UNIVERSITY of York

This is a repository copy of *AD-8 for detection of dementia across a variety of healthcare settings*.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/id/eprint/151361/

Version: Published Version

**Article:**
Hendry, Kirsty, Green, Claire, McShane, Rupert et al. (6 more authors) (2019) AD-8 for detection of dementia across a variety of healthcare settings. Cochrane Database of Systematic Reviews. CD011121. ISSN: 1469-493X

https://doi.org/10.1002/14651858.CD011121.pub2

**Reuse**
Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

**Takedown**
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

White Rose university consortium
Universities of Leeds, Sheffield & York

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk/

Cochrane Library

Cochrane Database of Systematic Reviews

# AD-8 for detection of dementia across a variety of healthcare settings (Review)

Hendry K, Green C, McShane R, Noel-Storr AH, Stott DJ, Anwer S, Sutton AJ, Burton JK, Quinn TJ

Hendry K, Green C, McShane R, Noel-Storr AH, Stott DJ, Anwer S, Sutton AJ, Burton JK, Quinn TJ.
AD-8 for detection of dementia across a variety of healthcare settings.
*Cochrane Database of Systematic Reviews* 2019, Issue 3. Art. No.: CD011121.
DOI: 10.1002/14651858.CD011121.pub2.

**www.cochranelibrary.com**


Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

WILEY

## TABLE OF CONTENTS


HEADER . . . 1
ABSTRACT . . . 1
PLAIN LANGUAGE SUMMARY . . . 2
SUMMARY OF FINDINGS FOR THE MAIN COMPARISON . . . 4
BACKGROUND . . . 5
OBJECTIVES . . . 7
METHODS . . . 7
RESULTS . . . 10
Figure 1. . . . 11
Figure 2. . . . 13
Figure 3. . . . 15
Figure 4. . . . 17
Figure 5. . . . 19
Figure 6. . . . 21
ADDITIONAL SUMMARY OF FINDINGS . . . 22
DISCUSSION . . . 27
AUTHORS' CONCLUSIONS . . . 28
ACKNOWLEDGEMENTS . . . 29
REFERENCES . . . 29
CHARACTERISTICS OF STUDIES . . . 34
DATA . . . 54
Test 1. AD-8 informant cutpoint 1. . . . 54
Test 2. AD-8 informant cutpoint 2. . . . 55
Test 3. AD-8 informant cutpoint 3. . . . 55
Test 4. AD-8 informant cutpoint 7. . . . 56
Test 5. Sensitivity analysis removing low average age. . . . 56
Test 6. Sensitivity analysis removing disease-specific study. . . . 57
Test 7. English language studies - cutpoint 2. . . . 57
Test 8. English language studies - cutpoint 3. . . . 58
Test 9. Sensitivity analysis removing CDR 0.5 - cutpoint 2. . . . 58
Test 10. Sensitivity analysis removing CDR 0.5 - cutpoint 3. . . . 59
ADDITIONAL TABLES . . . 59
APPENDICES . . . 60
CONTRIBUTIONS OF AUTHORS . . . 77
DECLARATIONS OF INTEREST . . . 77
SOURCES OF SUPPORT . . . 78
DIFFERENCES BETWEEN PROTOCOL AND REVIEW . . . 78




**Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.**

[Diagnostic Test Accuracy Review]

# AD-8 for detection of dementia across a variety of healthcare settings

Kirsty Hendry[1], Claire Green[1], Rupert McShane[2], Anna H Noel-Storr[2], David J Stott[1], Sumayya Anwer[3], Alex J Sutton[4], Jennifer K Burton[5], Terry J Quinn[1]

[1]Institute of Cardiovascular and Medical Sciences, University of Glasgow, Glasgow, UK. [2]Radcliffe Department of Medicine, University of Oxford, Oxford, UK. [3]Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK. [4]Department of Health Sciences, University of Leicester, Leicester, UK. [5]Academic Geriatric Medicine, Institute of Cardiovascular and Medical Sciences, University of Glasgow, Glasgow, UK

Contact address: Terry J Quinn, Institute of Cardiovascular and Medical Sciences, University of Glasgow, New Lister Building, Glasgow Royal Infirmary, Glasgow, G4 OSF, UK. Terry.Quinn@glasgow.ac.uk.

**Editorial group:** Cochrane Dementia and Cognitive Improvement Group.
**Publication status and date:** New, published in Issue 3, 2019.

**Citation:** Hendry K, Green C, McShane R, Noel-Storr AH, Stott DJ, Anwer S, Sutton AJ, Burton JK, Quinn TJ. AD-8 for detection of dementia across a variety of healthcare settings. *Cochrane Database of Systematic Reviews* 2019, Issue 3. Art. No.: CD011121. DOI: 10.1002/14651858.CD011121.pub2.

Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

## ABSTRACT

### Background

Dementia assessment often involves initial screening, using a brief tool, followed by more detailed assessment where required. The AD-8 is a short questionnaire, completed by a suitable 'informant' who knows the person well. AD-8 is designed to assess change in functional performance secondary to cognitive change.

### Objectives

To determine the diagnostic accuracy of the informant-based AD-8 questionnaire, in detection of all-cause (undifferentiated) dementia in adults. Where data were available, we described the following: the diagnostic accuracy of the AD-8 at various predefined threshold scores; the diagnostic accuracy of the AD-8 for each healthcare setting and the effects of heterogeneity on the reported diagnostic accuracy of the AD-8.

### Search methods

We searched the following sources on 27 May 2014, with an update to 7 June 2018: ALOIS (Cochrane Dementia and Cognitive Improvement Group), MEDLINE (Ovid SP), Embase (Ovid SP), PsycINFO (Ovid SP), BIOSIS Previews (Thomson Reuters Web of Science), Web of Science Core Collection (includes Conference Proceedings Citation Index) (Thomson Reuters Web of Science), CINAHL (EBSCOhost) and LILACS (BIREME). We checked reference lists of relevant studies and reviews, used searches of known relevant studies in PubMed to track related articles, and contacted research groups conducting work on the AD-8 to try to find additional studies. We developed a sensitive search strategy and used standardised database subject headings as appropriate. Foreign language publications were translated.

### Selection criteria

We selected those studies which included the AD-8 to assess for the presence of dementia and where dementia diagnosis was confirmed with clinical assessment. We only included those studies where the AD-8 was used as an informant assessment. We made no exclusions in relation to healthcare setting, language of AD-8 or the AD-8 score used to define a 'test positive' case.


**Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.**

**Data collection and analysis**

We screened all titles generated by electronic database searches, and reviewed abstracts of potentially relevant studies. Two independent assessors checked full papers for eligibility and extracted data. We extracted data into two-by-two tables to allow calculation of accuracy metrics for individual studies. We then created summary estimates of sensitivity, specificity and likelihood ratios using the bivariate approach and plotting results in receiver operating characteristic (ROC) space. We determined quality assessment (risk of bias and applicability) using the QUADAS-2 tool.

**Main results**

From 36 papers describing AD-8 test accuracy, we included 10 papers. We utilised data from nine papers with 4045 individuals, 1107 of whom (27%) had a clinical diagnosis of dementia. Pooled analysis of seven studies, using an AD-8 informant cut-off score of two, indicated that sensitivity was 0.92 (95% confidence interval (CI) 0.86 to 0.96); specificity was 0.64 (95% CI 0.39 to 0.82); the positive likelihood ratio was 2.53 (95% CI 1.38 to 4.64); and the negative likelihood ratio was 0.12 (95% CI 0.07 to 0.21). Pooled analysis of five studies, using an AD-8 informant cut-off score of three, indicated that sensitivity was 0.91 (95% CI 0.80 to 0.96); specificity was 0.76 (95% CI 0.57 to 0.89); the positive likelihood ratio was 3.86 (95% CI 2.03 to 7.34); and the negative likelihood ratio was 0.12 (95% CI 0.06 to 0.24).

Four studies were conducted in community settings; four were in secondary care (one in the acute hospital); and one study was in primary care. The AD-8 has a higher relative sensitivity (1.11, 95% CI 1.02 to 1.21), but lower relative specificity (0.51, 95% CI 0.23 to 1.09) in secondary care compared to community care settings.

There was heterogeneity across the included studies. Dementia prevalence rate varied from 12% to 90% of included participants. The tool was also used in various different languages. Among all the included studies there was evidence of risk of bias. Issues included the selection of participants, conduct of index test, and flow of assessment procedures.

**Authors' conclusions**

The high sensitivity of the AD-8 suggests it can be used to identify adults who may benefit from further specialist assessment and diagnosis, but is not a diagnostic test in itself. This pattern of high sensitivity and lower specificity is often suited to a screening test. Test accuracy varies by setting, however data in primary care and acute hospital settings are limited. This review identified significant heterogeneity and risk of bias, which may affect the validity of its summary findings.

## PLAIN LANGUAGE SUMMARY

**How accurate is the AD-8 informant questionnaire for diagnosing dementia in all healthcare settings?**

**Why is recognising dementia important?**

Many people are living with dementia but have never had the condition diagnosed. Not recognising dementia when it is present (a false negative result) may deny people access to social support, medications, and financial assistance. It also prevents the individual and their family from planning for the future. However, incorrectly diagnosing dementia when it is not present (a false positive result) can cause distress or fear and lead to additional investigations which can waste resources.

