreted without knowledge of the results of the reference standard? Unclear

Did the study provide a clear pre-specified definition of what was considered to be a "positive" result of the index test? Yes

Was the index test performed by a single operator or interpreted by consensus in a joint session? Yes

Unclear
Low
DOMAIN 3: Reference Standard

Is the reference standards likely to correctly classify the target condition? Yes

Were the reference standard results interpreted without knowledge of the results of the index tests? Unclear

Unclear
Low
DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard? Yes

Did all patients receive the same Yes

Prestia 2013 (ADNI) *(Continued)*

reference standard?

Were all patients included in the analysis? No

High
Prieto del Val 2016
Study characteristics

Patient sampling **Primary objectives:** investigate to what extent oscillatory EEG changes during memory encoding and/or retrieval enhance the accuracy of MTA in predicting conversion from amnesic MCI to AD

Study population: participants were recruited from senior citizens' associations, normal community health screening and hospital outpatient services.

Selection criteria: inclusion and exclusion criteria:

1. aged 50-80 years old
2. MMSE scores between 24 and 30
3. subjective memory complaints corroborated by the informant
4. objective memory loss confirmed by the Spanish version of the Logical Memory subtest extracted from the WMS-3rd Edition (scorings 1.5 SD below the age-appropriate mean)
5. absence of impairment in cognitive areas other than memory as revealed by neuropsychological testing
6. CDR global score of 0.5 (QD)
7. normal independence function
8. no criteria for dementia
9. no presence of neurological or psychiatric illness

Those participants with periventricular and/or deep white matter lesions in MRI as revealed by scores ≥ 2 on the Fazekas scale were excluded from the study. Individuals with a history of stroke and/or significant cerebrovascular conditions, clinically significant sensory impairment, past or current alcohol abuse, or those consuming medication known to affect memory, were not allowed to participate. None of the participants were taking cholinesterase inhibitors, and/or psychiatric medication at the time of recruiting or during the study.

Participants were excluded from the analysis if any psychiatric or neurological illness other than AD was present, and if participants presented with a systemic illness or signs of organ failure.

Included patients met core clinical criteria for MCI due to AD with an intermediate level of certainty ([Albert 2011](#))

Study design: prospective study

Patient characteristics and setting

Clinical presentation: amnesic MCI

Age mean (SD): MCI who progressed to AD: 69.7 ± 6.5 years; MCI non-converters to AD: 68.4 ± 7.1 years

Gender (% men): MCI who progressed to AD: 44%; MCI non-converters to AD: 28%

Education years mean (SD): MCI who progressed to AD: 7.7 ± 6.3 years; MCI non-converters to AD: 7.3 ± 5.1 years

ApoE4 carriers (%): not reported

Neuropsychological tests: employed; MMSE mean (SD): MCI who progressed to AD: 26.4 ± 2.7 ; MCI non-converters to AD: 26.8 ± 2.1

Prieto del Val 2016 (Continued)

Clinical stroke excluded: yes
Co-morbidities: not reported
Number enrolled: 34
Number available for analysis: 34

Setting: Pablo de Olavide University

Country: Spain
Period of study: not specified

Index tests	<p>Index test: MRI automated method for estimation of hippocampal and amygdala volume</p> <p>Manufacturer: Philips</p> <p>Tesla strength: 1.5 T</p> <p>Assessment methods: MRI data were preprocessed using Freesurfer v5.3. Removal of non-brain tissues was manually performed to increase the accuracy of segmentation. Volumetric measures were obtained for left and right sides of the hippocampus and amygdala</p> <p>Description of positive cases definition by index test as reported: not specified. The main outcome measure was the averaged AUC; additionally averaged overall accuracy, sensitivity and specificity based on the cut-off value maximised with the Youden index was computed (Youden 1950)</p> <p>Examiners: no details about radiologist</p> <p>Interobserver variability: not reported</p>		
Target condition and reference standard(s)	<p>Target condition: AD</p> <p>Prevalence of AD in the sample: 16/34 (47% of cases included in the analysis)</p> <p>Stable MCI or converted to other dementia: 18/34 (53%); 18 stable MCI</p> <p>Reference standard: NINCDS-ADRDA criteria and DSM IV criteria. AD participants had to further present MMSE scores ranging from 12-28 and a CDR global score of 1 (mild dementia).</p> <p>Mean clinical follow-up: 2 years</p>		
Flow and timing	<p>Withdrawals and losses to follow-up: none reported</p> <p>Uninterpretable MRI results have not been reported</p>		
Comparative			
Key conclusions by the authors	The results support the idea that synaptic integrity/function in the posterior cingulate cortex is affected during prodromal AD and has the potential of improving early detection when combined with MRI biomarkers.		
Conflict of interests	The study authors declare no competing financial interests		
Notes	<p>Source of funding: the work was supported by research grants from the Spanish Ministry of Economy and Competitiveness (PSI2014-55747-R, SAF2011-25463); the Regional Ministry of Innovation, Science and Enterprise, Junta de Andalucía (P12-CTS-2327); and CIBERNED (CB06/05/1111)</p> <p>2 x 2 table: data from the published article. In order to avoid duplication, we only used the results of amygdala for this review.</p>		
Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			

Prieto del Val 2016 (Continued)

Was a consecutive or random sample of patients enrolled? No

Was a case-control design avoided? Yes

Did the study avoid inappropriate exclusions? No

High
Low
DOMAIN 2: Index Test All tests

Were the index test results interpreted without knowledge of the results of the reference standard? Unclear

Did the study provide a clear pre-specified definition of what was considered to be a "positive" result of the index test? No

Was the index test performed by a single operator or interpreted by consensus in a joint session? Unclear

High
Low
DOMAIN 3: Reference Standard

Is the reference standards likely to correctly classify the target condition? Yes

Were the reference standard results interpreted without knowledge of the results of the index tests? Yes

Low
Low
DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard? Yes

Prieto del Val 2016 (Continued)

Did all patients receive the same reference standard? Yes

Were all patients included in the analysis? Yes

Low

Rhodus-Meester 2016

Study characteristics

Patient sampling	<p>Primary objectives: evaluate how a clinical decision support systems such as the PredictAD tool can aid clinicians to integrate biomarker evidence to support AD diagnosis</p> <p>Study population: patients with MCI from the Amsterdam Dementia Cohort, who had visited the Alzheimer center at the VU University Medical Center (VUMC) between 2000 and 2012.</p> <p>Selection criteria: MCI were included if a MMSE score was present, if both MRI and CSF biomarkers were available, if a follow-up of at least 2 years was conducted. Exclusion criteria not reported in the published article. According to the Amsterdam Dementia Cohort protocol, all MCI were assessed in order to identify a potential neurodegenerative disease. In Amsterdam Dementia Cohort vascular contribution to dementia conversion was considered (Van der Flier 2014). 23 people with MCI that progressed to another dementia were excluded from the study, 40 cases without MRI were excluded from the study</p> <p>Study design: prospective longitudinal study (VUMC cohort)</p>
Patient characteristics and setting	<p>Clinical presentations: MCI was diagnosed using Petersen's criteria (Petersen 2004); in addition all participants fulfilled the core clinical criteria of the NIA-AA for MCI (Albert 2011)</p> <p>Age mean (SD): MCI who progressed to AD: 72 ± 7; stable MCI: 68 ± 6</p> <p>Gender (% men): MCI who progressed to AD: 45%; stable MCI: 69%</p> <p>Education years mean (SD): MCI who progressed to AD: 5 ± 1; stable MCI: 5 ± 1</p> <p>ApoE4 carriers (%): not stated</p> <p>Neuropsychological tests: employed; MMSE mean (SD): MCI who progressed to AD: 26 ± 3; stable MCI 27 ± 2</p> <p>Clinical stroke excluded: unclear: infarcts were permitted but it is not specified if clinical or radiological. On FLAIR MRI white matter hyperintensities were rated using Fazekas scale, lacunes were counted and defined as deep lesions with low signal on T1-weighted sequences and high signal on T2-weighted sequences. Microbleeds were counted on T2 star sequences.</p> <p>Co-morbidities: not specified</p> <p>Number enrolled: 171</p> <p>Number available for analysis: 171</p> <p>Setting: Amsterdam Dementia Cohort from VUMC</p> <p>Country: Netherlands</p> <p>Period: 2000-2012</p>

Rhodius-Meester 2016 (Continued)

Language: English

Index tests	<p>Index test: MRI visual method for estimation of MTA</p> <p>Manufacturer: Siemens Magnetom Impact and Sonata, GE Healthcare Signa HDXT</p> <p>Tesla strength: 1.0 or 1.5</p> <p>Assessment methods: all scans were visually rated. Visual rating of MTA was performed on coronal T1-weighted images according to Scheltens 1997. The PredictAD tool was also performed to judge the combined biomarkers as indicative of AD pathophysiology.</p> <p>Description of positive cases definition by index test as reported: MTA averaged left and right score ≥ 1.5 was considered pathologic (Van de Pol 2014).</p> <p>Examiners: a trained rater evaluated all scans. Images were evaluated again in a consensus meeting with an experienced neuroradiologist.</p> <p>Interobserver variability: inter- and intra-rater weighted kappa's of at least 0.80 for MTA was required.</p>
Target condition and reference standard(s)	<p>Target condition: AD</p> <p>Prevalence of AD in the sample: 104/171 (61% of participants included in the analysis)</p> <p>Stable MCI or converted to other dementia: 67 (39%) stable MCI</p> <p>Reference standards: NINCDS-ADRDA criteria (McKhann 1984) and NIA-AA core clinical criteria (McKhann 2011)</p> <p>Median clinical follow-up: 3 years</p>
Flow and timing	<p>Withdrawals explained and losses to follow-up: none</p> <p>Uninterpretable MRI results have not been reported</p>
Comparative	
Key conclusions by the authors	<p>The ability of the PredictAD tool to identify AD pathophysiology was comparable to individual biomarkers. The PredictAD tool has the advantage that assigns likelihood to all participants, regardless of missing or conflicting data, allowing clinicians to integrate biomarker data in daily practice</p>
Conflict of interests	<p>Study authors' disclosures available online (www.j-alz.com/manuscript-disclosures/15-0548r1).</p>
Notes	<p>Source of funding: VUMC Alzheimer centre is supported by Alzheimer Nederland and Stichting VUMC fonds. The clinical database structure was developed with funding from Stichting Dioraphte. Other grants for the project: grant no. 733050201, grant agreement 611005, grant agreements 601055 (VPH-DARE@IT) and 224328 (PredictAD)</p> <p>2 x 2 table: data from the published article; we only considered the MTA score for this review.</p>

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		

Rhodijs-Meester 2016 (Continued)

Was a case-control design avoided? Yes

Did the study avoid inappropriate exclusions? No

High
Low
DOMAIN 2: Index Test All tests

Were the index test results interpreted without knowledge of the results of the reference standard? Unclear

Did the study provide a clear pre-specified definition of what was considered to be a "positive" result of the index test? Yes

Was the index test performed by a single operator or interpreted by consensus in a joint session? Yes

Unclear
Low
DOMAIN 3: Reference Standard

Is the reference standards likely to correctly classify the target condition? Yes

Were the reference standard results interpreted without knowledge of the results of the index tests? Yes

Low
Low
DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard? Yes

Did all patients receive the same reference standard? Yes

Rhodijs-Meester 2016 (Continued)

Were all patients included in the analysis? Yes

Low

VanderFlier 2005

Study characteristics

Patient sampling

Primary objectives: to investigate whether MRI-based volumes of whole brain, medial temporal lobe and white matter hyperintensities predict progression of cognitive decline in a sample of elderly people without dementia

Study population: elderly people without dementia who were consecutively referred to the memory clinic for cognitive complaints (MCI and subjective memory complaints)

Selection criteria: MCI

1. memory complaint
2. normal ADL
3. normal general cognitive function
4. abnormal memory for age
5. not demented.

MCI was diagnosed when both the interview and neuropsychological tests results gave evidence of memory impairment in the absence of general cognitive decline.

Exclusion criteria: age < 60 years, baseline diagnosis of dementia, other neurologic or psychiatric comorbidity and abnormalities on MRI except white matter hyperintensities or an incidental small lacunar lesion (5 mm diameter)

Study design: prospective longitudinal study (VUMC cohort)

Patient characteristics and setting

Clinical presentations: amnesic MCI according to Petersen 1995 and 1999 ([Petersen 1999](#))

Age mean (SD): 75 ± 7 years

Gender (% men): 29%

Education years mean (SD): 10 ± 3

ApoE4 carriers (%): not stated

Neuropsychological tests: employed; MMSE mean (SD): 26.0 ± 2

Clinical stroke excluded: not specified

Co-morbidities: not reported

Number enrolled: 17

Number available for analysis: 15

Setting: VUMC cohort

Country: Netherlands

Period: not reported

Language: English

VanderFlier 2005 (Continued)

Index tests	<p>Index test: MRI manual method for estimation of hippocampus, medial temporal lobe volume; MRI semiautomated method for estimation of whole brain</p> <p>Manufacturer: Philips Medical Systems, Best, the Netherlands</p> <p>Tesla strength: 1.5</p> <p>Assessment methods: manual segmentation of the medial temporal lobe volume (including hippocampus and parahippocampal gyrus) was performed using the software DISPLAY (Pruessner 2002). The whole brain volume was segmented using in-house-developed semiautomated software (Division of Image Processing (LKEB), Department of Radiology, Leiden University Medical Center)</p> <p>Description of positive cases definition by index test as reported: not specified</p> <p>Examiners: all measurements were performed by a single rater who was blind to participant identity and diagnosis</p> <p>Interobserver variability: ICC for medial temporal lobe volume was 0.91. Interrater reliability of whole brain volume was 1.0</p>		
Target condition and reference standard(s)	<p>Target condition: AD</p> <p>Prevalence of AD in the sample: 9/15 (60% of participants included in the analysis)</p> <p>Stable MCI or converted to other dementia: 6 (40%) non converter to AD dementia</p> <p>Reference standards: NINCDS-ADRDA criteria (McKhann 1984). Diagnosis was made in a multidisciplinary consensus meeting</p> <p>Mean clinical follow-up: 1.8 years</p>		
Flow and timing	<p>Withdrawals explained and losses to follow-up: 2/17 participants (12%) were lost to follow-up, reasons unclear (reasons for losses in MCI group were reported together with those of the subjective memory complaint group: 1 participant died, 1 could not be traced, 3 refused to participate, 1 refused to participate in the follow-up study)</p> <p>Uninterpretable MRI results have not been reported</p>		
Comparative			
Key conclusions by the authors	<p>Results suggest that regional damage to the medial temporal lobes underlies initial mild cognitive impairment, whereas more global brain changes, such as whole brain atrophy and WMH [white matter hyperintensities], contribute to further progression of cognitive decline</p>		
Conflict of interests	Information not available		
Notes	<p>Source of funding: contract/grant sponsor: Internationale Stichting Alzheimer Onderzoek (ISAO).</p> <p>2 x 2 table: data to complete 2 x 2 table provided by the study authors</p>		
Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		

VanderFlier 2005 *(Continued)*

Was a case-control design avoided? Yes

Did the study avoid inappropriate exclusions? Yes

Low
Low
DOMAIN 2: Index Test All tests

Were the index test results interpreted without knowledge of the results of the reference standard? Yes

Did the study provide a clear pre-specified definition of what was considered to be a "positive" result of the index test? No

Was the index test performed by a single operator or interpreted by consensus in a joint session? Yes

High
Low
DOMAIN 3: Reference Standard

Is the reference standards likely to correctly classify the target condition? Yes

Were the reference standard results interpreted without knowledge of the results of the index tests? Yes