**What is the aim of the review?**

The aim of this Cochrane Review was to find out how accurate the AD-8 informant questionnaire is for detecting dementia in all healthcare settings. The researchers included 10 studies to answer this question, nine of which reported numerical information that could be used.

**What was studied in the review?**

The AD-8 questionnaire includes eight 'yes or no' questions, to be answered by someone who knows the person under investigation; for example, a relative, carer or close friend (sometimes described as an informant). The questions ask about whether the informant has noticed a change in the individual's memory and thinking abilities over the past years. A point is given for every item where they think the person's abilities have changed. Higher scores occur when more changes are noted by the informant. The AD-8 would not usually be used to make a final diagnosis of dementia, but it may help identify those who require further assessment.

**What are the main results of the review?**


**Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.**

The review included data from nine relevant studies, with a total of 4045 participants.

Seven of the studies used a score of two or more to indicate dementia. A score of two or more is the cut-off recommended for the AD-8. The results of these studies indicate that, in theory, if the AD-8 were to be used to diagnose dementia in a group of 1000 people across all healthcare settings, of whom 280 (28%) have dementia, an estimated 517 would have an AD-8 result indicating dementia is present and of these 259 (50%) will not have dementia. Of the 483 people with a result indicating that dementia is not present, 22 (5%) would be incorrectly classified as not having dementia

It is possible that the AD-8 may work differently in different settings, for example in hospital or General practice. In secondary care, the AD-8 produces more false positive results, but fewer false negative results, than when it is used in community settings. We could not directly compare the performance of the AD-8 in primary care as there was only one study from this setting.

**How reliable are the results of the studies in this review?**

In the included studies, the diagnosis of dementia was made by assessing all patients with a detailed clinical assessment. (In these studies, detailed clinical assessment was the gold standard we compared the AD-8 to.) This is likely to have been a reliable method for deciding whether patients really had dementia. However, there were some problems with how the studies were conducted. This may have resulted in the AD-8 appearing more accurate than it really is. The numbers described are an average across studies in the review. However, as estimates from individual studies varied we cannot be sure that the AD-8 will always produce these results.

**Who do the results of this review apply to?**

Studies included in the review were conducted in Brazil, China, Japan, Singapore, Taiwan, the UK and USA. Studies included those attending primary and secondary care services, and populations of older adults living in the community. Five studies used the English-language version of the AD-8. The percentage of people with a final diagnosis of dementia was between 12% and 90% (an average of 38%).

**What are the implications of this review?**

The studies included in this review suggest the AD-8 can identify adults who may have a diagnosis of dementia, who would benefit from specialist assessment and diagnosis. If the AD-8 was used alone to diagnose dementia, the chance of wrongly diagnosing someone with dementia when they do not actually have it is high (50% of those whose AD-8 score suggests they have dementia). This makes the AD-8 an unsuitable single diagnostic test as it would potentially create anxiety and distress. The chance of missing a diagnosis of dementia is much lower (5% of those whose AD-8 score suggests they do not have dementia when they actually have it). This group will miss out on opportunities to plan their future care and would not be eligible to be assessed for treatment with medicines. These findings should be considered when deciding whether or not to use the AD-8 to test for dementia.

**How up-to-date is this review?**

The review authors searched for and used studies published up to June 2018.


**Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.**

Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

## SUMMARY OF FINDINGS FOR THE MAIN COMPARISON *[Explanation]*

| Study ID | Country | Subjects (n) | Age (Years) (Mean, SD) | AD-8 language | Cut-off scores reported | Reference standard |
|---|---|---|---|---|---|---|
| Chan 2016 | Singapore | 309 | | Malay | 3 | Dementia (CDR and DSM-IV) |
| Chio 2017 | Taiwan | 153 | 76.9 (8.6) | Unclear | 1, 2, 3, 4 | Early cognitive impairment (CDR) |
| Correia 2011 | Brazil | 109 | | Portuguese | 2, 3 | Dementia (DSM-IV) |
| Galvin 2006 | US | 255 | 73.3 (11.3) | English | 2, 3 | Various criteria for dementia |
| Galvin 2007 | US | 325 | 76.8 (8.9) | English | 1, 2, 3 | Dementia (CDR) |
| Jackson 2016 | UK | 125 | 84.4 | English | 3, 7 | Pre-delirium dementia (DSM-IV) |
| Larner 2015 | UK | 212 | Median 64.5 (IQR 12.7) | English | 2 | Dementia (DSM-IV) |
| Meguro 2015 | Japan | 572 | | Japanese | 2 | Dementia (CDR) |
| Razavi 2014 | US | 186 | 77.8 (8.2) | English | 2 | Various criteria for dementia |
| Yang 2016 | China | 2063 | 79.5 (7.6) | Unclear | 2 | Dementia (NIA-AA criteria) |

CDR: Clinical Dementia Rating scale
DSM-IV: Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition
NIA-AA: National Institute on Aging and Alzheimer's Association

# BACKGROUND

Dementia is a substantial and growing public health concern (Livingston 2017; Prince 2013). As an example, depending on case definition employed, contemporary estimates of dementia prevalence in the United States are in the range 2.5 to 4.5 million individuals (up to 1.6% of the population; 6.5% of those aged over 65) (Hebert 2013). Changes in population demographics will be accompanied by increases in global dementia incidence and prevalence (Ferri 2005). Although the magnitude of the increase in prevalent dementia may have been overestimated in previous prediction models (Matthews 2013), there is no doubt that absolute numbers of older adults with dementia will increase substantially in the short- to medium-term future (Ferri 2005).

A key element of effective management in dementia is a firm diagnosis. Recent guidelines place emphasis on early diagnosis to facilitate improved management and enable informed discussions and planning with patients and carers. The benefits of screening for cognitive decline are debated (Brunet 2012); however, in certain healthcare systems, screening or case-finding has already been introduced for certain groups, e.g. unscheduled hospital admissions of older adults (Burn 2018; Robinson 2015).

Given the projected global increase in dementia prevalence, there is potential tension between the clinical requirements for robust diagnosis at the individual patient level and the need for equitable, easy access to diagnosis at a population level. The ideal would be expert, multidisciplinary assessment informed by various supplementary investigations (neuropsychology; neuroimaging or other biomarkers). This approach is only really feasible in a specialist memory service and is not suited to population screening or case-finding at scale.

In practice, a two-stage process is often employed. This involves initial 'triage' assessments, suitable for use by non-specialists, to select those who require second-stage, further detailed assessment (Boustani 2003).

Various tools for initial cognitive screening have been described (Brodaty 2002; Folstein 1975). Regardless of the methods employed, there is scope for improvement, with observational work suggesting that many people with dementia are not diagnosed (Bradford 2009; Chodosh 2004; Lang 2017; Valcour 2000).

Screening assessment often takes the form of brief, direct cognitive testing. Such an approach will only provide a 'snapshot' of cognitive function. However, a defining feature of dementia is cognitive or neuropsychological change over time. People with cognitive problems themselves may struggle to make an objective assessment of personal change and so an attractive approach is to question collateral sources with sufficient knowledge of the person. These informant-based interviews aim to assess change in function retrospectively.

An instrument prevalent in research and clinical practice, particularly in North America, is the AD-8 informant questionnaire (Galvin 2005). The authors of the tool describe it as the AD-8, without further definition of the acronym (Galvin 2005). However, it is also described as the Ascertaining Dementia interview (Correia 2011), Ascertain Dementia 8 questionnaire (Chen 2018) and the Ascertain Dementia 8-item Informant Questionnaire (Chen 2017) in the literature. This screening/triage tool will be the focus of this review.

A number of properties can be described for a clinical assessment (reliability, responsiveness, feasibility). For our purposes, the test property of greatest interest is diagnostic test accuracy (DTA) ( Cordell 2013).

Although we describe test accuracy of the AD-8 for dementia diagnosis, the AD-8 is not suitable for establishing a clinical diagnosis of dementia when used in isolation. The AD-8 can be described in different ways dependent on the setting and purpose of testing.

## Target condition being diagnosed

The target condition for this DTA review is all-cause dementia (clinical diagnosis).

Dementia is a syndrome characterised by cognitive or neuropsychological decline sufficient to interfere with usual functioning. The neurodegeneration and clinical manifestations of dementia are progressive and at present there is no 'cure', although numerous interventions to slow or arrest cognitive decline have been described, including pharmacological (Birks 2018; McShane 2006), and non-pharmacological (Bahar-Fuchs 2013).

Dementia remains a clinical diagnosis, based on history from the person and suitable collateral sources and direct examination, including cognitive assessment. There is no universally-accepted, ante-mortem, gold-standard diagnostic strategy. We have chosen expert clinical diagnosis as our gold standard (reference standard) for describing AD-8 test properties, as we believe this is most in keeping with current diagnostic criteria and best practice.