Low
Low
DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard? Yes

Did all patients receive the same reference standard? Yes

Were all patients included in the analysis? No

High

Visser 1999
Study characteristics

Patient sampling	<p>Primary objectives: to determine whether the medial temporal lobe is atrophic in participants with MCI, and whether atrophy of this structure is a better predictor of dementia than memory dysfunction</p> <p>Study population: participants affected by minimal dementia in the Amsterdam Study of the Elderly (AMSTEL)</p> <p>Selection criteria: participants from AMSTEL: for the brain-imaging substudy 73 participants with minimal dementia were randomly selected, 33 were asked to participate and 28 agreed. 8 cases were not included because of missing baseline MRI. Exclusion criteria: not reported</p> <p>Study design: prospective longitudinal study</p>
Patient characteristics and setting	<p>Clinical presentations: minimal dementia according to CAMDEX criteria (CAMDEX: Cambridge Examination of Mental Disorders of the Elderly). The diagnosis of minimal dementia was made when the DSM-III-R criteria of dementia were not met, but based on an overall clinical impression, there was limited and variable impairment in cognitive and social functioning such as difficulty with learning and recalling events, a tendency to misplace possessions, and minor errors in orientation. Similar entities are QD or a score of 0.5 on the CDR scale, and “mild cognitive impairment” or a score of 3 on the global deterioration scale (Reisberg 1982)</p> <p>Age mean (SD): minimal dementia who progressed to AD: 79 ± 4; stable minimal dementia: 78 ± 7</p> <p>Gender (% men): minimal dementia who progressed to AD: 33%; stable minimal dementia: 0%</p> <p>Education years mean (SD): minimal dementia who progressed to AD: 7.1 ± 2.3; stable minimal dementia: 8 ± 2.3</p> <p>ApoE4 carriers (%): not stated</p> <p>Neuropsychological tests: employed; MMSE mean (SD): minimal dementia who progressed to AD: 23.1 ± 1.2; stable minimal dementia: 21 ± 1.8</p> <p>Clinical stroke excluded: not specified</p> <p>Co-morbidities: not reported</p> <p>Number enrolled: 20</p> <p>Number available for analysis: 13</p> <p>Setting: population-based study of mental functioning in non-institutionalised persons (AMSTEL cohort: minimal dementia group)</p> <p>Country: Netherlands</p> <p>Period: not reported</p> <p>Language: English</p>
Index tests	<p>Index test: MRI manual method for estimation of hippocampus, parahippocampal gyrus and lateral temporal lobe; MTA score was also available</p> <p>Manufacturer: Teslacon II (Technicare, Solon, Ohio)</p> <p>Tesla strength: 0.6 Tesla</p> <p>Assessment methods: volumetry was carried out on a SUN workstation with software developed in-house. The hippocampus, parahippocampal gyrus, and intracranial area, as a measure of the intracranial volume, were outlined by hand. A seed function was used for the temporal lobe. The volumes of the hippocampus and parahippocampal gyrus were subtracted from the volume of the total temporal lobe to give the volume of the lateral temporal lobe</p>

Visser 1999 (Continued)

Description of positive cases definition by index test as reported: not specified

Examiners: all measurements were carried out by 1 rater who was blinded to the participants' age, diagnosis, and sex

Interobserver variability: not reported. The average difference between the 1st and 2nd measurement of the brain structures on 10 scans was $-0.08 \pm 0.29 \text{ cm}^3$ for the parahippocampal gyrus, $-0.07 \pm 0.20 \text{ cm}^3$ for the hippocampus, $0.14 \pm 1.1 \text{ cm}^3$ for the lateral temporal lobe, and $-1.9 \pm 1.4 \text{ cm}^3$ for the intracranial area

Target condition and reference standard(s)

Target condition: AD

Prevalence of AD in the sample: 9/13 (69% of participants included in the analysis)

Stable minimal dementia or converted to other dementia: 4 (31%) stable MCI

Reference standards: NINCDS-ADRDA criteria (McKhann 1984).

Mean clinical follow-up: 3 years

Flow and timing

Withdrawals explained and losses to follow-up: 7/20 participants with minimal dementia (35%): participants dropped out before the 1st assessment

Uninterpretable MRI results not reported

Comparative

Key conclusions by the authors

Severe MTA was present in some participants who had MCI at baseline and subsequently developed dementia due to AD. The absence of MTA, however, does not exclude the development of dementia. Memory impairment was a better predictor for dementia than atrophy of the medial temporal lobe, but the combination of the two increased diagnostic accuracy

Conflict of interests

Information not available

Notes

Source of funding: not reported

2 x 2 table: data to complete 2 x 2 table provided by the study authors

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

Unclear

Low

DOMAIN 2: Index Test All tests

Were the index test results interpreted with-

Yes

Visser 1999 (Continued)
 out knowledge of the
 results of the refer-
 ence standard?

Did the study provide
 a clear pre-specified
 definition of what was
 considered to be a
 "positive" result of the
 index test? No

Was the index test per-
 formed by a single op-
 erator or interpreted
 by consensus in a joint
 session? Yes

High

Low

DOMAIN 3: Reference Standard

Is the reference stan-
 dards likely to correct-
 ly classify the target
 condition? Yes

Were the reference
 standard results inter-
 preted without knowl-
 edge of the results of
 the index tests? Yes

Low

Low

DOMAIN 4: Flow and Timing

Was there an appro-
 priate interval be-
 tween index test and
 reference standard? Yes

Did all patients receive
 the same reference
 standard? Yes

Were all patients in-
 cluded in the analysis? No

High

Visser 2002

Study characteristics

Patient sampling **Primary objectives:** to investigate whether MTA predicted outcome in patients with MCI and whether assessment of the medial temporal lobe could increase the diagnostic accuracy of age and delayed recall for outcome

Visser 2002 (Continued)

Study population: patients with MCI were selected from the Maastricht Memory Clinic, a university affiliated outpatient clinic for patients with cognitive impairments. Patients were referred to the clinic by a general practitioner, a neurologist, or a psychiatrist

Selection criteria: inclusion criteria were a score on the Global Deterioration Scale of 2 or 3. Exclusion criteria were age < 50 years, a baseline diagnosis of dementia according to the DSM-IV criteria, sensory impairment, psychosis, panic disorder, bipolar disorder, a score on the Hamilton depression rating scale-17 items (HDRS) > 22, or cognitive problems in relation to cerebrovascular events, neurodegenerative diseases (for example, Parkinson's disease or Huntington's disease), brain neoplasm, head trauma, drug intoxication, alcohol misuse, hypothyroid or hyperthyroid function, or vitamin deficiency

Study design: prospective longitudinal study

Patient characteristics and setting

Clinical presentations: MCI definition based on Global Deterioration Score according to [Reisberg 1982](#)

Age mean (SD): 65 ± 9.5 years

Gender (% men): 58 %

Education years mean (SD): 10.7 ± 3.2

ApoE4 carriers (%): not stated

Neuropsychological tests: employed; MMSE mean (SD): 27.7 ± 1.8

Clinical stroke excluded: yes (conversely vascular disorders were permitted if not related to the onset of cognitive impairment)

Co-morbidities: MCI had vascular risk factors or vascular disorders such as hypertension (diastolic blood pressure > 95, systolic blood pressure > 170 on a single measurement, or treatment for hypertension) (N = 8), total cholesterol serum concentrations > 6.0 mmol/L (N = 5), smoking (N = 6), angina pectoris (N = 1), transient ischaemic attack (N = 2), and lacunar infarction on MRI (N = 2)

Number enrolled: 31

Number available for analysis: 30 for the MTA estimation, 29 for the hippocampal volume estimation

Setting: : Maastricht Memory Clinic

Country: Netherlands

Period: not reported

Language: English

Index tests

Index test: MRI manual method for estimation of hippocampal and parahippocampal gyrus volumes, MRI visual method for estimation of MTA

Manufacturer: Gyroscan ACS-II, Philips

Tesla strength: 1.5

Assessment methods: data were transferred to a SUN workstation and the ROI were measured with ShowImage (developed at the Department of Clinical Physics and Informatics, Vrije Universiteit, Amsterdam, The Netherlands). The brain structures were manually traced with a mouse-driven cursor. Measurements were done with reference to an anatomical atlas ([Duvernoy 1988](#)). Visual method according to [Scheltens 1992](#).

Description of positive cases definition by index test as reported: volumes and MTA rating were classified on the basis of the zeta score in tertiles; tertiles are based on a reference population of healthy participants. A low tertile score was indicative of a small volume. A cut-off was not specified.

Examiners: the manual method for estimation of the hippocampal and parahippocampal gyrus volumes was performed by 1 rater. The MTA was scored by a neurologist. All raters were blinded to all clinical information.

Visser 2002 (Continued)

Interobserver variability: regarding the manual method the ICC between the first and second measurement was 0.95 for the hippocampus, 0.92 for the parahippocampal gyrus; regarding the visual method, the intrarater reliability was substantial ($\kappa = 0.70$)

Target condition and reference standard(s)

Target condition: AD

Prevalence of AD in the sample: 7/30 (23% of participants included in the analysis)

Stable MCI or converted to other dementia: 23 (77%) non-AD dementia (20 stable MCI and 3 participants converted to non-AD dementia)

Reference standards: NINCDS-ADRDA criteria (McKhann 1984)

The diagnosis at follow-up was made by an experienced neuropsychiatrist who was unaware of the results of the baseline assessment including the MRI data.

Mean clinical follow-up: 1.9 years

Flow and timing

Withdrawals explained and losses to follow-up: 1/31 cases (3%) for missing information on the presence or absence of AD. 1 lost for the manual method

Uninterpretable MRI results have not been reported

Comparative

Key conclusions by the authors

The ability to detect patients at high risk for AD among those with MCI increases when data on age and memory function are combined with measures of MTA. Volumetry of the hippocampus is preferred, but qualitative rating of MTA is a good alternative.

Conflict of interests

Information not available

Notes

Source of funding: not reported

The MRI scan of 1 participant was not available for volumetry and only the qualitative rating was performed for this scan.

2 x 2 table: data available from study authors

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
------	--------------------	--------------	------------------------

DOMAIN 1: Patient Selection

Was a consecutive or random sample of patients enrolled?	No		
--	----	--	--

Was a case-control design avoided?	Yes		
------------------------------------	-----	--	--

Did the study avoid inappropriate exclusions?	No		
---	----	--	--

High

Low

DOMAIN 2: Index Test All tests

Visser 2002 (Continued)

Were the index test results interpreted without knowledge of the results of the reference standard?

Yes

Did the study provide a clear pre-specified definition of what was considered to be a "positive" result of the index test?

No

Was the index test performed by a single operator or interpreted by consensus in a joint session?

Yes

High
Low
DOMAIN 3: Reference Standard

Is the reference standards likely to correctly classify the target condition?

Yes

Were the reference standard results interpreted without knowledge of the results of the index tests?

Yes

Low
Low
DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard?

Yes

Did all patients receive the same reference standard?

Yes

Were all patients included in the analysis?

No

High

Wang 2006

Study characteristics

Patient sampling

Primary objectives: to evaluate the correlations of ApoE genotype, cognitive performance, medial temporal structure volumes, and clinical outcome in amnesic MCI

Study population: MCI consecutively recruited from neurological clinics at Taipei Veterans General Hospital

Selection criteria: inclusion criteria:

1. memory complaint, preferably corroborated by an informant
2. objective memory impairment (scores of free delayed recall on Selective Reminding Test or Complex Figure Text > 1 SD less than the data from the healthy elderly control group)
3. normal general cognitive function
4. intact ADL
5. did not meet dementia criteria.

Each MCI patient had a CDR score of 0.5. Exclusion criteria: evidence of other neurological, psychiatric or systemic conditions that can cause cognitive impairment (e.g. stroke, alcoholism, major depression). Individuals with evidence of structural brain alterations, such as masses, cortical stroke, multiple subcortical lacunae or prominent periventricular white matter changes, were also excluded

Study design: prospective longitudinal study

Patient characteristics and setting

Clinical presentations: amnesic MCI according to Petersen's criteria (Petersen 1999).

Age mean (SD): MCI who progressed to AD: 78 ± 5; stable MCI: 76 ± 4

Gender (% men): MCI who progressed to AD: 63%; stable MCI: 79%

Education years mean (SD): MCI who progressed to AD: 11.7 ± 5.7; stable MCI: 11.7 ± 3.3

ApoE4 carriers (%): MCI who progressed to AD: 33.3 % ; stable MCI: 22.9 %

Neuropsychological tests: employed; MMSE mean (SD): MCI who progressed to AD: 24.4 ± 2.1; stable MCI: 26.6 ± 2.6

Clinical stroke excluded: yes

Co-morbidities: not reported

Number enrolled: 58

Number available for analysis: 58

Setting: Taipei Veterans General Hospital (Memory Clinic)

Country: Taiwan

Period: August 1999-July 2003

Language: English

Index tests

Index test: MRI manual method for estimation of hippocampus and amygdala volumes

Manufacturer: Magnetom Vison, Siemens

Tesla strength: 1.5

Assessment methods: anatomical boundaries of the hippocampus and amygdala adopted were those defined by Lehericy (Lehericy 1994).

Wang 2006 (Continued)

Description of positive cases definition by index test as reported: not specified

Examiners: a single neuroradiologist performed the assessment, blinded to clinical and neuropsychological data.

ICC for intra-rater (test–retest) agreement from 12 images were 0.98 for the hippocampus and the 0.97 for the amygdala.

Interobserver variability: not provided

Target condition and reference standard(s)

Target condition: AD

Prevalence of AD in the sample: 19/58 (33% of enrolled participants)

Stable MCI or converted to other dementia: 39 (67%) stable MCI

Reference standards: NINCDS-ADRDA criteria (Mckhann 1984)

Neurologists blind to the neuroimaging volumetric measurement results made clinical diagnoses at baseline. It is unclear if the blindness was respected in the follow-up.

Mean clinical follow-up: 1.8 years

Flow and timing

Withdrawals explained and losses to follow-up: none reported

Uninterpretable MRI results have not been reported

Comparative

Key conclusions by the authors

Both cognitive performance and hippocampal volume predicted decline of MCI patients. However, when time-to-conversion was considered as a principal outcome factor, global cognitive performance had greater significant influence on conversion time than hippocampal volume and ApoEε4 was not a significant predictive factor for dementia due to AD

Conflict of interests

Information not available

Notes

Source of funding: this study was partially supported by research grants from the National Science Council (NSC91-2314-B-010-014, NSC 92-2314-B-010-031) and Taipei Veterans General Hospital (V310)

2 x 2 table: data to complete 2 x 2 table provided by the study authors

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?	No		
---	----	--	--

High

Low

DOMAIN 2: Index Test All tests

Wang 2006 (Continued)

Were the index test results interpreted without knowledge of the results of the reference standard? Yes

Did the study provide a clear pre-specified definition of what was considered to be a "positive" result of the index test? No

Was the index test performed by a single operator or interpreted by consensus in a joint session? Yes

High
Low
DOMAIN 3: Reference Standard

Is the reference standards likely to correctly classify the target condition? Yes

Were the reference standard results interpreted without knowledge of the results of the index tests? Yes

Low
Low
DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard? Yes

Did all patients receive the same reference standard? Yes

Were all patients included in the analysis? Yes

Low
Westman 2011
Study characteristics

Patient sampling **Primary objectives:** aims of this study were:

Westman 2011 (Continued)

1. to compare the ability of MTA visual assessment rating scales, a multivariate MRI classification method and manually measured hippocampal volumes to distinguish between participants with AD and healthy elderly controls
2. to assess how well the 3 techniques perform when predicting future conversion from MCI to AD

Study population: participants with MCI originated from AddNeuroMed project, recruited from local memory clinics of the 6 participating sites across Europe

Selection criteria: inclusion criteria:

1. MMSE score range 24-30
2. GDS score ≤ 5
3. age ≥ 65 years
4. medication stable
5. good general health.

Exclusion criteria:

1. meet the DSM- IV criteria for dementia
2. significant neurological or psychiatric illness other than AD
3. significant unstable systematic illness or organ failure.
4. scores 0.5 on CDR scale

Study design: prospective longitudinal study

Patient characteristics and setting

Clinical presentations: participants with MCI; it was preferable that the participant and informant report occurrence of memory problems

Age mean (SD): 74.0 \pm 6

Gender (% men): 49%

Education years mean (SD): 8.7 \pm 4.3

ApoE4 carriers (%): not reported

Neuropsychological tests: employed; MMSE mean (SD): 27.2 \pm 1.6

Clinical stroke excluded: not specified

Co-morbidities: not reported

Number enrolled: 101

Number available for analysis: 101

Setting: AddNeuroMed cohort

Country: Finland, Italy, Greece, UK, Poland, France (AddNeuroMed)

Period: not reported

Language: English

Index tests

Index test: MRI manual method for estimation of hippocampal volume and MRI visual method for estimation of medial temporal lobe volume. MRI automated method was also used to generate regional volume and cortical thickness measures (57 variables).