Dementia diagnosis can be made according to various internationally-accepted diagnostic criteria, with exemplars being the World Health Organization, International Classification of Diseases (ICD) (ICD-10), and the American Psychiatric Association Diagnostic and Statistical Manual of Mental Disorders (DSM) (DSM-IV) for all-cause dementia and subtypes (Appendix 1). The label of dementia encompasses varying pathologies, of which Alzheimer's disease is the most common. Diagnostic criteria are available for specific dementia subtypes, e.g. NINCDS-ADRDA criteria for Alzheimer's dementia (McKhann 1984; McKhann 2011); McKeith criteria for Lewy body dementia (McKeith 2005); Lund criteria for frontotemporal dementias (McKhann 2001); and the NINDS-AIREN criteria for vascular dementia (Román 1993).


Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

## Index test(s)

Our index test was the AD-8 informant questionnaire (Galvin 2005) (Appendix 2; Appendix 3).

First published in 2005, the AD-8 is a screening tool which has been used to distinguish individuals with dementia or mild cognitive impairment (MCI) from those with normal cognitive function. It is designed to be administered to a relevant proxy, usually a relative or carer, in questionnaire form. The AD-8 is a brief screening tool. With only eight questions, it takes less than three minutes to complete and was developed to replace other lengthy informant questionnaires (Galvin 2006). The AD-8 was originally developed for administration in the English language but has been reproduced in other languages including Brazilian Portuguese (Correia 2011), Taiwanese (Yang 2011), and Korean (Ryu 2009).

The AD-8 items cover domains of judgement, hobby/activity level, repetitive conversations, learning ability, memory in relation to date/appointments, finances and daily thought processes. Informants indicate presence of change 'over several years' using responses of 'Yes, a change', 'No, no change' or 'NA, don't know'. Each 'yes' answer is scored one point. Final scores range from zero, where no change has been noticed by the informant, to eight, where change has been noted across all domains. The commonly employed threshold score for the AD-8 to differentiate cognitive from no cognitive impairment is greater than or equal to two out of eight (i.e. a 'yes' response for two or more items) (Galvin 2005).

The AD-8 has a number of features that potentially make it attractive for clinical and research use. The questions used have an immediacy and relevance that is likely to appeal to users. Assessment and (informant) scoring is brief, and as the scale is not typically interviewer-administered it requires minimal training in application and scoring. There are data to suggest that, compared to standard direct assessments, informant interviews may be less prone to bias from cultural norms and previous level of education (Jorm 2004). As diagnostic criteria for dementia make explicit reference to documenting decline and involving collateral informants, the potential utility of an informant interview tool such as the AD-8 is clear.

## Clinical pathway

Dementia develops over a trajectory of several years, and tests may be performed at different stages in the dementia pathway. In this review we considered any use of the AD-8 as an initial assessment for cognitive decline, and did not limit our inclusion criteria to a particular healthcare setting. We have operationalised the various settings where the AD-8 may be used as secondary care, primary care, and community.

In secondary-care settings, patients will have been referred for expert input but not exclusively due to memory complaints. Cognitive testing in secondary care involves two main groups: opportunistic screening of adults presenting as unscheduled admissions to hospitals, and those people referred to specific services for dementia, memory or psychiatry of older age. Both populations will have a high prevalence of cognitive disorders and mimics. In-hospital estimates of dementia prevalence vary from 15% to 42%, dependent on methodology used (Jackson 2017). Secondary-care patients are more likely to have had a degree of prior cognitive assessment than those in other settings, although we recognise that cognitive screening prior to referral to specialist services is neither consistent nor guaranteed (Menon 2011).

The prevalence of dementia in primary care is estimated to be 1.4% (Donegan 2017), and it has been projected that most General practitioners could expect one to two new cases each year, per physician (Iliffe 2009). However, cognitive testing in the General practice/primary-care setting is likely to be conducted in response to an individual person self-presenting because of subjective memory complaints, a common finding in older adults, estimated to be reported by 25% to 50% of those aged 65 years and over (Iliffe 2010). Using the AD-8 in this setting could be described as 'triage' or 'case finding' to determine individuals with objective concerns requiring further investigation.

In the community setting, the cohort is largely unselected.The estimated prevalence of dementia in those aged 65 years and over is 6.5% (Matthews 2013) and this age group accounts for ~18% of the population (Office for National Statistics 2018). This means dementia prevalence will be low compared to other settings. In the community setting, use of the AD-8 as a cognitive testing approach may be described as 'population screening'.

Most studies of test accuracy compare the test against contemporaneous reference standards (in this case, clinical dementia diagnosis). An alternative is to describe the test properties for detection of early, 'pre-clinical' problems that are formally diagnosed during prospective, longitudinal follow-up. This delayed verification approach is commonly employed in studies describing properties of dementia biomarkers, but may have utility for other test strategies such as informant interview.

## Rationale

There is no consensus on the optimal initial assessment for dementia, and choice is currently dictated by experience with a particular instrument, time constraints and training. A better understanding of the diagnostic properties of various strategies would allow for an informed approach to testing. Critical evaluation of the evidence base for screening tests or other diagnostic markers is of major importance. Without a robust synthesis of the available information there is the risk that future research, clinical practice and policy will be built on erroneous assumptions about diagnostic validity.

The AD-8 is commonly used in practice and research; it is used internationally and is one of only a few validated informant-based screening/diagnostic tools. A body of literature describing the test accuracy of the AD-8 in different settings is available, although some of these studies have been modest in size. Thus, systematic


**Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.**

review and, if possible, meta-analysis of the diagnostic properties of the AD-8 is warranted.
This review forms part of a body of work describing the diagnostic properties of commonly-used dementia tools (Appendix 4). At present we are conducting single-test reviews and meta-analyses. However, the intention is then to collate these data and perform an overview, to enable comparison of various test strategies.

# OBJECTIVES

To determine the diagnostic accuracy of the informant-based AD-8 questionnaire, in detection of all-cause (undifferentiated) dementia in adults.

## Secondary objectives

Where data were available, we described the following.

1. The diagnostic accuracy of the AD-8 at various predefined threshold scores (one, two, three). These thresholds have been chosen to represent the range of cut-off scores that are commonly used in practice and research; we have been inclusive in our choice of cut-off score to maximise available data for review.
2. The diagnostic accuracy of the AD-8 for each healthcare setting (community; primary care; secondary care).
3. The effects of heterogeneity on the reported diagnostic accuracy of the AD-8. Potential sources of heterogeneity that we intended to explore were: case mix of cohort; method of dementia diagnosis; method (language) of AD-8 assessment.

# METHODS

## Criteria for considering studies for this review

### Types of studies

We anticipated that the majority of the studies would be of AD-8 test properties compared against a contemporaneous clinical diagnosis of dementia in secondary-care settings. We included test studies performed in other healthcare settings, and classified settings as secondary care, primary care, or community.
Case-control studies are known to potentially overestimate properties of a test and we did not include such studies in the review (Davis 2013).
We also did not include case studies or samples with very small numbers (chosen as 10 participants, for the purposes of this review).
Where settings were mixed, for example, a population study 'enriched' with additional cases from secondary care, we did not consider such studies unless separate data were presented for participants from each setting. This design can be affected by similar biases to a case-control design.

### Participants

All adults (aged over 18 years) were eligible.
We did not predefine exclusion criteria relating to the 'case mix' of the population studied, but assessed this aspect of the study as part of our assessment of heterogeneity.

### Index tests

Studies must have included (not necessarily exclusively) the AD-8, used as an informant questionnaire.
The AD-8 has been translated into various languages to enable international administration. The properties of a translated AD-8 in a cohort of non-English speakers may differ from the properties of the original English-language questionnaire. We collected data on the principal language used for AD-8 assessment.
For this review we did not consider other cognitive screening/assessment tools. Where a paper described the AD-8 with in-study comparison against another screening tool, we included the AD-8 data only. Where the AD-8 was used in combination with another cognitive screening tool, we included the AD-8 data only.

### Target conditions

We included any clinical diagnosis of all-cause (unspecified) dementia. We did not require a definition of a particular dementia subtype, although we recorded this information where available.

### Reference standards

Our reference standard is clinical diagnosis of dementia. We recognise that clinical diagnosis itself has a degree of variability but this is not unique to dementia studies and does not invalidate the basic diagnostic test accuracy (DTA) approach.
The primary analysis for clinical diagnosis included all-cause (unspecified) dementia, using any recognised diagnostic criteria (for example, DSM-IV; ICD-10). We included use of the full Clinical Dementia Rating (CDR) Scale (Morris 1993) as an acceptable method for making a dementia diagnosis, with a CDR score greater than or equal to one for a clinical diagnosis of dementia. Dementia diagnosis may have specified a pathological subtype and we included all common dementia subtypes, e.g. NINCDS-ADRDA (Alzheimer's disease), Lund-Manchester (frontotemporal dementia), McKeith (Dementia with Lewy bodies), NINCDS-AIREN (vascular dementia) (McKeith 2005; McKhann 1984; McKhann 2001; Román 1993). We did not define preferred diagnostic criteria for rarer forms of dementia (e.g. alcohol-related; HIV-related;


Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

prion disease-related). These were considered under our rubric of 'all-cause' dementia, and not considered separately.

The label 'dementia' can span a range of disease severities, from mild disease to 'end stage'. The diagnostic properties of a tool will vary depending on disease stage, for example, true positives are more likely when disease is advanced and diagnosis is clear. For our primary analysis we included any dementia diagnosis, at any stage of disease.

Clinicians may use imaging, pathology or other data to aid diagnosis; however, we did not include diagnosis based only on these data without corresponding clinical assessment. We recognise that different iterations of diagnostic criteria may not be directly comparable, and that diagnosis may vary with the degree or manner in which the criteria have been operationalised (e.g. individual clinician versus algorithm versus consensus determination). We collected data on method and application of dementia diagnosis for each study and explored the validity of the dementia diagnosis as part of our 'Risk of Bias' assessment. We did not accept use of other (brief) direct performance tests in isolation as a basis for diagnosis.

We recognise that dementia diagnosis often comprises a degree of informant assessment. Thus there is potential for incorporation bias. We explored the potential effects of this bias through our 'Risk of bias' assessment.

## Search methods for identification of studies

We used a variety of information sources to ensure all relevant studies were included. We devised terms for electronic database searching in conjunction with the team at the Cochrane Dementia and Cognitive Improvement Group. As this AD-8 review forms part of a suite of reviews looking at informant scales, we have created a comprehensive search strategy designed to pick up all cognitive assessment scales (Quinn 2014); we complemented this generic search with searches specific to AD-8 terminology.

### Electronic searches

We searched the specialised register of the Cochrane Dementia and Cognitive Improvement Group, ALOIS (which includes both intervention and diagnostic accuracy studies), MEDLINE (Ovid SP), Embase (Ovid SP), BIOSIS (ISI Web of Science), Web of Science Core Collection (ISI Web of Science), PsycINFO (Ovid SP), CINAHL (EBSCOhost) and LILACS (Bireme). We designed similarly-structured search strategies and used search terms appropriate to each database. We used MeSH words and other controlled vocabulary where appropriate.

We also searched sources specific to diagnostic accuracy or systematic review:

· MEDION database (Meta-analyses van Diagnostisch Onderzoek www.mediondatabase.nl);

· DARE (Database of Abstracts of Reviews of Effects) and HTA Database (Health Technology Assessments Database), both *The Cochrane Library*;

· ARIF database (Aggressive Research Intelligence Facility; www.arif.bham.ac.uk).

See Appendix 5 and Appendix 6 for the search strategies run.

We did not apply language or date restrictions to the electronic searches, nor did we apply restrictions regarding publication status when assessing records for potential inclusion, including abstracts, conference proceedings and unpublished data. We used translation services where necessary (see: Acknowledgements), and used a translation proforma for data extraction (Appendix 7).

The Cochrane Dementia and Cognitive Impairment Group Information Specialist (ANS) ran the initial searches.The most recent search for this review was performed on 7 June 2018.

### Searching other resources

Grey literature and proceedings: we searched our chosen electronic databases and included relevant assessments of conference proceedings. Where possible we accessed theses or PhD abstracts from institutions known to be involved in prospective dementia studies.

Handsearching: we did not perform handsearching as there is little published evidence of the benefits of handsearching for diagnostic studies (Glanville 2010).

Reference lists: we checked the reference lists of all relevant studies and reviews in the field for further possible titles, and repeated the process until no new titles were found (Greenhalgh 2005).

Correspondence: We contacted research groups who have published or are conducting work on AD-8 for dementia diagnosis, informed by results of initial search (see: Acknowledgements).

We used relevant studies in PubMed to search for additional studies using the 'related article' feature. We examined key studies in citation databases, such as Science Citation Index and Scopus, to ascertain details of any further relevant studies.

## Data collection and analysis

### Selection of studies

One review author (ANS) screened all titles generated by initial electronic database searches for relevance. The initial search was a sensitive, generic search, designed to include all potential dementia screening tools. Two review authors (KH, CG) then independently screened all remaining titles for relevance. The two review authors inspected abstracts of selected titles and selected all potentially eligible studies for full-text review. They independently assessed full manuscripts against the inclusion criteria, and resolved disagreements by discussion or by involving an arbitrator where necessary.

Where a study included useable data but did not present these in the published manuscript, we contacted the authors directly to


**Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.**

request further information. If the same data were presented in more than one paper we included the primary paper only.
We have detailed the study selection process in a PRISMA flow diagram.

### Data extraction and management

We extracted data to a study-specific pro forma that included clinical/demographic details of the participants; details of setting; details of AD-8 administration and details of the dementia diagnosis process.
Where AD-8 data were presented for differing threshold scores, we extracted data for each threshold and collated these separately.
We extracted test accuracy data to a standard two-by-two table.
Two review authors (KH, CG), blinded to study identifiers, performed data extraction independently. They resolved disagreements by discussion, with the use of an arbitrator if necessary. As a further check of validity of search and data extraction, a third reviewer (SA) assessed all titles and papers from the most recent updated literature search.
For each included paper, we detailed the flow of participants (numbers recruited, included, assessed) in a flow diagram.

### Assessment of methodological quality

We assessed the methodological quality of each study using the QUADAS-2 tool (www.bris.ac.uk/quadas/quadas-2) (Appendix 8). This tool incorporates domains specific to participant selection; index test; reference standard; and participant flow. Each domain is assessed for risk of bias and the first three domains are also assessed for applicability. Certain key areas important for quality assessment are participant selection, blinding, and missing data. Following a group meeting of review authors, we created guidance for the application of QUADAS-2 to dementia screening assessments, specifically developing anchoring statements for QUADAS-based assessment that are suited to dementia test accuracy studies. This QUADAS-2 guidance was created through a multidisciplinary working group and has been extensively piloted. The process and resulting statements for assessment are described in Appendix 9.
We did not use QUADAS-2 data to form a summary quality score; we produced a narrative summary describing numbers of studies for which we found high/low/unclear risk of bias, or concerns regarding applicability with corresponding tabular and graphical displays.
Paired independent raters (KH, CG), blinded to each other's scores, performed both assessments. They resolved disagreements by further review and discussion, with recourse to a third-party arbitrator where necessary. A third reviewer assessed all titles selected from the most recent literature search.

### Statistical analysis and data synthesis

We were interested in the test accuracy of the AD-8 for the dichotomous variable 'dementia'/'no dementia'. Thus, we applied the current Cochrane DTA framework for analysis of a single test.
We extracted data from included papers to allow creation of a standard two-by-two data table showing dichotomised AD-8 test results (AD-8 positive or AD-8 negative), cross-classified with binary reference standard (dementia or no dementia).
We used Review Manager 5 software (RevMan 2012) to calculate sensitivity, specificity and 95% confidence intervals (CIs) from the two-by-two tables abstracted from the included studies. We used a threshold score of two or more on the AD-8 for primary analyses. If data at other thresholds were presented, we examined these in separate analyses (we were able to describe accuracy at cut-off scores of one, two, three, and seven). We presented individual study results graphically by plotting estimates of sensitivities and specificities as forest plots.
To allow for summary analysis, we used SAS release 9.4, in addition to Review Manager 5. Using the bivariate approach we described metrics of pooled sensitivity, specificity, positive and negative likelihood ratios, all with corresponding 95% confidence intervals. We plotted summary data in receiver operating characteristic (ROC) space, including 95% confidence regions. Where we were interested in comparative accuracy and where data allowed, for example differential accuracy by setting, we plotted summary accuracy data for each variable in shared ROC space and described relative sensitivity and specificity of one factor to the other.
We suspected papers would use the classical cross-sectional test accuracy study design. An alternative is the 'delayed verification' study design, i.e. where the AD-8 is performed at baseline and those without disease are prospectively followed up for development of incident dementia. No included studies took this approach, but if they had, we planned to use baseline (contemporaneous testing) data for our primary analysis.

### Investigations of heterogeneity

Heterogeneity is expected in DTA reviews, and 'traditional' measures of heterogeneity used in meta-analysis are not appropriate to DTA reviews.
The properties of a tool describe behaviour of the instrument under particular circumstances. We included all AD-8 studies in narrative review. We prespecified particular areas of potential heterogeneity, as follows.

#### AD-8 threshold score

We included data from all AD-8 threshold scores described in the included studies. Where data allowed, we collated scores at each cut-off score of test positivity, using meta-analytical techniques to create summary estimates of sensitivity and specificity.


Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

#### Healthcare setting

We suspected that healthcare setting would impact on properties of the test. Our primary analysis was across all settings; where data allowed, we performed sensitivity analysis comparing hospital-based (secondary care) settings and primary care or community settings.

#### Case mix

In the first instance we explored age, taking age over 65 years as a reference point. We anticipated that the majority of included participants in eligible studies would be aged over 65 years. The AD-8 may have different properties in younger cohorts and so we looked at age ranges within studies. We planned to grade studies with greater than 20% of included participants younger than 65 years as potentially unrepresentative. Where studies offered test accuracy at different age ranges, we chose the range that best represented our primary population of interest, namely older adults (aged over 65 years).