Manufacturer: 6 different MR systems: 4 GE, 1 Siemens and 1 Picker (information retrieved from AddNeuroMed protocol (Simmons 2009))

Tesla strength: 1.5 (Simmons 2009)

Westman 2011 (Continued)

Assessment methods: data acquisition for the AddNeuroMed study was designed to be compatible with the ADNI. detailed quality control carried out on all MIs according to the AddNeuroMed quality control procedure. Manual measurements were performed on a HERMES workstation. The visual rating assessment was performed according to [Scheltens 1992](#). Regarding the automated method, it utilised a pipeline developed by Fischl and Dale that produces regional cortical thickness (31 areas) and volumetric (23 areas) measures. The 57 variables were used for a multivariate analysis.

Description of positive cases definition by index test as reported: positive cases were defined only for the visual method (age-dependent cut-off)

Examiners: for the manual and visual methods, a single rater blinded to diagnosis performed the corresponding assessment.

Interobserver variability: for the manual method the ICC of the measurements were 0.93; for the visual method, the intra-rater reliability was 0.81 on right side and on left side 0.78. Weighted kappa was 0.93 on both sides

Target condition and reference standard(s)	<p>Target condition: AD</p> <p>Prevalence of AD in the sample: 19/101 (19% of enrolled participants)</p> <p>Stable MCI or converted to other dementia: 82 (81%) non-AD converters</p> <p>Reference standards: NINCDS-ADRDA criteria (McKhann 1984); not specified in the published article; this information is reported in Liu 2010, another AddNeuroMed study</p> <p>Mean clinical follow-up: 1 year</p>
Flow and timing	<p>Withdrawals explained and losses to follow-up: none reported</p> <p>Uninterpretable MRI results have not been reported</p>
Comparative	
Key conclusions by the authors	<p>Visual rating assessment of the medial temporal lobe gave similar prediction accuracy to multivariate classification and manual hippocampal volumes. This suggests a potential future role for computerised methods as a complement to clinical assessment of AD.</p>
Conflict of interests	<p>The study was supported by InnoMed (Innovative Medicines in Europe). No patents, products in development or marketed products to declare</p>
Notes	<p>Source of funding: InnoMed (FP6-2004-LIFESCIHEALTH-5)</p> <p>2 x 2 table: data from the published article; in order to avoid duplicate, only the results of the manual method were used for the review purpose. The automated method was excluded because provided a mixed index test (volume and thickness)</p>

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		

Westman 2011 (Continued)

Did the study avoid inappropriate exclusions? Yes

High
Low
DOMAIN 2: Index Test All tests

Were the index test results interpreted without knowledge of the results of the reference standard? Yes

Did the study provide a clear pre-specified definition of what was considered to be a "positive" result of the index test? No

Was the index test performed by a single operator or interpreted by consensus in a joint session? Yes

High
Low
DOMAIN 3: Reference Standard

Is the reference standards likely to correctly classify the target condition? Yes

Were the reference standard results interpreted without knowledge of the results of the index tests? Yes

Low
Low
DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard? Yes

Did all patients receive the same reference standard? Yes

Were all patients included in the analysis? Yes

Westman 2011 (Continued)

Low

Wolz 2011
Study characteristics

Patient sampling	<p>Primary objectives: to assess the improvement in classification accuracy that can be achieved by combining features from different structural MRI analysis techniques (hippocampal volume, tensor-based morphometry, cortical thickness and a novel technique based on manifold learning).</p> <p>Study population: participants with MCI included in ADNI for which a 1.5 T T1-weighted MRI scan at baseline was available</p> <p>Selection criteria: MCI included participants had MMSE scores between 24–30 (inclusive), a memory complaint, objective memory loss measured by education-adjusted scores on WMS Logical Memory II, a CDR of 0.5, absence of significant levels of impairment in other cognitive domains, essentially preserved ADL, and an absence of dementia. Exclusion criteria not reported in the published article. Exclusion criteria as reported in the ADNI protocol ADNI 2010 and Gaser 2013</p> <p>Study design: prospective longitudinal study (ADNI study)</p>
Patient characteristics and setting	<p>Clinical presentations: amnesic MCI according to ADNI 2010 and Petersen 2010</p> <p>Age mean (SD): MCI who progressed to AD: 75 ± 7; stable MCI: 75 ± 8</p> <p>Gender (% men): MCI who progressed to AD: 62%; stable MCI: 66%</p> <p>Education years mean (SD): MCI who progressed to AD: 15.7 ± 2.9; stable MCI: 15.6 ± 3.1</p> <p>ApoE4 carriers (%): MCI who progressed to AD: 66%; stable MCI: 39%</p> <p>Neuropsychological tests: MMSE mean (SD): MCI who progressed to AD: 26.6 ± 1.7; stable MCI 27.3 ± 1.8</p> <p>Clinical stroke excluded: not specified</p> <p>Co-morbidities: not reported</p> <p>Number enrolled: 405</p> <p>Number available for analysis: 405</p> <p>Setting: participants of the ADNI study</p> <p>Country: USA and Canada</p> <p>Period: follow-up period was stopped in July 2011</p> <p>Language: English</p>
Index tests	<p>Index test: MRI automated method for estimation of hippocampal volume; also other methods described</p> <p>Manufacturer: GE Healthcare, Philips Medical System, Siemens Medical Solution (Jack 2008b)</p> <p>Tesla strength: 1.5</p> <p>Assessment methods: hippocampal volumes were measured using an approach based on fast and robust multi-atlas segmentation (Lötjönen 2011). Details about cortical thickness, tensor-based morphometry and minifold-based learning methods were also reported. 2 different methods were used to perform classification based on individual features and their combination: linear discriminant analysis and support vector machines. Total volume of left and right hippocampus were combined in a single feature.</p>

Wolz 2011 (Continued)

Description of positive cases definition by index test as reported: classifiers were built on a training set composed of healthy and AD and used to classify the MRI images of MCI participants as more similar to healthy (negative cases) and more similar to AD (positive cases) .

Examiners: imaging interpretation reserved to an automatic classifier

Interobserver variability: not provided

Target condition and reference standard(s)

Target condition: AD

Prevalence of AD in the sample: 167/405 (41% of enrolled participants)

Stable MCI or converted to other dementia: 238 (59%) stable MCI

Reference standards: NINCDS-ADRDA criteria ([McKhann 1984](#)).

Mean clinical follow-up: 1.5 years

Flow and timing

Withdrawals explained and losses to follow-up: none reported

Uninterpretable MRI results have not been reported

Comparative

Key conclusions by the authors

A comprehensive analysis of MRI images combining multiple features improves classification accuracy and predictive power in detecting early AD. The most stable and reliable classification was achieved when combining all available features.

Conflict of interests

The study authors have declared no competing interests

Notes

Source of funding: U01 AG024904

2 x 2 table: data from the published article; we only used data regarding volumetric results obtained with linear discriminant analysis in the review

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
------	--------------------	--------------	------------------------

DOMAIN 1: Patient Selection

Was a consecutive or random sample of patients enrolled?	No		
--	----	--	--

Was a case-control design avoided?	Yes		
------------------------------------	-----	--	--

Did the study avoid inappropriate exclusions?	No		
---	----	--	--

High

Low

DOMAIN 2: Index Test All tests

Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
---	---------	--	--

Wolz 2011 (Continued)

Did the study provide a clear pre-specified definition of what was considered to be a "positive" result of the index test?

Yes

Was the index test performed by a single operator or interpreted by consensus in a joint session?

Unclear

Unclear

Low

DOMAIN 3: Reference Standard

Is the reference standards likely to correctly classify the target condition?

Yes

Were the reference standard results interpreted without knowledge of the results of the index tests?

Yes

Low

Low

DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard?

Yes

Did all patients receive the same reference standard?

Yes

Were all patients included in the analysis?

Yes

Low

Wood 2016
Study characteristics

Patient sampling

Primary objectives: test the hypothesis that allocentric spatial memory is predictive of conversion from MCI to dementia due to AD

Study population: patients with MCI recruited from the Cognitive Disorders Clinic, Hurstwood Park Neurological Centre, Haywards Heath, West Sussex, and from East Sussex Memory Assessment Service

Selection criteria: initial screening blood tests were undertaken to exclude reversible causes of cognitive impairment, such as vitamin B12 deficiency and thyroid dysfunction. Participants were excluded

Wood 2016 (Continued)

ed from the study if they had depression, other psychiatric diagnoses, a significant vascular lesion load on MRI or a Hachinski ischaemic score > 4

Study design: prospective longitudinal study

Patient characteristics and setting

Clinical presentations: multiple domain MCI with memory impairment, MCI definition according to Petersen 2004

Age mean (SD): MCI who progressed to AD: 72 ± 3; MCI non-AD converters: 65 ± 3

Gender (% men): MCI who progressed to AD: 78%; MCI non-AD converters: 67%

Education years mean (SD): MCI who progressed to AD: 12.1 ± 0.7; MCI non-AD converters: 11.0 ± 0.5

ApoE4 carriers (%): not stated

Neuropsychological tests: employed; MMSE mean (SD): MCI who progressed to AD: 27.89 ± 0.42; MCI non-AD converters: 27.33 ± 0.21

Clinical stroke excluded: not specified

Co-morbidities: not reported

Number enrolled: 15

Number available for analysis: 15

Setting: memory clinics in West and East Sussex

Country: UK

Period: not reported

Language: English

Index tests

Index test: MRI automated method for estimation of hippocampal volume

Manufacturer: Siemens Avanto scanner

Tesla strength: 1.5

Assessment methods: total hippocampal volumes, corrected for total intracranial volume, were measured using the FSL (version 5.0)/FIRST tool (FMRIB, Oxford Centre for Functional Magnetic Resonance Imaging of the Brain, Oxford, UK)

Description of positive cases definition by index test as reported: the threshold criteria for binary classification (MCI converters vs non-converters) were selected by maximising Youden's index (Youden 1950)

Examiners: no details

Interobserver variability: not provided

Target condition and reference standard(s)

Target condition: AD

Prevalence of AD in the sample: 9/15 (60% of enrolled participants)

Stable MCI or converted to other dementia: 6 (40%) MCI non-AD converters

Reference standards: NIA/AA criteria (McKhaan 2011). It was not specified if only the core clinical criteria were used to do the diagnosis. Neurologist was blinded to the baseline 4 Mountains test score.

Mean clinical follow-up: 2 years

Flow and timing

Withdrawals explained and losses to follow-up: none reported

Wood 2016 (Continued)

Uninterpretable MRI results have not been reported

Comparative

Key conclusions by the authors The study provided the first evidence supporting the hypothesis that testing of allocentric spatial memory (the 4 Mountains test) was predictive of conversion from MCI to AD dementia

Conflict of interests The study authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest

Notes **Source of funding:** UK National Institute for Health Research (NIHR), Cambridge NIHR Biomedical Research Centre, BBSRC Research Grant (BB/M008975/1).

2 x 2 table: data from the published article

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
------	--------------------	--------------	------------------------

DOMAIN 1: Patient Selection

Was a consecutive or random sample of patients enrolled?	No		
--	----	--	--

Was a case-control design avoided?	Yes		
------------------------------------	-----	--	--

Did the study avoid inappropriate exclusions?	No		
---	----	--	--

High
Low
DOMAIN 2: Index Test All tests

Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
---	---------	--	--

Did the study provide a clear pre-specified definition of what was considered to be a "positive" result of the index test?	No		
--	----	--	--

Was the index test performed by a single operator or interpreted by consensus in a joint session?	Unclear		
---	---------	--	--

High
Low
DOMAIN 3: Reference Standard

Is the reference standards likely to correctly classify the target condition?	Yes		
---	-----	--	--

Wood 2016 (Continued)

Were the reference standard results interpreted without knowledge of the results of the index tests?

Unclear

Unclear

Unclear

DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard?

Yes

Did all patients receive the same reference standard?

Yes

Were all patients included in the analysis?

Yes

Low

AA: Alzheimer's Association; **AD:** Alzheimer's disease; **ADL:** activities of daily living; **ADNI:** Alzheimer's Disease Neuroimaging Initiative; **ADPR/ADRC:** Alzheimer's Disease Patient Registry/Alzheimer's Disease Research Center; **ADRDA:** Alzheimer's Disease and Related Disorders Association; **AUC:** area under the curve; **B-ADL:** Bayer Activities of Daily Living; **CDR:** Clinical Dementia Rating; **CSF:** cerebrospinal fluid; **CVRS:** comprehensive visual rating scale; **DSM:** Diagnostic and Statistical Manual of Mental Disorders; **EEG:** electroencephalogram; **FDG:** fluorodeoxyglucose; **FTD:** frontotemporal dementia; **GDS:** Geriatric Depression Scale; **GE:** General Electric Medical Systems; **HCV:** hepatitis C virus; **IADL:** instrumental activities of daily living; **ICC:** intraclass correlation coefficient; **ICV:** intracerebroventricular injection; **LBD:** Lewy body dementia; **MCI:** mild cognitive impairment; **MMSE:** Mini Mental State Examination; **MRC:** Medical Research Council; **MRI:** magnetic resonance imaging; **MTA:** medial temporal lobe atrophy; **NIA:** National Institute on Aging; **NIH:** National Institutes of Health; **NINCDS:** National Institute of Neurological and Communicative Disorders and Stroke; **PET:** positron emission tomography; **QD:** questionable dementia; **ROI:** region of interest; **SD:** standard deviation; **SPECT:** single-photon emission computerised tomography; **VD:** vascular dementia; **WMS:** Wechsler Memory Scale; **WMS-R:** Wechsler Memory Scale – Revised

Characteristics of excluded studies [ordered by study ID]

Study	Reason for exclusion
Aguilar 2014	Index test not in line with inclusion criteria (mixed test: brain cortical thickness and volumetric measures)
Aksu 2011	Index test outside inclusion criteria (MRI test obtained from multiple volumetric measures)
Apostolova 2006	Not a DTA study
Apostolova 2014	ADNI study that reported on a cohort that overlapped with a cohort in another included paper (Wolz 2011)
Archer 2010	Insufficient DTA information (unable to construct 2 x 2 tables); unable to contact study authors
Ardekani 2017	Index test not in line with inclusion criteria (MRI-derived index: hippocampal volumetric integrity)
Bakkour 2009	Insufficient DTA information (unable to construct 2 x 2 tables); unable to contact study authors
Barnes 2014	Insufficient DTA information (unable to construct 2 x 2 tables); study authors did not answer data request

Study	Reason for exclusion
Beheshti 2016	Study design outside inclusion criteria (case-control study)
Beheshti 2017	Index test outside inclusion criteria (test obtained from multiple volumetric measures)
Bell-McGinty 2005	Index test not in line with inclusion criteria (voxel-based-morphometry)
Bernard 2014	Participants outside inclusion criteria (individuals aged ≥ 65 years randomly recruited from population electoral lists)
Blasko 2008	Index test not in line with inclusion criteria (data reported for MRI combined with plasma amyloid beta 42 and homocysteine)
Bombois 2008	Insufficient DTA information (unable to construct 2 x 2 tables); study authors did not answer data request
Borgio 2012	No reference standard
Boutet 2012	Not a DTA study (focus on comparison between visual assessment of MTA with automatic hippocampal volumetry for the classification of AD, MCI and cognitively normal)
Bouwmann 2007	High risk for duplication (study design and MRI parameters similar to Rhodius-Meester 2016 , population from the same setting VUMC, with a smaller sample size)
Brickman 2015	Participants outside inclusion criteria (the Washington Heights Inwood Columbia Aging Project (WHICAP), an ongoing longitudinal study of cognitive aging and dementia that recruited "non demented participants")
Bron 2014	Index test outside inclusion criteria (test obtained from multiple volumetric measures))
Bron 2015	Index test outside inclusion criteria (MRI test obtained from multiple volumetric measures)
Brueggen 2015	Study design outside inclusion criteria, retrospective study (participants were selected through a retrospective review of clinical records)
Brys 2009	Index test outside inclusion criteria (voxel-based-morphometry)
Brück 2013	Index test was not a volumetric MRI test (PET)
Buckley 2017	Index test outside inclusion criteria
Buckner 2005	Participants outside inclusion criteria (healthy and AD participants)
Callahan 2015	Index test not in line with inclusion criteria (data reported for MRI combined with neuropsychological tests)
Cardenas 2003	Not a DTA study (focus on atrophy rate comparison between healthy participants and patients affected by cognitive impairment)
Carmichael 2013	The study reported outcomes for number of structural MRI throughout participants' follow-up - not number of participants
Caroli 2015	Participants outside inclusion criteria (participant selection based on biomarkers)
Casanova 2013	Index test outside inclusion criteria (MRI-derived index: the AD Pattern Similarity (AD-PS) scores))
Cespedes 2017	Participants outside inclusion criteria (healthy control and MCI)