We anticipated that most studies would be of unselected adults, however we included studies that limited inclusion to a specific population, for example stroke survivors. For studies of selected populations, we assessed the validity of inclusion in a summary analysis with other included studies on a case-by-case basis.

We suspected that the majority of studies would focus on all-cause (undifferentiated) dementia or Alzheimer's disease, but included any dementia pathology in the review. We assessed generalisability according to age, comorbidity and dementia pathology, as part of our 'Risk of Bias' assessment. If data allowed, we planned to run sensitivity analyses removing studies with unrepresentative populations.

#### Criteria used to reach dementia diagnosis

We recorded the classification used (for example, DSM-IV; ICD-10) and the methodology used to reach dementia diagnosis (for example, individual assessment; group (consensus) assessment). We assessed the 'quality' of diagnosis at study level using the QUADAS-2 tool. If data allowed, we planned subgroup analyses comparing differing approaches to diagnosis.

#### Technical features of the testing strategy

Our focus was on language of assessment. We classified this as either English language or non-English-language tests. We performed subgroup analyses comparing the English-language AD-8 versus non-English-language AD-8.

### Sensitivity analyses

Where appropriate (i.e. if not already explored in our analyses of heterogeneity), and as data allowed, we explored the sensitivity of any summary accuracy estimates to other methodological aspects of the included studies. We performed a sensitivity analysis excluding a paper with a disease-specific patient group (those with delirium). We also performed a post-hoc sensitivity analysis excluding a paper whose test accuracy estimates were based on a more inclusive definition of dementia than was used in the other papers. We had planned to perform sensitivity analyses based on key aspects that could indicate risk of bias, such as nature of blinding and loss to follow-up, guided by the anchoring statements developed in our QUADAS-2 exercise. Due to the modest number of included studies, and the lack of papers deemed to have low risk of bias on all domains, we were unable to perform these analyses.

### Assessment of reporting bias

We did not investigate reporting bias because of current uncertainty about how it operates in test accuracy studies, and about the interpretation of existing analytical tools such as the funnel plot (van Enst 2014).

## RESULTS

### Results of the search

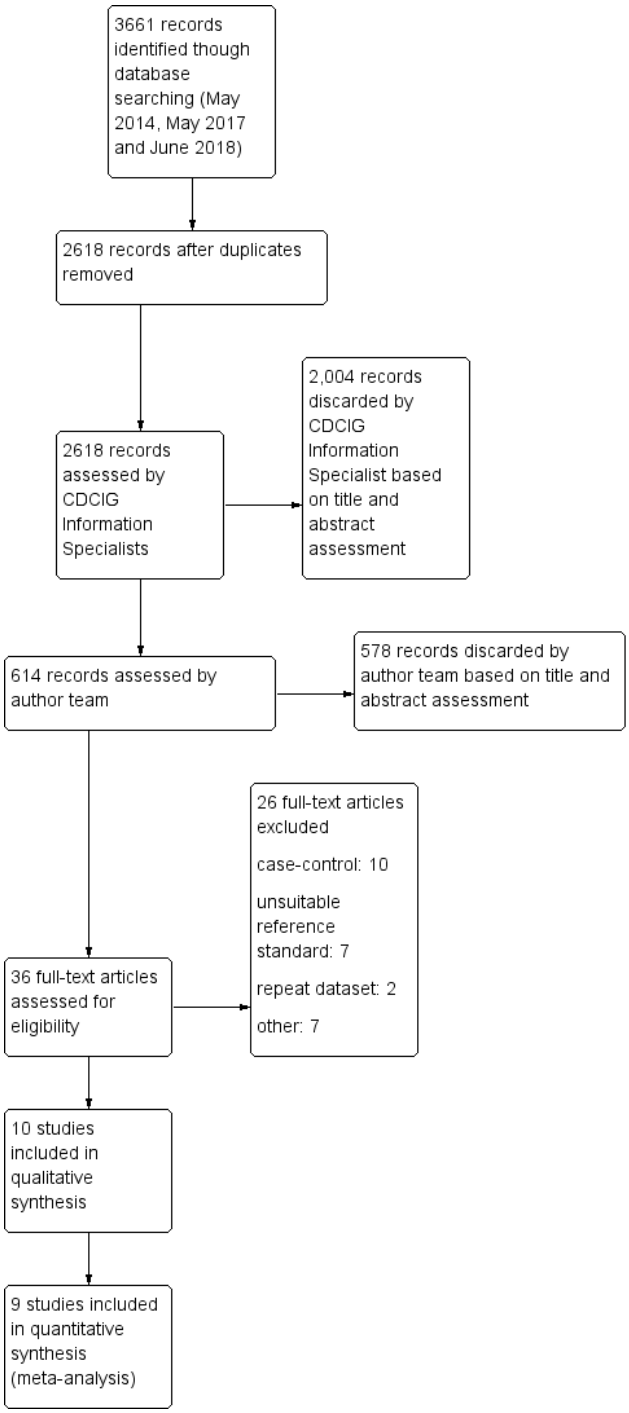
Our search resulted in 3661 records, from which we identified 36 full-text papers for eligibility. We included 10 studies, with a total of 4309 participants (Summary of findings 1).

We identified two studies which required translation (Munoz 2010; Pardo 2013). We contacted three authors to provide additional data, and one author provided this for incorporation in the review (see: Acknowledgements).

We excluded 26 papers. The most frequent reasons for exclusion were: case-control data (n = 10); unsuitable reference standard (n = 7) and repeat data set (n = 2) (see Characteristics of excluded studies). We excluded one study following repeated review and consultation with the review group Editor (Malmstrom 2009). The study subgroup, which could have been eligible, did not include any participants with a cognitive diagnosis that was compatible with our prespecified criteria for dementia (clinical diagnosis using a recognised clinical classification system and/or Clinical Dementia Rating (CDR) Scale score of one or more) As our review focuses on dementia, we chose to exclude this study. For a PRISMA flow diagram depiciting study selection, see Figure 1.


Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

**Figure 1. Study flow diagram.**





**Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.**

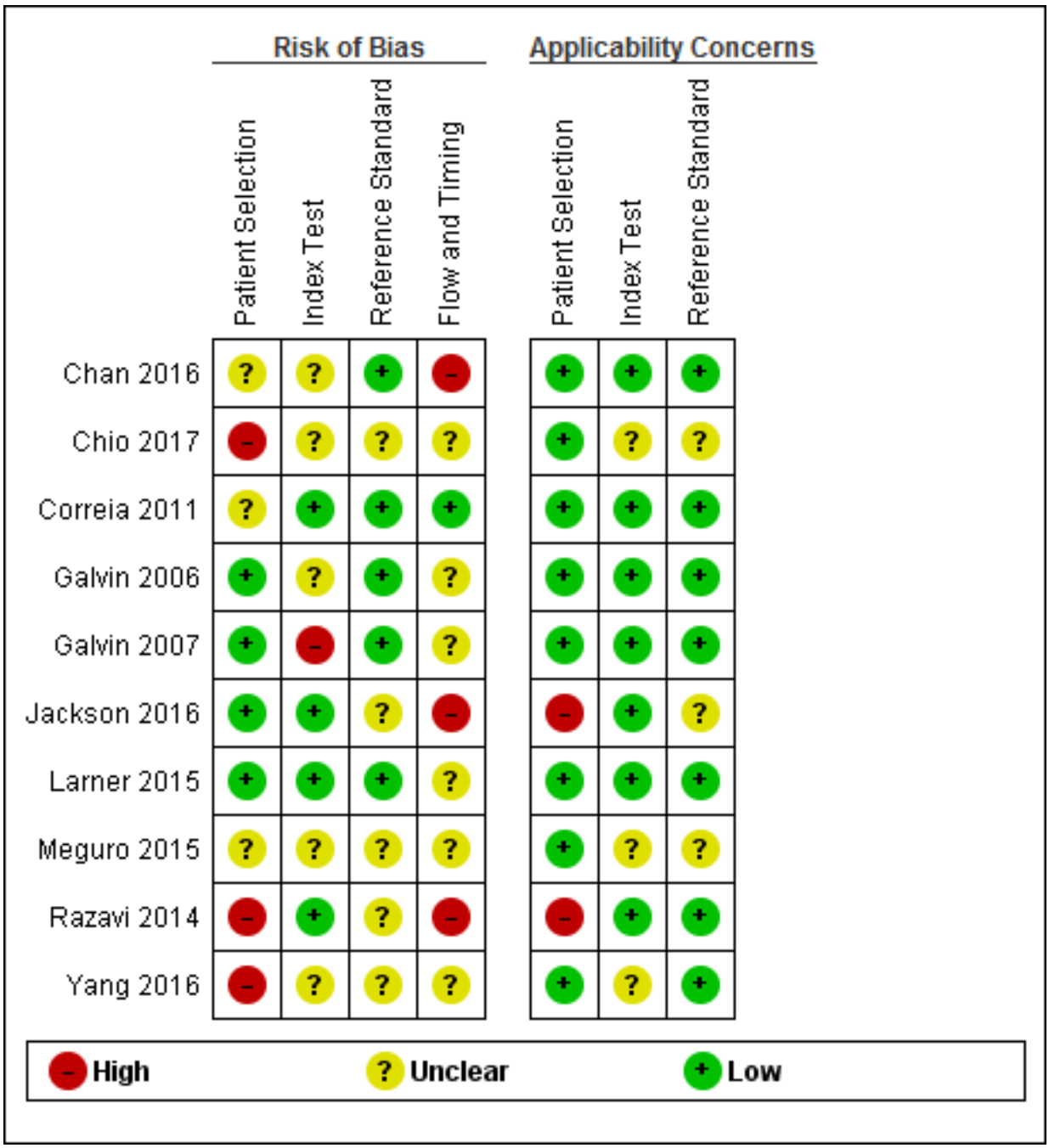
## Methodological quality of included studies

We described the risk of bias and applicability using the QUADAS-2 methodology (Appendix 8); our anchoring statements for the AD-8 are summarised in Appendix 9.

We did not rate any included study as being at low risk of bias for all the categories of QUADAS-2. For all domains there were papers where risk of bias was high or where reporting was insufficient, which necessitated an assessment of unclear risk of bias. There were particular issues around inappropriate exclusions of participants (for example, exclusion those who had dementia) and around patient flow (for example, substantial dropouts due to inability to complete test). We had fewer concerns around applicability. See: Figure 2.


Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

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





Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

## Findings

We have described the individual included studies in Characteristics of included studies and Table 1. We have also presented tabulated data for test accuracy by AD-8 threshold (Summary of findings 2), and by setting and language of administration (Summary of findings 3).

The total number of participants across the studies was 4309 (range: 109 to 2063). One study (n = 153 participants) identified cases of dementia (those with a CDR score of one or more), but did not provide the test accuracy data specific to this (Chio 2017). The focus of the paper was on identifying early cognitive impairment and so, while the study was eligible for inclusion in review, we could not include its data in the quantitative synthesis. We performed quantitative synthesis for nine studies, including a total of 4045 participants of whom 1107 (27%) had a clinical dementia diagnosis. This excludes 14 participants from Galvin 2006, who were missing a clinical diagnosis; one participant from Galvin 2007, who had no reference standard data; and 48 participants from Jackson 2016, plus 48 participants from Yang 2016, who were lacking complete follow-up assessment data. Dementia prevalence in the included studies varied from 12% to 90%.

Nine of the ten studies are included in the quantitative synthesis. Seven studies used a cut-off score of two to differentiate dementia from no dementia, and five reported data using a cut-off of three. We did not consider it was appropriate to combine data using different cut-off scores of test positivity. The cut-off scores used to define a 'positive' AD-8 included scores of one, two, three, and seven.

The included studies are international, including datasets from seven countries (Brazil, China, Japan, Singapore, Taiwan, the UK, the USA). Four different language versions of the AD-8 were used, although in two papers it was not clear if the authors had used a translated version.

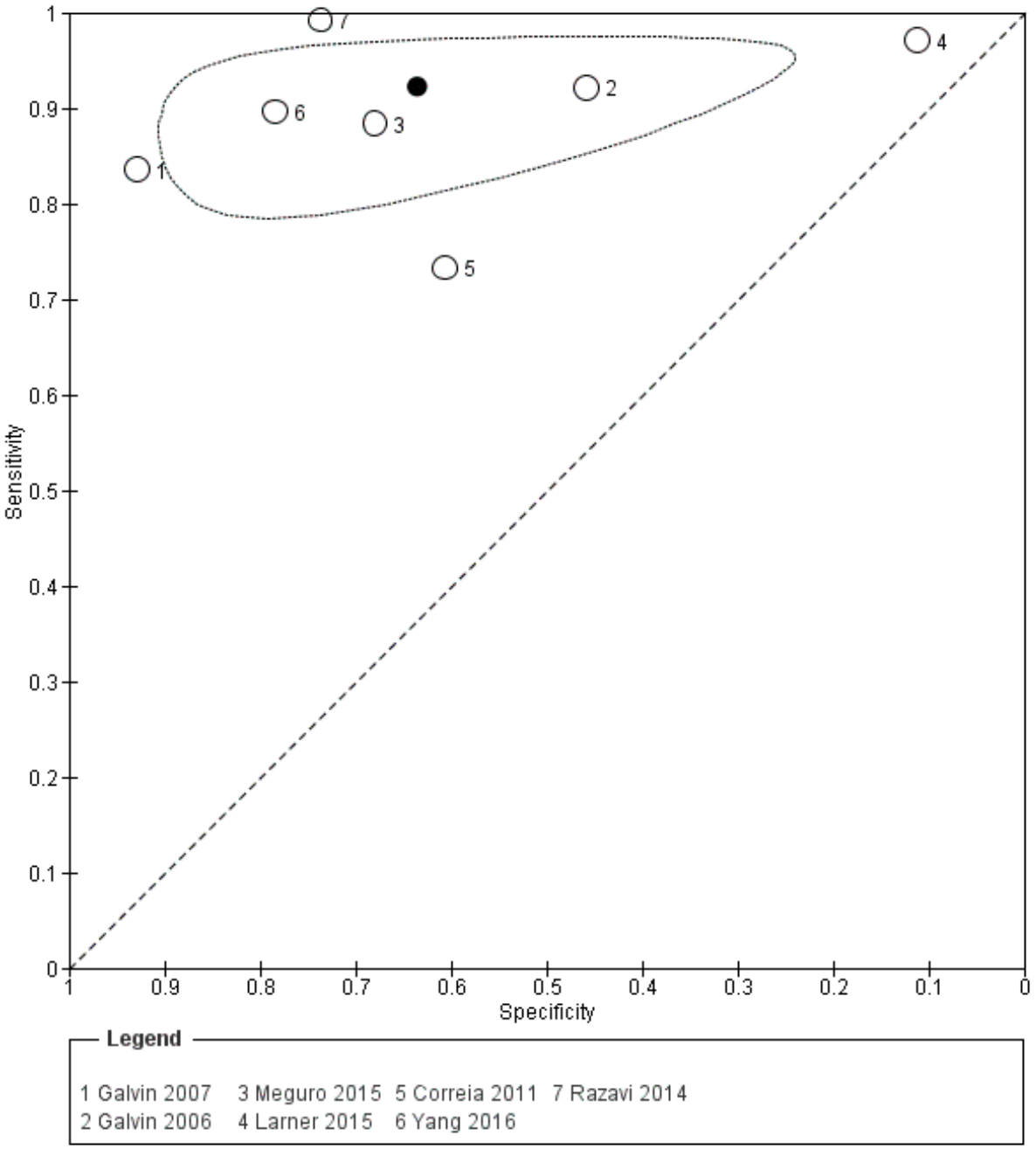
Using a cut-off score of two, there was a spread of sensitivity and specificity (sensitivity range: 73% to 99%; specificity range: 11% to 93%). Using a cut-off score of three, there was a spread of sensitivity and specificity (sensitivity range: 75% to 100%; specificity range: 40% to 91%). Table 1 provides a summary of the test accuracy for each study, presenting results at cut-off scores of two and three where available, ordered by dementia prevalence.

### AD-8 using an informant cut-off score of two

Seven studies (n = 3659) reported data using an AD-8 informant cut-off score of two (Correia 2011; Galvin 2006; Galvin 2007; Larner 2015; Meguro 2015; Razavi 2014; Yang 2016). Sensitivity was 0.92 (95% confidence interval (CI) 0.86 to 0.96); specificity was 0.64 (95% CI 0.39 to 0.82). The overall positive likelihood ratio was 2.53 (95% CI 1.38 to 4.64) and the negative likelihood ratio was 0.12 (95% CI 0.07 to 0.21). The summary describing test accuracy across the included studies in receiver operating characteristic (ROC) space is presented in Figure 3.


Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

**Figure 3. Summary ROC plot of AD-8 informant cut-off score 2. The dark point is a summary point, the other points individual studies; the broken line represents 95% confidence region.**





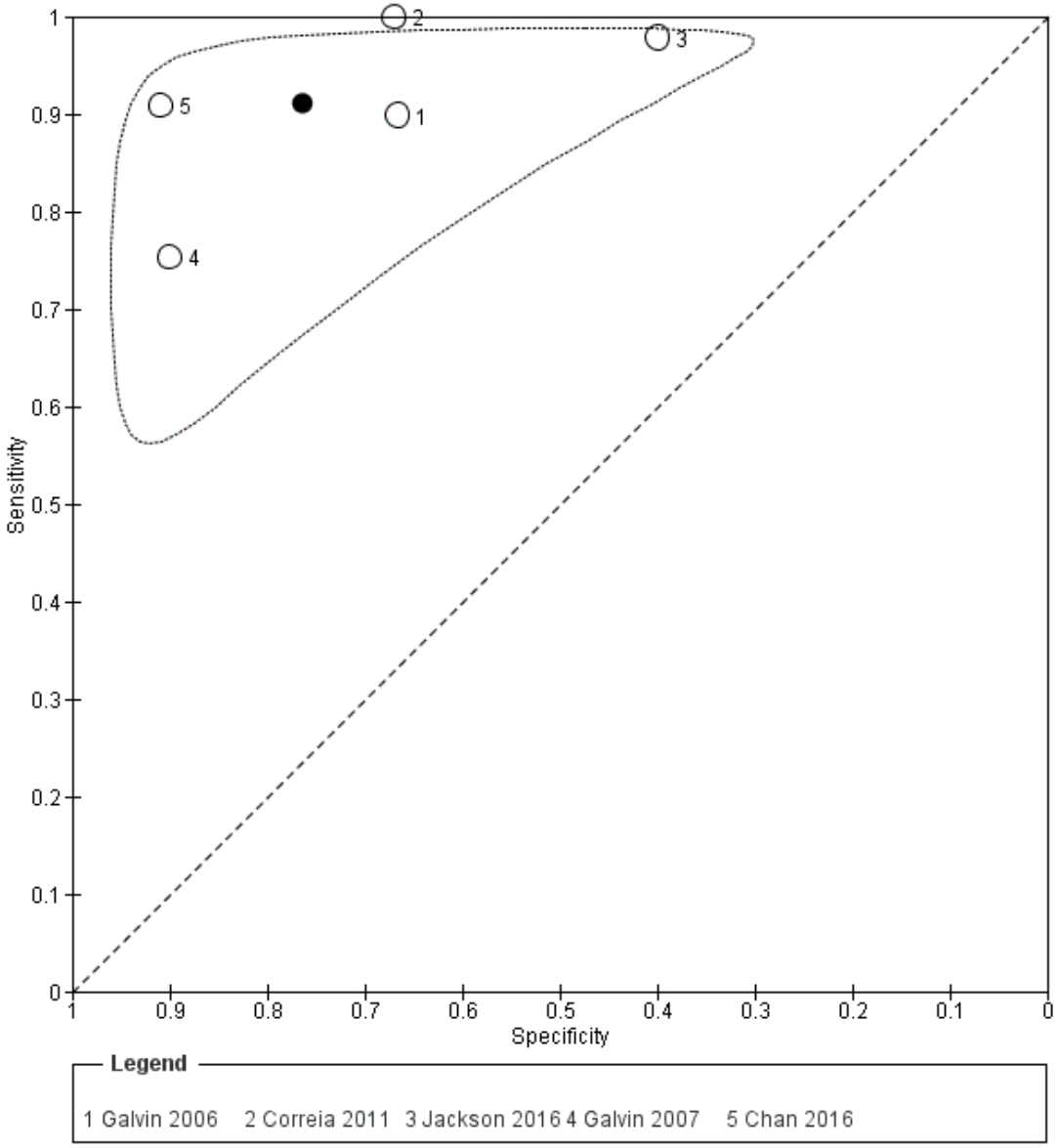
Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

### AD-8 using an informant cut-off score of three

Five studies (n = 1060) reported data using an AD-8 informant cut-off score of three (Chan 2016; Correia 2011; Galvin 2006; Galvin 2007; Jackson 2016). These gave a sensitivity of 0.91 (95% CI 0.80 to 0.96) and specificity of 0.76 (95% CI 0.57 to 0.89). The overall positive likelihood ratio was 3.86 (95% CI 2.03 to 7.34) and the negative likelihood ratio was 0.12 (95% CI 0.06 to 0.24). The summary describing test accuracy across the included studies in ROC space is presented in Figure 4 .


Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

**Figure 4. Summary ROC plot of AD-8 informant cut-off score 3. The dark point is a summary point, the other points individual studies; the broken line represents 95% confidence region.**





Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

## AD-8 test accuracy at other diagnostic thresholds for dementia

### AD-8 informant cut-off score of one

One study (n = 324) reported data using an AD-8 informant cut-off score of one (Galvin 2007). This had a sensitivity of 0.90 (95% CI 0.81 to 0.96) and specificity of 0.68 (95% CI 0.62 to 0.74).

### AD-8 informant cut-off score of seven

One study (n = 77) reported data using an AD-8 informant cut-off score of seven (Jackson 2016). This had a sensitivity of 0.83 (95% CI 0.69 to 0.92) and specificity of 0.90 (95% CI 0.73 to 0.98).

## Heterogeneity relating to setting

### Primary care

One study (n = 309) was conducted in a primary care setting, specifically the waiting areas of healthcare centres in Singapore (Chan 2016). Data were only reported using an AD-8 cut-off score of three. The sensitivity was 0.91 (95% CI 0.78 to 0.97) and specificity was 0.91 (95% CI 0.87 to 0.94).

### Community

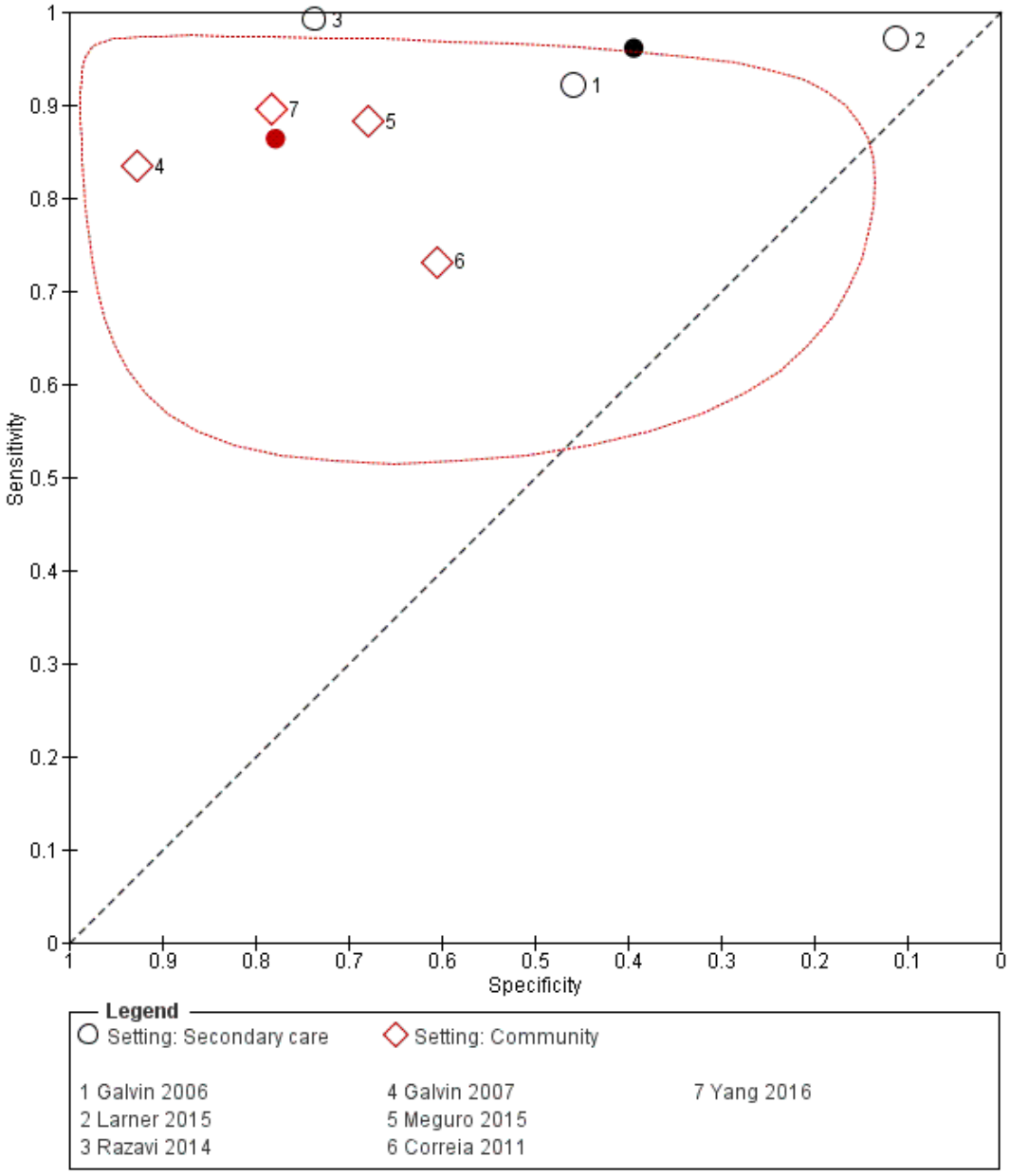
Four studies (n = 2937) were conducted in community settings (Correia 2011; Galvin 2007; Meguro 2015; Yang 2016).
Data were available at a cut-off score of two for all four studies. The sensitivity was 0.86 (95% CI 0.78 to 0.92) and specificity was 0.78 (95% CI 0.57 to 0.90). The overall positive likelihood ratio was 3.89 (95% CI 1.76 to 8.59) and the negative likelihood ratio was 0.18 (95% CI 0.10 to 0.31). The dementia prevalence ranged from 12% to 23%.