Study	Reason for exclusion
Cevik 2017	Index test outside inclusion criteria (test obtained from multiple volumetric measures)
Chan 2016	Index test was not a volumetric MRI (neuropsychological test)
Chao 2005	Insufficient DTA information (unable to construct 2 x 2 tables); study authors did not answer data request
Cheng 2012	Index test outside inclusion criteria (MRI test obtained from multiple volumetric measures)
Cheng 2015a	Index test outside inclusion criteria (MRI test obtained from multiple volumetric measures)
Cheng 2015b	Index test outside inclusion criteria (MRI test obtained from multiple volumetric measures)
Chertkow 2012	Insufficient description of methods and population (abstract)
Chetelat 2005	Index test outside inclusion criteria (voxel-based-morphometry)
Chincarini 2011	Index test outside inclusion criteria (MRI test obtained from multiple volumetric measures)
Chincarini 2014	Index test outside inclusion criteria (MRI test obtained from multiple volumetric measures)
Chincarini 2016	Index test outside inclusion criteria (annualised volume change)
Cho 2012	Index test was not a volumetric MRI test (brain cortical thickness)
Chow 2015	Not a DTA study (focus on comparison between 2 types of MRI)
Chu 2012	Index test outside inclusion criteria (MRI test obtained from multiple volumetric measures)
Chung 2016	Study design outside inclusion criteria (depressed vs non-depressed MCI participants)
Chupin 2009	ADNI study that reported on a cohort that overlapped with a cohort in another ADNI included paper (Wolz 2011)
Citak-Er 2017	Index test outside inclusion criteria (test obtained from multiple volumetric measures)
Clerx 2013b	Study design outside inclusion criteria (cross sectional study)
Clerx 2014	Insufficient description of methods and population (abstract)
Convit 2000	Insufficient DTA information (unable to construct 2 x 2 tables); unable to contact study authors
Cover 2016	Participants outside inclusion criteria (no MCI participants)
Cui 2011	Index test outside inclusion criteria (mixed test: brain cortical thickness and volumetric measures)
Cuignet 2011	ADNI study that reported on a cohort that overlapped with a cohort in another ADNI included paper (Wolz 2011)
Da 2013	Index test outside inclusion criteria (MRI-derived index: the spatial pattern of abnormalities for recognition of early AD) (SPARE-AD)
Damian 2013	Index test was not a volumetric MRI test (neuropsychological tests)
Davatzikos 2011	Index test outside inclusion criteria (SPARE-AD)

Study	Reason for exclusion
De Leon 1997	Study design outside inclusion criteria (no follow-up data in MCI)
de Leon 2006	Study design outside inclusion criteria (case-control study)
de Leon 2007	Participants outside inclusion criteria (baseline condition: MCI cases mixed with healthy control). Index test was voxel based morphometry
Den Heijer 2006	Participants outside inclusion criteria (healthy participants)
Desikan 2008	Study design outside inclusion criteria (retrospective selection of cases)
Devanand 2008	The study reported accuracy results of combined MRI hippocampal and entorhinal cortex volumes with cognitive test performance, informant report of functional impairment, apolipoprotein E genotype, and olfactory identification deficit
Devanand 2012	Insufficient DTA information (unable to construct 2 x 2 tables); study authors did not answer data request
Dickerson 2001	Study authors not able to supply data request to complete 2 x 2 tables
Dickerson 2004	Index test outside inclusion criteria (not a volumetric MRI test)
Dickerson 2013	ADNI study that reported on a cohort that overlapped with a cohort in another included paper (Wolz 2011)
Douaud 2013	Index test not in line with inclusion criteria (voxel-based-morphometry)
Doyle 2014	Index test not in line with inclusion criteria (MRI-derived index: (Ordinal Regression Characteristic Index of Dementia score)
Duara 2008	Insufficient DTA information (unable to construct 2 x 2 tables); study authors did not answer data request
Duchesne 2009	Study design outside inclusion criteria (retrospective selection of cases)
Duchesne 2014	Index test not in line with inclusion criteria (MRI-derived index, grey matter+determinant)
Duchesne 2015	Index test not in line with inclusion criteria (MRI-derived index, the disease evaluation factor (DEF index))
Dukart 2016	Index test not in line with inclusion criteria (MRI index test obtained from multiple volumetric measures)
Dyrba 2015	Reference standard outside inclusion criteria (AD diagnosis based on CSF biomarkers)
Eckerström 2015	Index test not in line with inclusion criteria (mixed test: both manual and automated techniques, no separate data for the 2 techniques)
Egli 2015	Study design outside inclusion criteria (retrospective selection of cases)
El Fakhri 2003	Insufficient DTA information (unable to construct 2 x 2 tables); unable to contact study authors
El Fakhri 2004	Insufficient DTA information (unable to construct 2 x 2 tables); unable to contact study authors
Ellis 2014	Not a DTA study (focus on cognitive tests)

Study	Reason for exclusion
Eskildsen 2013	Index test was not a volumetric MRI test (cortical thickness)
Eskildsen 2015	Index test not in line with inclusion criteria (MRI-derived index: scoring by Non local Image Patch Estimator (SNIPE))
Evans 2010	Insufficient DTA information (unable to construct 2 x 2 tables). Unable to contact the study author
Ewers 2012	High risk for duplication (study design and MRI parameters similar to Gaser 2013 , population from the same setting ADNI, with a smaller sample size)
Fan 2008	Index test not in line with inclusion criteria (voxel-based-morphometry)
Fellgiebel 2006	Insufficient DTA information (unable to construct 2 x 2 tables); study authors did not answer data request
Fjell 2010	Not a DTA study (focus on relationships between baseline biomarkers and change on CDR score)
Fleisher 2008	Insufficient DTA information (unable to construct 2 x 2 tables); unable to contact study authors
Fouquet 2009	Index test was not a volumetric MRI test (FDG-PET)
Franko 2013	Index test not in line with inclusion criteria (subvolumes of hippocampus)
Gao 2018	Index test outside inclusion criteria (data reported for MRI combined with FDG-PET and participants' information)
Gavidia 2017	Index test not in line with inclusion criteria (MRI-derived index: MRI-based residuals)
Gavrilova 2008	Not a DTA study (focus on neuropsychological test)
Geroldi 2006	Insufficient DTA information (unable to construct 2 x 2 tables); study authors did not answer data request
Gomar 2011	ADNI study that reported on a cohort that overlapped with a cohort in another included paper (Ledig 2018 and Gaser 2013)
Goryawala 2015	Study design outside inclusion criteria (no follow-up data in MCI participants)
Grothe 2013	Insufficient DTA information (unable to construct 2 x 2 tables). Unable to contact the study author
Grundman 2002	Study design outside inclusion criteria (RCT designed to evaluate the efficacy of either vitamin E, donepezil, or placebo to delay progression of MCI to dementia)
Guo 2017	Index test outside inclusion criteria (MRI index test obtained from multiple volumetric measures)
Guzman 2013	Not a DTA study (focus on correlation between biomarkers)
Gómez-Sancho 2018	ADNI study that reported on a cohort that overlapped with a cohort in another ADNI included paper (Ledig 2018)
Hall 2015	Index test not in line with inclusion criteria (data reported for the Disease State Index prediction model)
Hall 2015b	Index test not in line with inclusion criteria (data reported for the Disease State Index prediction model)

Study	Reason for exclusion
Hamalainen 2007	Index test outside inclusion criteria (voxel-based-morphometry)
Hamalainen 2008	Index test outside inclusion criteria (voxel-based-morphometry)
Heister 2011	ADNI study that reported on a cohort that overlapped with a cohort in another ADNI included paper (Wolz 2011)
Henneman 2009	Not a DTA study (focus on association of baseline variables with hippocampal atrophy rate)
Henry-Feugeas 2008	Insufficient DTA information (unable to construct 2 x 2 tables); unable to contact study authors
Hensel 2005	Insufficient DTA information (unable to construct 2 x 2 tables); unable to contact study authors
Hinrichs 2011	Index test not in line with inclusion criteria (data reported for MRI test combined with PET, no separate data for MRI)
Hu 2016	Index test outside inclusion criteria (MRI index test obtained from multiple volumetric measures)
Huang 2017	Index test outside inclusion criteria (test obtained from multiple volumetric measures)
Inui 2017	Index test outside inclusion criteria (test obtained from multiple volumetric measures)
Jack 2004	Insufficient DTA information (unable to construct 2 x 2 tables); study authors did not answer data request
Jack 2005	Insufficient DTA information (unable to construct 2 x 2 tables); study authors did not answer data request
Jack 2008a	Study design outside inclusion criteria (retrospective selection of cases)
Jack 2009	Not a DTA study (focus on ventricular expansion rate in different clinical groups: cognitively normal, MCI, AD)
Jacobs 2011	Participants outside inclusion criteria (baseline condition: MCI cases mixed with healthy control)
Jie 2013	Index test not in line with inclusion criteria (data reported for MRI test combined with PET, no separate data for imaging)
Kalin 2017	Study design outside inclusion criteria (case-control study)
Kaneko 2005	Study design outside inclusion criteria (case study)
Kantarci 2005	Insufficient DTA information (unable to construct 2 x 2 tables); unable to contact study authors
Kantarci 2009	Insufficient DTA information (unable to construct 2 x 2 tables); unable to contact study authors
Karas 2008	Insufficient DTA information (unable to construct 2 x 2 tables); study authors did not answer data request
Kaye 1997	Participants outside inclusion criteria (healthy participants)
Kaye 2005	Not a DTA study (focus on volume changes in different phases of cognitive decline)
Khan 2015b	Index test not in line with inclusion criteria (mixed test; brain cortical thickness and volumetric measures)

Study	Reason for exclusion
Killiany 2000	Participants outside inclusion criteria (participants with normal cognition; participants with “questionable AD”)
Kim 2017	Study design outside inclusion criteria (retrospective selection of cases)
Kloppel 2015	Insufficient DTA information (unable to construct 2 x 2 tables); unable to contact study authors
Kong 2014	Index test not in line with inclusion criteria (data reported for MRI test combined with neuropsychological and genetic tests)
Korf 2004	Insufficient DTA information (unable to construct 2 x 2 tables); study authors did not answer data request
Korolev 2016	Index test not in line with inclusion criteria (mixed test: brain cortical thickness and volumetric measures)
Kovacevic 2009	Participants outside inclusion criteria (6-month follow-up)
Krashenyi 2016	Study design outside inclusion criteria (no follow-up data in MCI participants)
Laforce 2010	Study design outside inclusion criteria (retrospective study)
Lan 2017	Not a DTA study (focus on survival analysis)
Landau 2010	High risk for duplication (study design and MRI parameters similar to Wolz 2011 , population from the same setting ADNI, with a smaller sample size)
Lebedev, 2014	Index test outside inclusion criteria (MRI test obtained from multiple volumetric measures)
Lee 2015	Index test outside inclusion criteria (data reported for MRI test combined with neuropsychological and genetic tests)
Lehman 2013	ADNI study that reported on a cohort that overlapped with a cohort in another included paper (Pereira 2014)
Leung 2010	Not a DTA study and focused on technical aspects of the test
Leung 2013	Not a DTA study and focused on comparisons of atrophy rate among healthy participants, MCI and AD dementia. ADNI study
Li 2012	Not a DTA study (focus on correlation between atrophy and cognitive functions in AD and MCI patients)
Li 2014a	Index test outside inclusion criteria (mixed test: brain cortical thickness and volumetric measures)
Li 2014b	Participants outside inclusion criteria (MCI patients stable during follow-up)
Lillemark 2014	ADNI study that reported on a cohort that overlapped with a cohort in another ADNI included paper (Ledig 2018 and Wolz 2011)
Lin 2018	Index test outside inclusion criteria (test obtained from multiple volumetric measures)
Lindemer 2015	Index test was not a volumetric MRI test (brain white matter signal abnormalities)

Study	Reason for exclusion
Liu 2013	High risk for duplication (study design and MRI parameters similar to Pereira 2014 , population from setting ADNI with a smaller sample size compared to Pereira 2014 , that combine patients from ADNI and AddNeuroMed cohorts)
Liu 2014a	Index test outside inclusion criteria (MRI test obtained from multiple volumetric measures)
Liu 2014b	Index test outside inclusion criteria (MRI test obtained from multiple volumetric measures)
Llano 2011	Insufficient DTA information (unable to construct 2 x 2 tables); unable to contact study authors
Long 2016	Study design outside inclusion criteria (no follow-up data in MCI participants)
Lopez 2016	Index test outside inclusion criteria (data reported for MRI test combined with neuropsychological tests and electroencephalogram)
Luo 2016	Study design outside inclusion criteria (no follow-up data in MCI participants)
Ma 2016	Not a DTA study (comparison between MCI and AD; MCI and normal controls)
MacDonald 2013	Population outside inclusion criteria (participants classified as MCI from baseline to 12 months and prediction of conversion evaluated between 12 and 24 months)
Mah 2015	Study design outside inclusion criteria (anxious vs non-anxious MCI participants)
Mangialasche 2013	Index test outside inclusion criteria (data reported for MRI test combined with vitamin E plasma levels)
Manning 2014	Not a DTA study (focus on hippocampal atrophy rate and comparison between epsilon 4 carriers and non-carriers)
Martínez-Torteya 2015	Study design outside inclusion criteria (case-control study)
Maruyama 2004	Not a DTA study (focus on biomarker comparison between stable and progressive MCI)
Mascalchi 2016	Index test outside inclusion criteria (magnetisation transfer imaging)
Massaro 2004	Study design outside inclusion criteria (no follow-up data in MCI participants)
McEvoy 2009	Index test outside inclusion criteria (mixed test: brain cortical thickness and volumetric measures)
McEvoy 2011	Insufficient description of methods and population (abstract)
Meguro 2016	Index test outside inclusion criteria (data reported for MRI test combined with PET)
Meyer 2005a	Study design outside inclusion criteria (no follow-up data in MCI participants)
Meyer 2005b	Study design outside inclusion criteria (cross sectional study)
Meyer 2007	Study design outside inclusion criteria (no follow-up data in MCI participants)
Miller 2008	Index test was not a volumetric MRI test (functional MRI)
Minhas 2017	High risk for duplication (study design and MRI parameters similar to Wolz 2011 and Ledig 2018 , population from the same setting ADNI, with a smaller sample size)

Study	Reason for exclusion
Moradi 2015	Index test outside inclusion criteria (imaging test obtained from multiple volumetric measures)
Moradi 2016	Index test was not a volumetric MRI test (neuropsychological tests)
Moretti 2015	Not a DTA study (focus on the relationship between EEG markers and the cortical thickness in participants with MCI)
Morra 2009	Not a DTA study (focus on comparison between Alzheimer's disease patients vs MCI patients vs healthy elderly control)
Mubeen 2017	Index test outside inclusion criteria (MRI-derived index)
Mungas 2002	No reference standard
Mungas 2005	Study design outside inclusion criteria (no follow-up data in MCI participants)
Nesteruk 2015	Insufficient DTA information (unable to construct 2 x 2 tables); unable to contact study authors
Nordlund 2005	Index test was not a volumetric MRI test (neuropsychological tests)
Ota 2015	Index test outside inclusion criteria (data reported for MRI test combined with PET)
Ota 2016	Index test outside inclusion criteria (data reported for MRI test combined with PET)
Overdorp 2014	Study design outside inclusion criteria (retrospective selection of cases)
Park 2013	Participants outside inclusion criteria (healthy participants)
Park 2015	Study design outside inclusion criteria (follow-up of MCI participants was in relation to reversion to normal cognitive function)
Peng 2015	Not a DTA study (focus on correlation of hippocampal volume and cognitive performances in MCI participants and AD)
Perani 2015	Study design outside inclusion criteria (retrospective selection of cases)
Persson 2017	Insufficient DTA information (unable to construct 2 x 2 tables); unable to contact study authors
Peters 2014	Index test was not a volumetric MRI test (cortical thickness measure)
Petersen 2010	Index test was not a volumetric MRI (CSF biomarker)
Prasad 2011	Insufficient DTA information (unable to construct 2 x 2 tables); unable to contact study authors
Prestia 2015	High risk for duplication (study design and MRI parameters similar to Prestia 2013 , population from the same setting Brescia, VUMC and Stockholm, with an equal sample size)
Prins 2013	Study design outside inclusion criteria. A post hoc analyses of placebo participants with MCI who participated in a previously published clinical trial
Qiu 2014	ADNI study. Insufficient description of methods and population. Unable to contact the study authors
Querbes 2009	Index test was not a volumetric MRI test (cortical thickness measure)

Study	Reason for exclusion
Raamana 2015	Index test was not a volumetric MRI test (cortical thickness measure)
Rana 2017	Participants outside inclusion criteria (no MCI participants)
Redolfi 2015	Index test was not a volumetric MRI test (cortical thickness measure)
Richard 2013	Study design outside inclusion criteria (case-control study)
Risacher 2010	Index test outside inclusion criteria (voxel-based-morphometry)
Ritter 2016	Participants outside inclusion criteria (clinical uncertain cognitive impairment in elderly patients hospitalised for acute condition)
Runtti 2014	Index test outside inclusion criteria (MRI test obtained from multiple volumetric measures)
Salvatore 2018	Index test outside inclusion criteria (MRI test obtained from multiple volumetric measures)
Sambuchi 2015	Not a DTA study (focus on MRI alterations in subjective cognitive impairment)
Schmitter 2014	Index test outside inclusion criteria (MRI test obtained from multiple volumetric measures)
Schuff 2009	Not a DTA study (focus on atrophy rate in healthy elderly, MCI and AD)
Shaffer 2013	Study design outside inclusion criteria (retrospective selection of cases)
Sheng 2017	Not a DTA study (focus on comparison between MCI vs healthy control)
Sluimer 2009	Study design outside inclusion criteria (prediction of AD risk based on atrophy rate)
Smith 2008	Insufficient DTA information (unable to construct 2 x 2 tables); unable to contact study authors
Sohn 2015	Index test outside inclusion criteria (voxel-based-morphometry)
Song 2013	Reference standard outside inclusion criteria (MMSE and the AD Assessment Scale-cognitive sub-scale)
Sousa 2015	Not a DTA study (focus on neural and behavioural substrates of disorientation in MCI and AD)
Sousa 2016	Index test was not a volumetric MRI test (neuropsychological tests)
Spulber 2010	Index test outside inclusion criteria (annualised atrophy rate)
Spulber 2013	Index test outside inclusion criteria (mixed test, brain cortical thickness and volumetric measures)
Staekenborg 2009	Insufficient DTA information (unable to construct 2 x 2 tables); unable to contact study authors
Stephan 2015	Participants outside inclusion criteria (baseline condition not MCI)
Stonnington 2018	Participants outside inclusion criteria (no MCI participants)
Stoub 2005	Participants outside inclusion criteria (baseline condition MCI mixed with healthy control)
Suk 2014	Index test outside inclusion criteria (MRI test obtained from multiple volumetric measures)

Study	Reason for exclusion
Suppa 2015a	High risk for duplication (study design and MRI parameters similar to Wolz 2011 , population from the same setting ADNI, with a smaller sample size)
Suppa 2015b	Participants outside inclusion criteria (baseline condition no MCI)
Susanto 2015	Not a DTA study (focus on the trajectories of cognitive performance, grey matter volume (GMV), and CSF biomarker during the AD course)
Sørensen 2016	ADNI study that reported on a cohort that overlapped with a cohort in another ADNI included paper (Wolz 2011)
Tang 2014	Study design outside inclusion criteria (case-control study)
Tang 2015	High risk for duplication (study design and MRI parameters similar to Wolz 2011 , population from the same setting ADNI, with a smaller sample size)
Tapiola 2008	Insufficient DTA information (unable to construct 2 x 2 tables); study authors did not answer data request
Tarnanas 2014	Index test was not a volumetric MRI test (cortical thickness measure)
Teipel 2015	Index test outside inclusion criteria (data reported for MRI test combined with PET)
Ten Kate 2017a	Index test outside inclusion criteria. Gray matter volume of the hippocampal, temporal, parietal, and frontal regions was assessed by voxel-based morphometry
Tosun 2010	Study design outside inclusion criteria (no follow-up data in MCI participants)
Trzepacz 2014	Index test outside inclusion criteria (imaging test obtained from multiple volumetric measures)
Trzepacz 2016	Study design outside inclusion criteria (no follow-up data in MCI participants)
Van Maurik 2016	Insufficient description of methods and population (abstract)
Van Rossum 2012	Participants outside inclusion criteria (MCI with positive CSF biomarker)
Vannini 2007	Index test not in line with inclusion criteria (data reported for functional MRI)
Varon 2011	Not a DTA study (focus on correlation between biomarkers)
Varon 2015	ADNI study that reported on a cohort that overlapped with a cohort in another ADNI included paper (Pereira 2014 ; Wolz 2011 ; Ledig 2018)
Vasta 2016	High risk for duplication (study design and MRI parameters similar to Wolz 2011 , population from the same setting ADNI, with a smaller sample size)
Vemuri 2009	Index test outside inclusion criteria (MRI-derived index: the Structural Abnormality Index (STAND) score)
Verma 2018	Index test not in line with inclusion criteria (data reported for a biomarker that combines cognitive and MRI atrophy markers)
Villemagne 2013	Not a DTA study (focus on longitudinal data of amyloid β ($A\beta$) deposition, cerebral atrophy, and cognitive decline during the AD pathology)

Study	Reason for exclusion
Vos 2012	High risk for duplication (study design and MRI parameters similar to Clerx 2013a , population from the same setting DESCRIPA+VUMC, with a smaller sample size: 153 participants)
Vos 2013	Insufficient description of methods. Unable to construct 2 x 2 tables. Study authors did not answer request for data
Wahlund 2003	Not a DTA study (focus on correlation between biomarkers)
Walhovd 2010	Index test outside inclusion criteria (test obtained from multiple volumetric measures)
Wang 2009	Not a DTA study and focused on comparison of structural volume changes between MCI converters and MCI non converters
Wang 2016	Study design outside inclusion criteria (case-control design)
Wee 2013	Index test outside inclusion criteria (MRI test obtained from multiple volumetric measures)
Wei 2016	Index test outside inclusion criteria (mixed test, brain cortical thickness and volumetric measures)
Weise 2015	Not a DTA study (focus on congruence between biomarker)
Westman 2012	Index test outside inclusion criteria (mixed test, brain cortical thickness and volumetric measures)
Whitwell 2007	Participants outside inclusion criteria (only MCI who developed AD during follow-up)
Whitwell 2008	Index test not in line with inclusion criteria (Voxel Based Morphometry Study)
Willette 2014	Index test was not a volumetric MRI test (surface-based-morphometry)
Wolk 2009	Study design outside inclusion criteria (no follow-up data in MCI participants)
Wolz 2010	Study design outside inclusion criteria (prediction of AD risk based on atrophy rate)
Xu 2015	Index test outside inclusion criteria (test obtained from multiple volumetric measures)
Xu 2016	Index test outside inclusion criteria (MRI test obtained from multiple volumetric measures)
Yamaguchi 2002	Study design outside inclusion criteria (retrospective selection of cases)
Yang 2012	High risk for duplication (study design and MRI parameters similar to Ledig 2018 , population from the same setting ADNI, with a smaller sample size)
Ye 2012	Index test outside inclusion criteria (mixed test, brain cortical thickness, surface area, and volumetric measures)
Yi 2016	Study design outside inclusion criteria (participants were selected through a retrospective review of clinical records)
Young 2013	Index test outside inclusion criteria (MRI test obtained from multiple volumetric measures)
Youssofzadeh 2017	Index test outside inclusion criteria (MRI test obtained from multiple volumetric measures)
Yu 2012	ADNI study that reported on a cohort that overlapped with a cohort in another ADNI included paper (Ledig 2018)

Study	Reason for exclusion
Yu 2014	Index test outside inclusion criteria (MRI test obtained from multiple volumetric measures)
Yun 2015	Study design outside inclusion criteria (case-control study)
Zhang 2012a	Index test outside inclusion criteria (MRI test obtained from volumetric measures)
Zhang 2012b	High risk for duplication (study design and MRI parameters similar to Pereira 2014 , population from the same setting ADNI, with a smaller sample size)
Zheng 2015	Index test was not a structural MRI test (cortical thickness)
Zhou 2014	Study design outside inclusion criteria (case-control study)
Zhou 2019	Index test outside inclusion criteria (test obtained from multiple volumetric measures)

AD: Alzheimer's Disease; **ADNI:** Alzheimer's Disease Neuroimaging Initiative; **CSF:** cerebrospinal fluid; **DTA:** diagnostic test accuracy; **EEG:** electroencephalogram; **FDG:** fluorodeoxyglucose; **MCI:** mild cognitive impairment; **MMSE:** Mini-Mental State Examination; **MRI:** magnetic resonance imaging; **MTA:** medial temporal lobe atrophy; **PET:** positron emission tomography; **RCT:** randomised controlled trial; **VUMC:** University Medical Center Amsterdam

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 Hippocampus total	22	2209
2 Hippocampus left	8	525
3 Hippocampus right	8	673
4 Medial temporal lobe total	7	1077
5 Medial temporal lobe left	1	15
6 Medial temporal lobe right	1	15
7 Lateral ventricles	5	1077
8 Entorhinal cortex total	4	529
9 Entorhinal cortex left	3	199
10 Entorhinal cortex right	2	159
11 Whole brain	4	424
12 Medial temporal gyrus	1	343

Test	No. of studies	No. of participants
13 Lateral temporal lobe total	1	13
14 Lateral temporal lobe right	1	29
15 Amygdala total	2	401
16 Amygdala left	2	401
17 Amygdala right	3	435
18 Cortical grey matter	1	343

Test 1. Hippocampus total.

Test 2. Hippocampus left.

Test 3. Hippocampus right.

Test 4. Medial temporal lobe total.

Test 5. Medial temporal lobe left.

Test 6. Medial temporal lobe right.

Test 7. Lateral ventricles.

Test 8. Entorhinal cortex total.

Test 9. Entorhinal cortex left.

Test 10. Entorhinal cortex right.

Test 11. Whole brain.

Test 12. Medial temporal gyrus.

Test 13. Lateral temporal lobe total.

Test 14. Lateral temporal lobe right.

Test 15. Amygdala total.

Test 16. Amygdala left.

Test 17. Amygdala right.

Test 18. Cortical grey matter.

ADDITIONAL TABLES

Table 1. Participants: sociodemographic and clinical characteristics

Study Country	Multi- centric ^a	Age (years) mean ± SD	Number of par- ticipants ^b (% female)	Education (years) mean ± SD	Baseline MMSE mean ± SD	Mean fol- low-up (years)	No. of MCI converters to AD dementia (%)	No. of sta- ble MCI (%)	No. of MCI who convert- ed to other demen- tia (%)
Carmichael 2007 USA	No	86.6 ± 5.9	29 (69%)	-	91.6 ± 5.5 ^c	3.2	12 (45%)	12 (41%)	4 (14%)
Caroli 2007 Italy	No	70.2 ± 6.7	23 (43%)	9.7 ± 4.75	26.9 ± 2.0	1.6	9 (39%)	14 (61%)	1 (4%) ^d
Clerx 2013a Europe	Yes	70.6 ± 7.6	328 (52%)	10.0 ± 3.8	27.0 ± 2.5	2.0	91 (28%)	225 (69%)	12 (3%) ^e
deToledo-Morell 2004 USA	No	81.7 ± 6.9	27 (56%)	16.4 ± 3.1	27.3 ± 1.8	3.0	10 (37%)	17 (63%)	-
Devanand 2007 USA	No	66.8 ± 9.7	139 (56%)	15.2 ± 4.2	27.5 ± 2.2	3.0	35 (25%)	104 (75%)	2 (1%) ^d
Eckerstrom 2008 Sweden	No	67.9 ± 6.7	42 (57%)	11.4 ± 3.6	-	2.0	13 (31%)	21 (50%)	8 (19%)
Eckerstrom 2013 Sweden	No	69.6 ± 6.9	42 (57%) (34 included in analysis)	10.2 ± 3.2	27.7 ± 2.6	2.0	13 (31%)	21 (50%)	8 (19%)
Erten-Lyons 2006 USA	No	86.9 ± 6.6	37 (70%)	13.7 ± 3.7	27.3 ± 1.5	7.6	22 (59%)	14 (38%)	1 (3%)
Frolich 2017 Germany	Yes	65.7 ± 9	115 (42%)	9.5 ± 1.9	27.0 ± 2.1	2.2	28 (24%)	87 (76%)	-
Galton 2005	No	63.7 ± 9.9	29 (48%)	-	26.9 ± 2.4	1.6	11 (38%)	18 (62%)	2 (6%) ^d

Table 1. Participants: sociodemographic and clinical characteristics (Continued)

UK									
Gaser 2013	Yes	75.2 ± 6.9	195 (33%)	16.0 ± 2.7	27.0 ± 1.8	3.0	133 (68%)	62 (32%)	-
USA, Canada									
Herukka 2008	No	71.2 ± 4.5	21 (67%)	-	-	4.2	8 (38%)	13 (62%)	-
Finland									
Jack 2000	No	77.6 ± 8.2	43 (53%)	13.6 ± 3.2	25.7 ± 3.3	3.0	18 (42%)	25 (58%)	-
USA									
Jang 2018	Yes	71.3 ± 7.4	340 (47%)	16 (14-18 ^f)	29 (27-29) ^f	3.0 ^f	69 (20%)	271 (80%)	-
USA, Canada									
Khan 2015	Yes	74.9 ± 6.9	447 (40%)	14.2 ± 4.5	27.0 ± 1.4	1.0	90 (20%)	357 (80%)	-
USA, Canada, Europe									
Ledig 2018	Yes	74.3 in MCI c, 74.4 in MCI nc ^f	343 (41%)	-	26 in MCI c, 28 in MCI non c ^f	2.0	177 (52%)	166 (48%)	-
USA, Canada									
Liu 2010	Yes	73.6 ± 5.8	100 (53%)	9.0 ± 4.0	27.0 ± 2.0	1.0	21 (21%)	79 (79%)	-
Europe									
Monge Argilés 2014	No	72.9 ± 6.9	30 (60%)	-	23.5 ± 2.0	2.0	15 (50%)	15 (50%)	-
Spain									
Nesteruk 2016	No	63.2 ± 9.6	40 (55%)	13.9 ± 2.9	27.5 ± 1.7	2.0	9 (22%)	31 (78%)	-
Poland									
Ong 2015	Yes	72.7 ± 6.6	45 (-)	13.6 ± 3.7	27.3 ± 1.9	2.0	20 (44%)	21 (47%)	4 (9%)
Australia									
Pereira 2014	Yes	74.9 ± 7.3	480 (40%)	13.9 ± 4.6	27.0 ± 1.4	1.0	95 (20%)	385 (80%)	-
USA, Canada, Europe									
Platero 2019	Yes	74.1 ± 5.2	97 (63%)	8.5 ± 4.3	26.5 ± 2.7	3.0	36 (37%)	61 (63%)	-

Table 1. Participants: sociodemographic and clinical characteristics (Continued)

Spain										
Prestia 2013	Yes	66.2 ± 9.4	73 (56%)	-	27.2 ± 1.5	2.4	29 (40%)	44 (60%)	Not reported ^e	
Italy, Netherlands, Sweden										
Prestia 2013 (ADNI)	Yes	73.6 ± 8.6	93 (47%)	-	26.9 ± 1.7	2.7	18 (50%)	18 (50%)	Not reported ^e	
USA, Canada, Italy (only data from Italy were used, see Table 2) (36 included in analysis)										
Prieto del Val 2016	Spain	Yes	69.0 ± 7.0	34 (65%)	7.5 ± 5.7	26.6 ± 2.4	2.0	16 (47%)	18 (53%)	-
Rhodus-Meester 2016	Netherlands	No	70.6 ± 7.3	171 (46%)	5.0 ± 1.0	26.7 ± 1.9	3.0 ^f	104 (61%)	67 (39%)	23 ^e
Netherlands										
VanderFlier 2005	Netherlands	No	75.0 ± 7.0	15 (71%)	10.0 ± 3.0	26.0 ± 2.0	1.8	9 (60%)	6 (40%)	-
Visser 1999	Netherlands	No	78.8 ± 4.5	13 (54%)	7.4 ± 2.3	22.4 ± 2.3	3.0	9 (69%)	4 (31%)	-
Visser 2002	Netherlands	No	64.9 ± 9.5	29 (42%)	10.7 ± 3.2	27.7 ± 1.8	1.9	7 (23%)	20 (67%)	3 (1%)
Wang 2006	Taiwan	No	76.3 ± 4.0	58 (26%)	11.8 ± 4.3	25.9 ± 2.9	1.8	19 (33%)	39 (67%)	-
Taiwan										
Westman 2011	Europe (6 countries)	Yes	74.0 ± 5.8	101 (52%)	8.7 ± 4.3	27.2 ± 1.6	1.0	19 (19%)	82 (81%)	-
Europe (6 countries)										
Wolz 2011	USA, Canada	Yes	74.7 ± 7.9	405 (35%)	15.6 ± 3.2	27.0 ± 1.9	1.5	167 (41%)	238 (59%)	-
USA, Canada										
Wood 2016	UK	Yes	69.1 ± 4.5	15 (27%)	11.7 ± 1.0	27.7 ± 1.3	2.0	9 (60%)	6 (40%)	-
UK										
AD: Alzheimer's disease; MCI: mild cognitive impairment; MCI c: MCI converted to AD; MCI nc: MCI not converted to AD; MMSE: Mini Mental State Examination; SD: standard deviation										

^aAll studies were conducted at memory clinics or tertiary centres.
^bNumber of participants reported in this table are those used in the meta-analysis.
^cModified Mini Mental State Examination.
^dCases excluded from the analysis.
^eCases excluded a priori from the study.
^fMedian value was available instead of mean.

Table 2. Included and excluded studies assessed for overlapping risk

Study	Data sets	Study period	MRI (Tesla)	MRI technique (V, M, A ^a)	MRI region ^b	No. of participants	Fol-low-up mean years	Participants' overlapping risk with other included studies	Decision on inclusion or exclusion of the MRI region
Bouwmann 2007	VUMC	2001-2004	1	V-Scheltens	MTL	59	1.6	Rhodius-Meester 2016	Excluded
Caroli 2007	Brescia (Italy)	2002-2005	1	V-Scheltens	MTL	23	1.6	No	Included
			1	M-DISPLAY	H total, H left, H right	23	1.6	No	Included
Chupin 2009	ADNI	NR	1.5	sA-SNT	H total	210	1.5	Wolz 2011	Excluded
			1.5	A-SACHA	H total	210	1.5	Wolz 2011	Excluded
Clerx 2013a	DESCRIPA + VUMC	NR	1-1.5	V-Scheltens	MTL	328	2	No	Included
				M-Show_I-mages 3.7.0	H total	328	2	No	Included
				A-LEAP	H total	328	2	No	Included
				SIENAX	LV	328	2	No	Included
Cuignet 2011	ADNI	NR	1.5	A-SACHA	H total	104	1.5	Wolz 2011	Excluded
				A-Freesurfer	H total	104	1.5	Wolz 2011	Excluded
Dickerson 2013	ADNI	NR	1.5	A-Freesurfer	H total	111	3	Wolz 2011	Excluded
Eckerstrom 2008	Gothenburg (Sweden)	NR	0.5	M-Hip- posegm	H total	42	2	No	Included

Table 2. Included and excluded studies assessed for overlapping risk (Continued)

Eckerstrom 2013	Gothenburg (Sweden)	NR	0.5	M-Hip- posegm	H left and right	42	2	No	Included
Ewers 2012	ADNI	NR	1.5	A-Freesurfer	H left and right	130 (45)	2.3	Gaser 2013	Excluded
Gaser 2013	ADNI	NR	1.5	A-Freesurfer	H left and right	195	1	No	Included
					H left	195	3	No	Included
					H right	195	3	Ledig 2018	Excluded
Gomar 2011	ADNI	Down- loaded from AD- NI on Au- gust 3, 2009	1.5	A-Freesurfer	H left	320	2	Gaser 2013	Excluded
					H right			Ledig 2018	Excluded
					LV			Ledig 2018	Excluded
					WB			Ledig 2018	Excluded
Gómez-Sancho 2018	ADNI	NR	1.5	A-Freesurfer	H total	183	3	Ledig 2018	Excluded
Heister 2011	ADNI	October 14, 2010	1.5	A-Neuro- Quant	H total	192	3	Wolz 2011	Excluded
Jang 2018	ADNI	Data down- loaded in De- cember 2017	3	V-CVRS scale (Sheltens for MTL)	MTL	340	3	Pereira 2014	Excluded
					GCA (more than one region)	340		No	Excluded
					LV	340		No	Included
Khan 2015c	ADNI + AddNeu- roMed	NR	1.5	A-Freesurfer	H total	447	1	Wolz 2011 and Liu 2010	Included
Landau 2010	ADNI	NR	1.5	A-Freesurfer	H total	85	2	Wolz 2011	Excluded
Ledig 2018	ADNI	NR	1.5-3	A-MALPEM	H total	343	2	Wolz 2011	Excluded
					H right	343	2	Gaser 2013	Included
					EC total	343	2	No	Included
					A left	343	2	No	Included

Table 2. Included and excluded studies assessed for overlapping risk *(Continued)*

					A right	343	2	No	Included
					A total	343	2	No	Included
					MTG	343	2	No	Included
					WB	343	2	No	Included
					LV	343	2	No	Included
					cGM	343	2	No	Included
Lehman 2013	ADNI	Down-loaded from ADNI in June 2011	1.5	V-Schel tens	MTL	394	3	Pereira 2014	Excluded
Lillemark 2014	ADNI	NR	1.5	A-Freesurfer	WB	240	1	Ledig 2018	Excluded
					H total			Wolz 2011	Excluded
Liu 2010^c	AddNeuroMed	NR	1.5	A-Fischl	H total	100	1	Khan 2015	Included
Liu 2013	ADNI	NR	NR	V-Schel tens	MTL	387	3	Pereira 2014	Excluded
Minhas 2017	ADNI	NR	1.5	A-Freesurfer	H total	52	3	Wolz 2011	Excluded
					EC total	52	3	Ledig 2018	Excluded
					LV	52	3	Ledig 2018	Excluded
Pereira 2014	ADNI	NR	1.5-3	V-Schel tens	MTL	480	1	No	Included
	AddNeuroMed								
Prestia 2013	Brescia (Italy) + VUMC + Stockholm	NR	1 1.5 3	A-Freesurfer	H total (the smallest between left and right H)	73	2.4	No	Included

Table 2. Included and excluded studies assessed for overlapping risk (Continued)

Prestia 2013 (ADNI)	ADNI	NR	1.5-3	A-Freesurfer	H total	57	3	Wolz 2011	Excluded
	ADNI	NR	1.5-3	sA-SNT	H total	57	3	Wolz 2011	Excluded
	Brescia	From 2006	1	A-Freesurfer	H total	36	2.2	Prestia 2013	Included
	Brescia	From 2006	1	M-DISPLAY	H total	36	2.2	No	Included
Prestia 2015	Brescia +	NR	1	A-Freesurfer	H total	73	2.4	Prestia 2013	Excluded
	VUMC +		1.5						
	Stockholm		3						
Rhodus- Meester 2016	VUMC	2000-2012	1-1.5	V-Scheltens	MTL	171	3	No	Included
Sørensen 2016	ADNI	28 September 2012	1.5	A-Freesurfer	H total	233	2	Wolz 2011	Excluded
Suppa 2015a	ADNI	NR	1.5	A-VBM +mask	H total	198	1	Wolz 2011	Excluded
							2		
							3		
Tang 2015	ADNI	NR	1.5	A-Freesurfer	H total	222	3	Wolz 2011	Excluded
VanderFlier 2005	VUMC	NR	1.5	M-DISPLAY	H total, H left, H right, MTL total, MTL left, MTL right	15	1.8	No	Included
Varon 2015	ADNI	27 June 2013	1.5	A-FreeSurfer	H total	89	3.2	Wolz 2011	Excluded
				V-Sheltens	MTA			Pereira 2014	Excluded
				A-FreeSurfer	EC total			Ledig 2018	Excluded
Vasta 2016	ADNI	NR	1.5	A-Freesurfer	H total	121	1.5	Wolz 2011	Excluded
Visser 1999	AMSTEL study	NR	0.6	M-developed in	H total LTL	13	3	No No	Included

Table 2. Included and excluded studies assessed for overlapping risk (Continued)

Study	Study location	Risk of bias	Sample size	MRI technique		Region	Number of participants	Standard deviation	Reference	Status
				Visual	Manual					
Visser 2002	Maastricht Memory Clinic	NR	1.5	V-Sheltens	MTL	30	1.9	No	Included	
				M-Showl-mage	H total	30	1.9	No	Included	
Vos 2012	DESCRIPA+ VUMC	2003-2005	1-1.5	A-LEAP	H total	153	2	Clerx 2013a	Excluded	
Westman 2011	AddNeuroMed	NR	1.5	V-Scheltens	MTL	101	1	Pereira 2014	Excluded	
				M-HERMES software	H total	101	1	No	Included	
Wolz 2011 ^c	ADNI	Follow-up stopped in 2011	1.5	A-Lotjonen (fast and robust multi-atlas segmentation)	H total	405	1.5	Khan 2015	Included	
Yang 2012	ADNI	NR	1.5	A-Freesurfer	H total	111	2	Wolz 2011	Excluded	
			1.5	A-Freesurfer	LV	111	2	Ledig 2018	Excluded	
Yu 2012	ADNI	June 2010	1.5	NR	EC	63	2	Ledig 2018		
					LV	63				
					H left	63				
					H right	63				
Zhang 2012b	ADNI	NR	3	V-Scheltens	MTL	53	2	Pereira 2014	Excluded	

ADNI: Alzheimer's Disease Neuroimaging Initiative; **MMSE:** Mini Mental State Examination; **NR:** not reported; **SD:** standard deviation; **VUMC:** University Medical Centre, Amsterdam

^aMRI technique: **V:** visual; **M:** manual; **A:** automated

^bMRI region: **A:** amygdala; **cGM:** cortical grey matter; **EC:** entorhinal cortex; **GCA:** global cortical atrophy; **H:** hippocampus; **MTL:** medial temporal lobe; **LV:** lateral ventricles; **MTG:** medial temporal gyrus; **WB:** whole brain.

^cUncertain risk of overlap between these studies (Khan 2015 did not specify the number of participants in both ADNI and AddNeuroMed studies).

Table 3. Index test: description and common abbreviations

Study	Manufacturer of MRI scanners	Field strength (Tesla)	Brain regions ^a	MRI-B or MRI-L ^b	Technique: visual; quantitative manual; quantitative semi-automated or automated
Carmichael 2007^c	General Electric	1.5	LV, WB	MRI-B + MRI-L	Quantitative automated
Caroli 2007^c	General Electric	1.0	H left, H right, H total	MRI-B	Quantitative manual
			MTL	MRI-B	Visual
Clerx 2013a	Siemens, Philips	1 or 1.5	H total	MRI-B	Quantitative manual
			H total	MRI-B	Quantitative automated
			MTL	MRI-B	Visual
			LV	MRI-B	Quantitative automated
deTolledo-Morell 2004^c	General Electric	1.5	H total, EC total	MRI-B	Quantitative manual
Devanand 2007^c	General Electric	1.5	H left, H right, H total, EC left, EC right, EC total	MRI-B	Quantitative manual
Eckerstrom 2008^c	Philips	0.5	H total	MRI-B + MRI-L	Quantitative manual
Eckerstrom 2013	Philips	0.5	H left, H right	MRI-B	Quantitative manual
Erten-Lyons 2006^c	Not reported	1.5	H total, LV, WB	MRI-B + MRI-L	Quantitative semiautomated
Frolich 2017	Siemens, Philips	1.5	H total	MRI-B	Quantitative automated
Galton 2005	General Electric	1.5	H left, H right, LTL right	MRI-B	Visual
Gaser 2013	Several (ADNI scanners)	1.5	H left H right	MRI-B	Quantitative automated Quantitative automated ^d
Herukka 2008^c	Siemens	1.5	H left, H right, H total, EC left, EC right, EC total	MRI-B	Quantitative manual
Jack 2000^c	General Electric	1.5	H total	MRI-B + MRI-L	Quantitative manual
Jang 2018	Several (ADNI scanners)	3	MTL GCA	MRI-B MRI-B	Visual ^d Visual ^d

Table 3. Index test: description and common abbreviations (Continued)

			LV	MRI-B	Visual
Khan 2015	Several (ADNI and AddNeuroMed scanners)	1.5	H total	MRI-B	Quantitative automated
Ledig 2018	Several (ADNI scanners)	1.5-3	H total	MRI-B	Quantitative automated ^d
			H right, A total, A left, A right, MTG, EC total, WB, LV, cGM	MRI-B- MRI-L	Quantitative automated
Liu 2010	Several (ADNI and AddNeuroMed scanners)	1.5	H total	MRI-B	Quantitative automated
Monge Argilés 2014	General Electric	1.5	MTL	MRI-B	Visual
Nesteruk 2016	Toshiba	1.5	H left, H right, EC left	MRI-B	Quantitative automated
Ong 2015	Not specified	Not reported	H total	MRI-B + MRI-L	Quantitative automated
Pereira 2014	Several (ADNI and AddNeuroMed scanners)	1.5 or 3	MTL	MRI-B	Visual
Platero 2019	General Electric	1.5	H total	MRI-B	Quantitative automated
Prestia 2013 (ADNI)	Several (ADNI scanners)	1.5 or 3	H total	MRI-B	Quantitative automated ^d and semiautomated ^d
	Philips (TOMC)	1.0	H total	MRI-B	Quantitative manual and automated
Prestia 2013^c	PHILIPS, Siemens (TOMC, VUmc, KUHH)	1.0 or 1.5 or 3.0	H total (the smallest between left and right volumes)	MRI-B	Quantitative automated
Prieto del Val 2016	Philips	1.5	A right	MRI-B	Quantitative automated
Rhodius-Meester 2016	Siemens, General Electric	1.0 or 1.5	MTL	MRI-B	Visual
VanderFlier 2005^c	Philips	1.5	H left, H right, H total, MTL left, MTL right, MTL total	MRI-B	Quantitative manual
			WB	MRI-B	Quantitative semiautomated
Visser 1999^c	Teslacon II (Technicare)	0.6	H total, LTL	MRI-B	Quantitative manual
Visser 2002^c	Philips	1.5	H total	MRI-B	Quantitative manual
			MTL	MRI-B	Visual
Wang 2006^c	Siemens	1.5	H left, H right, H total, A left,	MRI-B	Quantitative manual

Table 3. Index test: description and common abbreviations (Continued)

A right, A total					
Westman 2011	Several (AddNeuroMed scanners)	1.5	H total	MRI-B	Quantitative manual
			> 1 region	MRI-B	Quantitative automated ^d
			MTL	MRI-B	Visual ^d
Wolz 2011	Several (ADNI scanners)	1.5	H total	MRI-B	Quantitative automated
Wood 2016	Siemens	1.5	H total	MRI-B	Quantitative automated

^a**A:** amygdala; **cGM:** cortical grey matter; **EC:** entorhinal cortex; **GCA:** global cortical atrophy; **H:** hippocampus; **MTL:** medial temporal lobe; **LTL:** lateral temporal lobe; **LV:** lateral ventricles; **WB:** whole brain.

^b**MRI-B:** MRI-baseline; **MRI-L:** MRI-longitudinal.

^cData received from the study authors.

^dData not used for the analysis (see Table 2).

Table 4. Numbers of participants in subgroup analysis

Brain region	Number of studies	Sample size	Converted to AD dementia (%)	Sensitivity (95% CI)	Specificity (95% CI)	LR+ (95% CI)	LR- (95% CI)
Hippocampus total^a	22	2209	687 (31%)	0.73 (0.64 to 0.80)	0.71 (0.65 to 0.77)	2.53 (2.09 to 3.06)	0.38 (0.29 to 0.50)
Automatic or semiauto- matic technique	11	1698	531 (31%)	0.59 (0.48 to 0.70)	0.66 (0.56 to 0.74)	1.72 (1.25 to 2.36)	0.62 (0.46 to 0.85)
Manual technique	13	551	156 (31%)	0.82 (0.69 to 0.90)	0.74 (0.67 to 0.81)	3.21 (2.42 to 4.27)	0.31 (0.14 to 0.70)
≥ 3 years' follow-up	8	413	156 (38%)	0.71 (0.54 to 0.84)	0.76 (0.67 to 0.82)	2.94 (2.11 to 4.11)	0.38 (0.22 to 0.64)
< 3 years' follow-up	14	1796	513 (29%)	0.74 (0.65 to 0.81)	0.69 (0.61 to 0.76)	2.39 (1.90 to 3.00)	0.31 (0.21 to 0.47)
≥ 70 years old	16	1796	566 (32%)	0.73 (0.64 to 0.81)	0.69 (0.62 to 0.75)	3.10 (2.15 to 4.48)	0.41 (0.16 to 1.03)
< 70 years old	6	413	121 (29%)	0.72 (0.54 to 0.84)	0.77 (0.67 to 0.84)	3.10 (2.15 to 4.48)	0.41 (0.16 to 1.03)
Hippocampus left	8	359	113 (31%)	0.71 (0.62 to 0.79)	0.76 (0.67 to 0.83)	2.95 (2.14 to 4.06)	0.38 (0.28 to 0.51)
Hippocampus right	8	359	113 (31%)	0.81 (0.73 to 0.88)	0.71 (0.61 to 0.80)	2.82 (2.01 to 3.96)	0.23 (0.11 to 0.46)
Medial temporal lobe to- tal	7	1077	330	0.64 (0.53 to 0.73)	0.65 (0.51 to 0.76)	1.81 (1.41 to 2.32)	0.56 (0.46 to 0.67)
Entorhinal cortex total	4	529	229	range: 0.50 to 0.88	range: 0.60 to 1.00	Not computed since no meta- analyses was conducted	

Table 4. Numbers of participants in subgroup analysis (Continued)

Lateral ventricles	5	1077	371	0.57 (0.49 to 0.65)	0.64 (0.59 to 0.70)	1.61 (1.39 to 1.87)	0.66 (0.57 to 0.78)
Whole brain	4	424	220	range: 0.33 to 0.92	range: 0.41 to 1.00	Not computed since no meta-analyses was conducted	

AD: Alzheimer's disease; **CI:** confidence interval; **LR+:** positive likelihood ratio; **LR-:** negative likelihood ratio.

^a Two studies (Clerx 2013a; Prestia 2013 (ADNI)) used both manual and automatic techniques for total hippocampal volume.

APPENDICES

Appendix 1. The NINCDS-ADRDA criteria for clinical diagnosis of Alzheimer's disease

I. The criteria for the clinical diagnosis of PROBABLE Alzheimer's disease include:

- dementia established by clinical examination and documented by the Mini-Mental Test, Blessed Dementia Scale, or some similar examination, and confirmed by neuropsychological tests;
- deficits in two or more areas of cognition;
- progressive worsening of memory and other cognitive functions;
- no disturbance of consciousness;
- onset between ages 40 and 90, most often after age 65; and
- absence of systemic disorders or other brain diseases that in and of themselves could account for the progressive deficits in memory and cognition.

II. The diagnosis of PROBABLE Alzheimer's disease is supported by:

- progressive deterioration of specific cognitive functions such as language (aphasia), motor skills (apraxia), and perception (agnosia);
- impaired activities of daily living and altered patterns of behavior;
- family history of similar disorders, particularly if confirmed neuropathologically; and
- laboratory results of:
 - * normal lumbar puncture as evaluated by standard techniques,
 - * normal pattern or nonspecific changes in EEG, such as increased slow-wave activity, and
 - * evidence of cerebral atrophy on CT with progression documented by serial observation.

III. Other clinical features consistent with the diagnosis of PROBABLE Alzheimer's disease, after exclusion of causes of dementia other than Alzheimer's disease, include:

- plateaus in the course of progression of the illness;
- associated symptoms of depression, insomnia, incontinence, delusions, illusions, hallucinations, catastrophic verbal, emotional, or physical outbursts, sexual disorders, and weight loss;
- other neurologic abnormalities in some patients, especially with more advanced disease and including motor signs such as increased muscle tone, myoclonus, or gait disorder;
- seizures in advanced disease; and
- CT normal for age.

IV. Features that make the diagnosis of PROBABLE Alzheimer's disease uncertain or unlikely include:

- sudden, apoplectic onset;
- focal neurologic findings such as hemiparesis, sensory loss, visual field deficits, and incoordination early in the course of the illness; and
- seizures or gait disturbances at the onset or very early in the course of the illness.

V. Clinical diagnosis of POSSIBLE Alzheimer's disease:

- may be made on the basis of the dementia syndrome, in the absence of other neurologic, psychiatric, or systemic disorders sufficient to cause dementia, and in the presence of variations in the onset, in the presentation, or in the clinical course;
- may be made in the presence of a second systemic or brain disorder sufficient to produce dementia, which is not considered to be the cause of the dementia; and
- should be used in research studies when a single, gradually progressive severe cognitive deficit is identified in the absence of other identifiable cause.

VI. Criteria for diagnosis of DEFINITE Alzheimer's disease are:

- the clinical criteria for probable Alzheimer's disease and
- histopathologic evidence obtained from a biopsy or autopsy.

VII. Classification of Alzheimer's disease for research purposes should specify features that may differentiate subtypes of the disorder, such as:

- familial occurrence;

- onset before age of 65;
- presence of trisomy-21; and
- coexistence of other relevant conditions such as Parkinson's disease.

(McKhann 1984)

Appendix 2. The NIA-AA criteria for the diagnosis of mild cognitive impairment (MCI)

Diagnostic category	Biomarker probability of AD aetiology	A β (PET or CSF)	Neuronal injury (tau, FDG, sMRI)
MCI-core clinical criteria	Uninformative	Conflicting/indeterminant/untested	Conflicting/indeterminant/untested
MCI due to AD – intermediate likelihood	Intermediate	Positive	Untested
MCI due to AD – high likelihood	Highest	Untested Positive	Positive Positive
MCI – unlikely due to AD	Lowest	Negative	Negative

A β : amyloid beta peptide; **AD**: Alzheimer's disease; **CSF**: cerebrospinal fluid; **FDG**: fluorodeoxyglucose; **MCI**: mild cognitive impairment; **sMRI**: structural magnetic resonance imaging; **PET**: positron emission tomography

Appendix 3. Sources searched and search strategies

Below is a table detailing the searches run for this review followed by a search narrative.

Source	Search strategy	Hits retrieved
Cochrane Dementia and Cognitive Improvement Group Specialised Register for DTAs (Date of most recent search: 29 January 2019)	MRI OR sMRI OR "magnetic resonance" OR "MR scan*" OR vMRI OR "volumetric MR"	January 2012: 51
		December 2012: 0
		July 2016: 8
		June 2017: 3
		January 2019: 5
		TOTAL: 67
MEDLINE (Ovid SP) (Ovid MEDLINE Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Ovid MEDLINE Daily and Ovid	1. *Hippocampus/ 2. Hippocampus/pa [Pathology] 3. ((MTL or "medial temporal limbic" or "medial temporal lobe") adj4 atrophy).ti,ab. 4. Gyrus Cinguli/pa [Pathology] 5. Parahippocampal Gyrus/pa [Pathology] 6. (hippocamp* adj4 atrophy).ti,ab.	January 2012: 5966
		December 2012: 543
		July 2016: 1599
		June 2017: 898
		January 2019: 961

(Continued)

MEDLINE) 1946 to present (Date of most re- cent search: 29 January 2019)	7. or/1-6 8. exp Dementia/ 9. Cognition Disorders/ 10. (alzheimer* or dement* or AD).ti,ab. 11. ((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. 12. MCI.ti,ab. 13. ACMI.ti,ab. 14. ARCD.ti,ab. 15. SMC.ti,ab. 16. CIND.ti,ab. 17. BSF.ti,ab. 18. AAMI.ti,ab. 19. LCD.ti,ab. 20. QD.ti,ab. 21. AACD.ti,ab. 22. MNCD.ti,ab. 23. MCD.ti,ab. 24. (nMCI or aMCI or mMCI).ti,ab. 25. ("N-MCI" or "A-MCI" or "M-MCI").ti,ab. 26. ("pre-clinical AD" or "pre-clinical Alzheimer*" or "preclinical AD" or "preclinical alzheimer*").ti,ab. 27. ("dementia prodrome" or "pre-clinical dementia" or "preclinical dementia" or "pre-clinical ADD" or "preclinical ADD").ti,ab. 28. or/8-27 29. 7 and 28 30. Magnetic Resonance Imaging/ 31. ("magnetic resonance imaging" or MRI*).ti,ab. 32. ("MR imag*" or "MR scan*").ti,ab. 33. ((structural adj2 MR*) or (volum* adj2 MR*) or "sMRI" or "vMRI").ti,ab. 34. or/30-33 35. 29 and 34 36. (di or pa or du).fs. 37. or/34,36 38. 7 and 28 and 37 39. exp Dementia/di [Diagnosis]	TOTAL: 9967
--	--	--------------------

(Continued)

40. 39 and (34 or 38)
41. 7 and 39
42. or/38,40-41
43. (animals not (humans and animals)).sh.
44. 42 not 43

Embase (Ovid SP) 1974 to present (Date of most recent search: 29 January 2019)	1. *Hippocampus/	January 2012: 4367
	2. ((MTL or "medial temporal limbic" or "medial temporal lobe") adj4 atroph*).ti,ab.	December 2012: 527
	3. cingulate gyrus/	July 2016: 1815
	4. Parahippocampal Gyrus/	June 2017: 1025
	5. (hippocamp* adj4 atrophy).ti,ab.	January 2019: 1451
	6. or/1-5	TOTAL: 9185
	7. exp Dementia/	
	8. (alzheimer* or dement* or AD).ti,ab.	
	9. ((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.	
	10. MCI.ti,ab.	
	11. ACMI.ti,ab.	
	12. ARCD.ti,ab.	
	13. SMC.ti,ab.	
	14. CIND.ti,ab.	
	15. BSF.ti,ab.	
	16. AAMI.ti,ab.	
	17. LCD.ti,ab.	
	18. QD.ti,ab.	
	19. AACD.ti,ab.	
	20. MNCD.ti,ab.	
	21. MCD.ti,ab.	
	22. (nMCI or aMCI or mMCI).ti,ab.	
	23. ("N-MCI" or "A-MCI" or "M-MCI").ti,ab.	
	24. ("pre-clinical AD" or "pre-clinical Alzheimer*" or "preclinical AD" or "preclinical alzheimer*").ti,ab.	
	25. ("dementia prodrome" or "pre-clinical dementia" or "preclinical dementia" or "pre-clinical ADD" or "preclinical ADD").ti,ab.	
	26. or/7-25	
	27. 6 and 26	

(Continued)

28. *Magnetic Resonance Imaging/
29. ("magnetic resonance imaging" or MRI*).ti,ab.
30. ("MR imag*" or "MR scan*").ti,ab.
31. ((structural adj2 (magnetic or MR*)) or (volum* adj2 (magnetic or MR*)) or "sMRI" or "vMRI").ti,ab.
32. or/28-31
33. 27 and 32
34. (di or pa or du).fs.
35. 32 or 34
36. 6 and 26 and 35
37. exp Dementia/di
38. 37 and (32 or 36)
39. 6 and 37
40. or/36,38-39

PsycINFO (Ovid SP) 1806-January week 4 2019 (Date of most recent search: 29 January 2019)	1. *Hippocampus/	January 2012: 719
	2. ((MTL or "medial temporal limbic" or "medial temporal lobe") adj4 atrophy).ti,ab.	December 2012: 113
	3. Cingulate Cortex/	July 2016: 1354
	4. "gyrus cinguli".ti,ab.	June 2017: 164
	5. "parahippocampal gyrus".ti,ab.	January 2019: 219
	6. (hippocamp* adj4 atrophy).ti,ab.	TOTAL: 2569
	7. or/1-6	
	8. exp Dementia/	
	9. exp Cognitive Impairment/	
	10. (alzheimer* or dement* or AD).ti,ab.	
	11. ((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.	
	12. MCI.ti,ab.	
	13. ACMI.ti,ab.	
	14. ARCD.ti,ab.	
	15. SMC.ti,ab.	
	16. CIND.ti,ab.	
	17. BSF.ti,ab.	
	18. AAMI.ti,ab.	
	19. LCD.ti,ab.	
	20. QD.ti,ab.	

(Continued)

21. AACD.ti,ab.
22. MNCD.ti,ab.
23. MCD.ti,ab.
24. (nMCI or aMCI or mMCI).ti,ab.
25. ("N-MCI" or "A-MCI" or "M-MCI").ti,ab.
26. ("pre-clinical AD" or "pre-clinical Alzheimer*" or "preclinical AD" or "preclinical alzheimer*").ti,ab.
27. ("dementia prodrome" or "pre-clinical dementia" or "preclinical dementia" or "pre-clinical ADD" or "preclinical ADD").ti,ab.
28. or/8-27
29. 7 and 28
30. Magnetic Resonance Imaging/
31. ("magnetic resonance imaging" or MRI*).ti,ab.
32. ("MR imag*" or "MR scan*").ti,ab.
33. ((structural adj2 MR*) or (volum* adj2 MR*) or "sMRI" or "vMRI").ti,ab.
34. or/30-33
35. 29 and 34

BIOSIS Citation Index 1926 to present (ISI Web of Science)	Topic=(Hippocampus OR "hippocampal atrophy" OR (MTL AND atroph*) OR ("medial temporal" AND atroph*) OR ("whole brain" AND atroph*)) AND Topic=(dementia* OR alzheimer* OR BPSD OR lewy OR "cognit* impair*" OR MCI OR VCI OR AD OR ACMI OR ARCD OR SIND OR AAMI OR AACD OR MNCD OR "CDR 0.5") AND Topic=(MRI OR sMRI OR "magnetic resonance" OR "MR scan*" OR vMRI OR "volumetric MR")	January 2012: 1499
(Date of most recent search: 29 January 2019)	Timespan=All Years. Databases=BIOSIS Previews.	December 2012: 176
	Lemmatization=On	July 2016: 482
		June 2017: 230
		January 2019: 396
		TOTAL: 2783
Web of Science Core Collection (1945-present) (ISI Web of Science)	Topic=(Hippocampus OR "hippocampal atrophy" OR (MTL AND atroph*) OR ("medial temporal" AND atroph*) OR ("whole brain" AND atroph*)) AND Topic=(dementia* OR alzheimer* OR BPSD OR lewy OR "cognit* impair*" OR MCI OR VCI OR AD OR ACMI OR ARCD OR SIND OR AAMI OR AACD OR MNCD OR "CDR 0.5") AND Topic=(MRI OR sMRI OR "magnetic resonance" OR "MR scan*" OR vMRI OR "volumetric MR")	January 2012: 2149
(Date of most recent search: 29 January 2019)	Timespan=All Years. Databases=BIOSIS Previews.	December 2012: 266
	Lemmatization=On	July 2016: 811
		June 2017: 588
		January 2019: 925
		TOTAL: 4739
LILACS (BIREME)	RNM OR "magnetic resonance imag\$" OR "MR imaging" OR "imagens de ressonância nuclear magnética" OR "MR scan" OR MRI OR sMRI OR "structural MR" OR "ressonância magnética" [Words] and Hippocampus OR hipocampo OR "hippocampal atrophy" OR "temporal lobe" OR "lóbulo temporal" OR MTL [Words] and Demências OR dementia OR demencias OR demência OR Alzheimer OR Alzheimers OR Alzheimer's OR cognitive OR	January 2012: 10
(Date of most recent search: 29 January 2019)		December 2012: 11
		July 2016: 0

(Continued)

cognitive OR cognitive OR cognition OR “déficit cognitive” OR cognición OR cognição OR
 Memória OR memory OR Memoria [Words] June 2017: 1
 January 2019: 3
TOTAL: 25

TOTAL before de-duplication **29,335**

TOTAL after de-duplication **24,272**

TOTAL after de-duplication and first-assessment by CDCIG information specialists

Jan 2012: 369
 Dec: 2012: 138
 Jul 2016: 1098
 Jun 2017: 439
 Jan 2019: 149
TOTAL: 2193

Search narrative:

The search uses three main concepts:

[A] What is being measured - the diagnostic marker (in this case the hippocampal atrophy) (combined: ln 7)

[B] The population we are interested in (in this case those with some objective cognitive impairment, no dementia - usually referred to as MCI) (combined: ln 28)

[C] The diagnostic test (structural MRI) (combined: ln 34)

The first combination employed for this strategy is a straight-forward [A] AND [B] AND [C] = line 35. On 3 December 2010 this combination retrieved 1459 hits. The majority of potentially relevant references to studies were identified using this simple A and B and C combination. However, some were missed.

The strategy then employs the use of diagnostic floating sub-headings (line 36). The three sub-headings used are: diagnosis; pathology, and diagnostic use. When **OR** combined with concept C (the diagnostic test MRI), this **expands** this concept to capture records that look at diagnosis. Line 37 shows this expansion (3513458 from 278767). This expanded element to concept C was then combined once again with using AND with concepts [A] and [B]. Line 28 shows the result of this new, broader combination = 4569.

The final part of the strategy then takes the outcome of conversion to dementia (line 39) and does three things:

1. AND combines it with the above combination A AND B AND (expanded) C
2. AND combines with just the original concept C (the MRI terms)
3. AND combines with concept A (what is being measured ie the hippocampal atrophy)

Line 42 brings these combinations together with OR

One limit is then applied which deducts those references in which animal only studies are described as these are not relevant to the review.

Appendix 4. Application of the QUADAS-2 tool for assessment of methodological quality of included studies

Domain 1 - Patient selection

Description	Describe methods of participant selection and characteristics of the included population
-------------	--

Type of bias assessed	Selection bias, spectrum bias
-----------------------	-------------------------------

Structural magnetic resonance imaging for the early diagnosis of dementia due to Alzheimer's disease in people with mild cognitive impairment (Review)

178

(Continued)

Review question	People with MCI (symptoms, clinical and neurological examinations, neuropsychological tests) tested by brain structural MRI for early diagnosis of AD
Information collected	Study objectives, study population, selection (inclusion/exclusion criteria), study design, clinical presentation, age, gender, number of enrolled and number available for analysis, setting, place and period of the study
Signalling question	Was a consecutive or random sample of participants enrolled?
Yes	If a consecutive sample or a random sample of eligible participants was included in the study
No	If a non-consecutive sample or a non-random sample of eligible participants was included in the study
Unclear	All studies that did not specify enrolment as a consecutive or random sample of patients were classified as 'no'; therefore none of the included studies were classified as 'unclear'
Signalling question	Did the study avoid inappropriate exclusions?
Yes	If all MCI participants with suspected AD were included, with an exception for those not able to undergo MRI (e.g. participants with metallic implants or claustrophobia) or for those with alternative diagnosis, e.g. a history of major stroke, a history of neurological or major psychiatric disease, in whom AD would not be suspected in clinical practice
No	If the study excluded participants with co-morbidities, e.g. with depression, diabetes, cardiovascular disease, or excluded participants on the basis of MRI findings, e.g. presence of ischaemic lesions, lacune
Unclear	If the study did not provide clear definition of inclusion/exclusion criteria and 'no' judgement was not applicable
Signalling question	Was a case-control design avoided?
Yes	If no selective recruitment of participants with a diagnosis of dementia of any type or MCI and a control group of healthy patients was done, or a nested case-control design systematically and randomly selected from a defined population cohort was used
No	If retrospective selection of participants with a diagnosis of dementia or MCI and a control group of healthy patients was reported. These studies were excluded; therefore none of the included studies were classified as 'no'
Unclear	All studies that did not provide clear definition of the study design were excluded; therefore none of the included studies were classified as 'unclear'
Risk of bias	Could the selection of participants have introduced bias?
High	If 'no' classification for any of the above 2 questions: 'Was a consecutive or random sample of participants enrolled?' and 'Did the study avoid inappropriate exclusions?'
Low	If 'yes' classification for above 3 questions
Unclear	If 'unclear' classification for the above question 'Did the study avoid inappropriate exclusion' and 'high risk' judgement was not applicable
Concerns about applicability	Are there concerns that included participants do not match the review question?

(Continued)

High	If the study population differed from the population defined in the review question in terms of clinical features and co-morbidity, e.g. studies with multiple sets of inclusion criteria with respect to clinical presentation, including participants who would not have undergone MRI in real practice
Low	If the study included a clinically relevant population who would have undergone MRI in real practice
Unclear	If this information was unclear

Domain 2 - Index test

Description	Describe the index test, how it was conducted and interpreted
Type of bias assessed	Test review bias, clinical review bias, interobserver variation bias
Review question	Volumetric imaging of the whole brain or ROI (hippocampus, ventriculi, entorhinal cortex, amygdale, medial temporal lobe, temporal lobe, cingulate gyrus) by either qualitative visual assessment or by quantitative volumetric measurements, including manually outlining the structure and semi-automated or automated computer-based methods
Information collected	Index test name, sequences, ROI, measurement techniques, magnetic field, description of positive case definition by index test as reported, examiners (numbers, level of expertise, blinding), interobserver variability

Signalling question
Were the index test results interpreted without knowledge of results of the reference standard?

Yes	If the volumetric imaging was conducted and interpreted before the clinical diagnosis of AD. If imaging results were interpreted at later date through the follow-up and the study reported a description of whether the interpretation of imaging was performed blind to the clinical diagnosis of AD.
No	If the volumetric imaging was interpreted retrospectively after clinical diagnosis of AD had been done and blinding was not reported
Unclear	If this information was unclear

Signalling question
Did the study provide a clear prespecified definition of what was considered to be a 'positive' result of the index test?

Yes	If study provided clear definition of positive MRI findings, and this was defined before execution/interpretation of MRI
No	If definition of positive MRI result was not provided, or if study described findings derived from MRI and not defined before its execution/interpretation
Unclear	If it was unclear whether the criteria were prespecified

Signalling question
Was the index test performed by a single operator or interpreted by consensus in a joint session?

Yes	If MRI was performed/interpreted by single operator or was interpreted after collegial discussion of the case
No	If MRI was performed/interpreted by various operators for different participants
Unclear	If this information was unclear

(Continued)

Risk of bias	Could the conduct or interpretation of the index test have introduced bias?
High	If 'no' classification for any of the above 3 questions
Low	If 'yes' classification for all the above 3 questions, or if 'unclear' classification for question 'Was the index test performed by a single operator or interpreted by consensus in a joint session?' and 'yes' classification for the remaining 2 questions
Unclear	If 'unclear' classification at least for the question 'Did the study provide a clear pre-specified definition of what was considered to be a 'positive' result of MRI?' and 'high risk' judgement was not applicable
Concerns about applicability	Are there concerns that the index test, its conduct or its interpretation differs from the review question?
High	We did not consider studies in which MRI looked at other target conditions not specified in the review (e.g. studies aimed at classifying brain atrophy in people with dementia); therefore, none of the included studies was classified as 'high concern'
Low	We considered all types of volumetric MRI modalities, i.e. by manual delineation of regional structures or by semi-automated or automated techniques, as eligible; therefore, all included studies were classified as 'low concern'
Unclear	Only studies with sufficient information on the volumetric MRI were included; therefore, none of the included studies was classified as 'unclear concern'
Domain 3 - Reference standard	
Description	Describe the reference standard, how it was conducted and interpreted
Type of bias assessed	Verification bias, bias in estimation of diagnostic accuracy due to inadequate reference standard
Review question	Target condition - AD; reference standard - clinical follow-up diagnosis of AD according to the criteria of the NINCDS-ADRA; McKhann 1984). Clinical follow-up of ≥ 1 year
Information collected	Target condition, prevalence of target condition in the sample, reference standard, description of positive case definition by reference test as reported, examiners (numbers, level of expertise, blinding)
Signalling question	Is the reference standard likely to correctly classify the target condition?
Yes	If the study reported diagnosis of AD according to the NINCDS-ADRA criteria with a mean clinical follow-up ≥ 1 year
No	If the study reported criteria for the diagnosis of AD not included in the review protocol, such as use of MRI or beta amyloid or combination of biomarkers. These studies were excluded; therefore none of the included studies were classified as 'no'
Unclear	If diagnostic criteria probably were consistent with our methods description, but not fully reported
Signalling question	Were reference standard results interpreted without knowledge of results of the index tests?
Yes	If clinicians diagnosing AD were unaware of the results of the MRI
No	If clinicians diagnosing AD were aware of the results of the MRI
Unclear	If this information was unclear

(Continued)

Risk of bias	Could the reference standard, its conduct or its interpretation have introduced bias?
High	If 'no' classification for the above question 'Were reference standard results interpreted without knowledge of results of the index tests?'
Low	If 'yes' classification for both of the above 2 questions
Unclear	If 'unclear' classification for either of the above 2 questions and 'high risk' judgement was not applicable
Concerns about applicability	Are there concerns that the target condition as defined by the reference standard does not match the question?
High	We excluded studies in which participants did not undergo at least 1 year's follow-up for clinical diagnosis of AD according to the NINCDS-ADRA criteria; therefore, none of the included studies were classified as 'high concern'
Low	In the light of inclusion criteria, all studies were classified as 'low concern'
Unclear	In the light of inclusion criteria, no included studies were classified as 'unclear concern'
Domain 4 - Flow and timing	
Description	Describe any participants who did not receive MRI or the reference standard, or who were excluded from the 2 x 2 table; describe the interval and any interventions between MRI and the reference standard
Type of bias assessed	Bias of diagnostic performance due to different reference standard, missing data
Review question	We had chosen an arbitrary minimum follow-up period of 12 months after MRI to assess whether AD is present. Studies indicate that annual rates of progression from MCI to clinical AD are approximately 10% to 15%
Information collected	Time interval between MRI and clinical diagnosis of AD, withdrawals and losses to follow-up (overall number reported and whether they were explained)
Signalling question	Was there an appropriate interval between index test and reference standard?
Yes	If follow-up period was reported and was at least 12 months
No	We excluded all studies for which the follow-up period was < 12 months; therefore, no included studies were classified as 'no' for this item
Unclear	If the time interval was not stated clearly but the study authors' description allowed one to assume that the interval was reasonably long
Signalling question	Did all participants receive the same reference standard?
Yes	If all MCI patients, or a random sample of them, who received MRI were followed up to receive verification of AD diagnosis according to the NINCDS-ADRA criteria
No	If some of the MCI participants who received MRI were diagnosed throughout the follow-up using different diagnostic criteria, e.g. some received diagnosis of AD according to the NINCDS-ADRA criteria and some according to the composite Dubois's criteria where MRI forms part of the Alzheimer diagnosis (incorporation bias)
Unclear	If this information was unclear

(Continued)

Signalling question	Were all participants included in the analysis?
Yes	If all participants were included in the analysis, or if participants were excluded because they did not meet inclusion criteria or if participants withdrew from the study or were lost to follow-up did not differ systematically from those who remained
No	If any participants were excluded from the analysis because of uninterpretable results, because of non random selection of participants who were followed, e.g. selection was associated with the results of MRI, or reasons for withdrawals were not explained
Unclear	No studies were classified as 'unclear' for this item
Risk of bias	Could the participant flow have introduced bias?
High	If 'no' classification for any of the above 3 questions
Low	If 'yes' classification for all of the above 3 questions
Unclear	If 'unclear' classification for any of the above 3 questions and 'high risk' judgement was not applicable

AD: Alzheimer's disease; **MCI:** mild cognitive impairment; **MRI:** magnetic resonance imaging; **NINCDS-ADRA:** National Institute of Neurological and Communication Disorders and Stroke-Alzheimer's Disease and Related Disorders Association; **ROI:** region of interest

CONTRIBUTIONS OF AUTHORS

Concept: GF, GCa, GFr

Title registration: GF

Protocol draft: GF, GCa, GFr

Protocol editing: GF, GCa, AGB, EC, CL, GCo, GFr

Data extraction: GL, GF, GCa, GC, AGB, EC, CL, GCo

Data entry: AGB, EC, CL, GC, GL

Data analysis: GCa, EL, GV

Drafting the review: GL, GC, GF, GCa, EC, EL, GV

Editing and revising the review: GL, GC, EC, EL, GCa, AGB, CL, GC, GFr, GV, GF

DECLARATIONS OF INTEREST

Gemma Lombardi: none

Giada Crescioli: none

Enrica Cavedo: none

Ersilia Lucenteforte: none

Giovanni Casazza: none

Alessandro-Giacco Bellatorre: none

Chiara Lista: none

Giorgio Costantino: none

Giovanni Frisoni: none

Gianni Virgili: none

Graziella Filippini: none

SOURCES OF SUPPORT

Internal sources

- IRCCS Foundation Neurological Institute 'C. Besta', Milano (Italy), Italy.

External sources

- The Italian Ministry of Health, Italy.

Italian Public Health Research. The Strategic Program 'Introducing new bio-markers in clinical practice for the early diagnosis of Alzheimer disease: methodological, clinical and organisational aspects for the National Health System'.

- NIHR, UK.

This review was supported by the National Institute for Health Research (NIHR), via a Cochrane Programme Grant to the Cochrane Dementia and Cognitive Improvement group. The views and opinions expressed therein are those of the authors and do not necessarily reflect those of the Systematic Reviews Programme, NIHR, National Health Service or the Department of Health

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

1. We modified the title according to the main topic of the review.
2. In the Background, we updated the sections on the target condition, index test and role of the index test, focusing on recent discussions on magnetic resonance imaging (MRI) to detect Alzheimer's disease dementia in mild cognitive impairment (MCI) and referring to recommendations from current guidelines. We removed older references.
3. Methods
 - a. We updated criteria for considering studies for this review as follows
 - i. Participants: following ad hoc observation, we included participants with MCI and other concomitant clinical conditions for whom clinicians would suspect Alzheimer's disease dementia and who would undergo neuroimaging in clinical practice. Exclusion of such participants could result in loss of generalisability of review results.
 - ii. Index test: accuracy data of the cingulate cortex region alone were not available in the included studies. Articles that reported a single result of MRI accuracy derived from multiple volumetric features (e.g. multiple regions of interest) were excluded because of a wide heterogeneity in the number and brain areas considered in such studies. For the same reason, we excluded studies which detected a pattern of brain atrophy involving multiple brain areas, e.g. MRI-derived index as the spatial pattern of abnormalities for recognition of early Alzheimer's disease (SPARE-AD) in the ADNI studies. Studies in which interpretation of neuroimaging was performed by an automatic classifier has been accepted only when MRI accuracy results were based on individual volumetric features (as in [Wolz 2011](#) or [Khan 2015](#)). If a study considered more than one classifier, we selected results obtained by the most performant classifier. Longitudinal changes of brain regions' volumes were not included in the review.
 - iii. Target condition: this review is only concerned with dementia due to Alzheimer's disease. We did not report diagnostic accuracy of MRI for other types of dementia. All the included studies reported the number of participants with MCI who converted to dementia due to Alzheimer's disease. Nine studies reported also the number of participants with MCI who converted to other types of dementia, but these patients were excluded from the analysis by authors in three studies, or if included, accuracy of MRI was not estimated for each type of dementia in the other six studies. ([Table 1](#)).
 - b. Assessment of methodological quality: we used the QUADAS-2 tool.
4. Analysis
 - a. We decided to estimate structural MRI accuracy of atrophy rates over time in a future review.
 - b. In the protocol we planned to evaluate patient spectrum (age of participants, amnesic versus non-amnesic MCI), duration of follow-up, MRI region of interest, and MRI techniques as sources of heterogeneity. We could only investigate the impact of MRI technique,

participants' age and length of follow-up because the included studies did not provide sufficient data or information to evaluate amnesic versus non-amnesic MCI, and MRI region of interest.

- c. Because very few studies reached a follow-up of more than three years, we changed the cut-off value for subgroup analyses on follow-up time from 'three years or less versus more than 3 years' to 'less than three years versus at least 3 years'.
 - d. In the protocol we planned to evaluate setting (referral centres versus population cohorts) and MRI Tesla as sources of heterogeneity. We performed no assessment of heterogeneity for setting and MRI Tesla because all included participants had been referred to tertiary centres and the majority of included studies used 1.5 Tesla.
 - e. In the protocol we planned to investigate the influence of study quality on accuracy estimates with a sensitivity analysis excluding studies at high risk of bias in order to establish if these studies have an effect on overall accuracy estimates. However, there were not enough studies to conduct these analyses since almost all studies had at least one domain at high risk of bias.
 - f. We did not calculate negative and positive predictive values as was planned since, given the low diagnostic accuracy of the test, we did not consider it useful to calculate them.
 - g. We had planned that if a study considered more than one classifier, we would select results obtained by the most performant classifier based on the Youden index (Youden 1950). This was not necessary in this version of the review.
5. We changed the list and order of review authors to reflect author contributions over time.