### Secondary care

Four studies (n = 716) were conducted in secondary care settings. This included three in outpatient neurology, cognitive function or memory clinics (Galvin 2006; Larner 2015; Razavi 2014), and one study of hospitalised individuals with delirium (Jackson 2016).
Data were available for three of the studies at a cut-off score of two (excluding Jackson 2016, which only reported data using a cut-off score of three). The sensitivity was 0.96 (95% CI 0.92 to 0.98) and specificity was 0.39 (95% CI 0.16 to 0.69). The overall positive likelihood ratio was 1.58 (95% CI 0.98 to 2.57) and the negative likelihood ratio was 0.10 (95% CI 0.03 to 0.33). The dementia prevalence ranged from 33% to 90%.
Comparing pooled test accuracy between the two settings, the AD-8 has a higher relative sensitivity (1.11, 95% CI 1.02 to 1.21), but lower relative specificity (0.51, 95% CI 0.23 to 1.09), in secondary care compared to community care settings (Figure 5).


Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

**Figure 5. Summary ROC plot of AD-8 informant cut-off score 2 comparing test accuracy in community versus secondary care settings. Red diamonds represent community studies and black circles those in secondary care. The dark point is a summary point, the other points individual studies; the broken line represents 95% confidence region. No confidence region calculable for secondary care due to lack of available data.**





Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

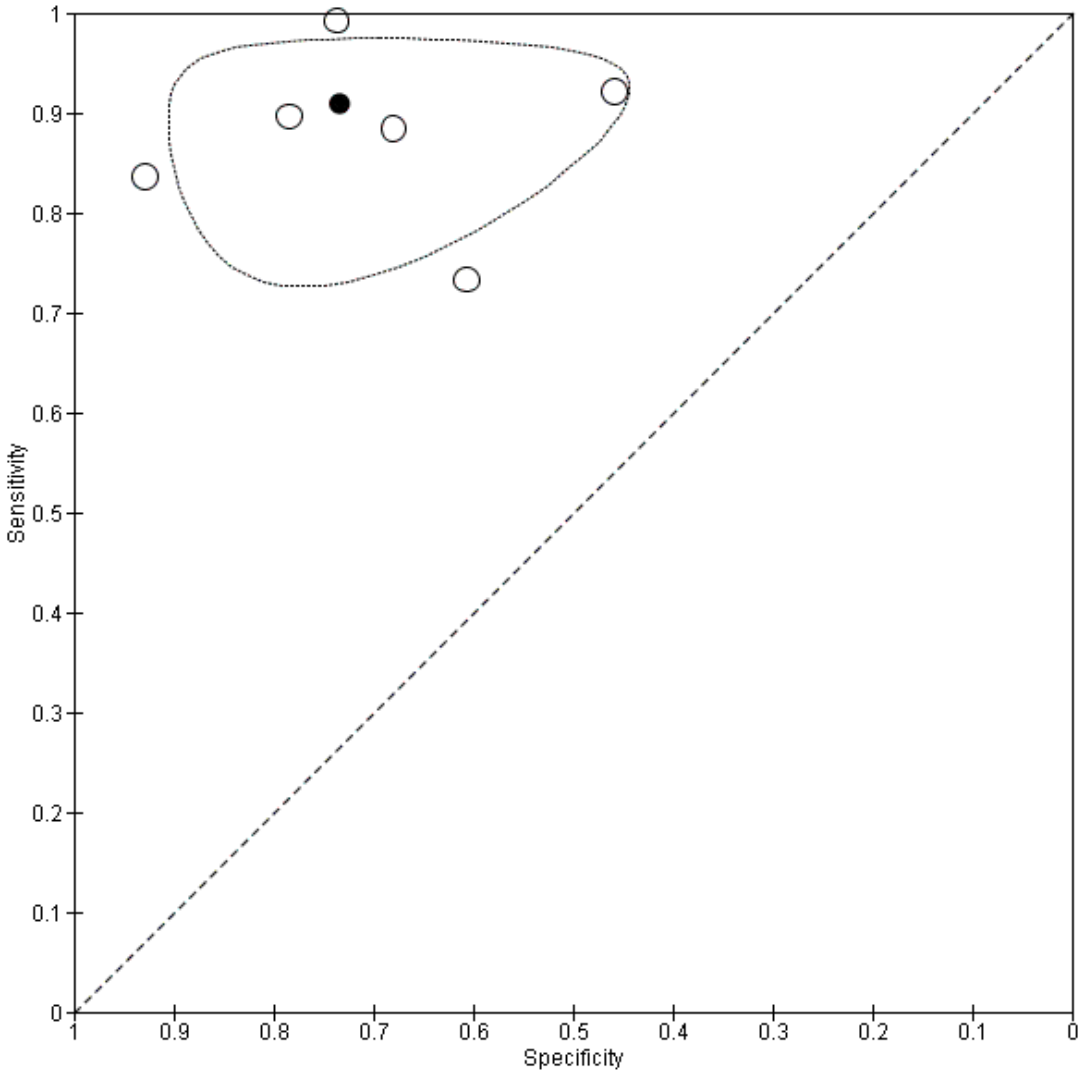
## Other sources of heterogeneity and sensitivity analyses

### Age

We performed a sensitivity analysis removing the study which included participants with a low mean or median age (less than 65 years) (Larner 2015). Using a cut-off score of two, test accuracy was similar after exclusion of this study, with an improvement in the specficity of the AD-8: sensitivity was 0.91 (95% CI 0.83 to 0.95); specificity was 0.73 (95% CI 0.59 to 0.84); the positive likelihood ratio was 3.42 (95% CI 2.12 to 5.49); and the negative likelihood ratio was 0.12 (95% CI 0.07 to 0.23) (Figure 6).


Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

**Figure 6. Summary ROC plot of sensitivity analysis removing low average age. The dark point is a summary point, the other points individual studies; the broken line represents 95% confidence region.**




#### Disease-specific populations

One study included a disease-specific population, individuals with a diagnosis of delirium (Jackson 2016). We performed a sensitivity analysis, removing this study from all the others using a cut-off score of three. Results showed a lower sensitivity and higher specificity than the pooled analysis result. Sensitivity was 0.88 (95% CI 0.78 to 0.94) and specificity was 0.83 (95% CI 0.69 to 0.91). The overall positive likelihood ratio was 5.04 (95% CI 2.82 to 9.01) and the negative likelihood ratio was 0.15 (95% CI 0.08 to 0.26).

#### Reference standard dementia diagnostic method

Three studies used the CDR as their reference standard (Chan 2016; Galvin 2007; Meguro 2015); three used the DSM-IV criteria (Correia 2011; Jackson 2016; Larner 2015); two studies used various criteria for dementia diagnosis (Galvin 2006; Razavi 2014); and one used the National Institute on Aging and Alzheimer's Association (NIA-AA) criteria (Yang 2016). Quanti-


Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

tative comparisons based on the diagnostic criteria used were not possible due to the small numbers of studies using each method.
One study took a more inclusive definition of dementia, including those with a CDR score of 0.5 in their summary test accuracy data (Galvin 2007). For the purposes of analysis, we used the total numbers with a clinical diagnosis of dementia (CDR score greater than or equal to one), in common with the approach used in all other studies and the CDR definitions (Morris 1993). We used the summary sensitivity and specificity data presented in the paper, however the prevalence of dementia was lower than that quoted due to our exclusion of those 102 participants scoring 0.5 from the disease-positive group (54% quoted in study reduced to 23%). As such, we performed a post-hoc sensitivity analysis to check the effects of including the data from this study on our summary estimates at cut-off scores of two and three.
Using a cut-off score of two, removing data from Galvin 2007 resulted in similar sensitivity (0.93, 95% CI 0.86 to 0.97); but lower specificity (0.56, 95% CI 0.34 to 0.76); a positive likelihood ratio of 2.10 (95% CI 1.28 to 3.43) and negative likelihood ratio of 0.12 (95% CI 0.06 to 0.25). Using a cut-off score of three, removing data from Galvin 2007 resulted in similar sensitivity (0.92, 95% CI 0.80 to 97); but lower specificity (0.71, 95% CI 0.48 to 0.87), with wider confidence intervals. The overall positive likelihood ratio was 3.19 (95% CI 1.61 to 6.33) and the negative likelihood ratio was 0.11 (95% CI 0.04 to 0.28).

#### Disease subtype

No studies considered a specific dementia subtype to facilitate sensitivity analysis based on this.

#### Language of AD-8 administration

The English-language version of the AD-8 was used in five studies (Galvin 2006; Galvin 2007; Jackson 2016; Larner 2015; Razavi 2014). Three studies named which specific non-English-language versions were used (Chan 2016; Correia 2011; Meguro 2015). In two studies the language of AD-8 administration was unclear (Chio 2017; Yang 2016).
The summary test accuracy of the English-language version, used at a cut-off score of two (n = 963) was calculated, resulting in sensitivity of 0.95 (95% CI 0.87 to 0.98); and specificity of 0.59 (95% CI 0.21 to 0.89). The overall positive likelihood ratio was 2.30 (95% CI 0.87 to 6.07) and the negative likelihood ratio was 0.08 (95% CI 0.03 to 0.22). Only three studies using a cut-off score of three used the English-language version, and thus we could not pool the data.


Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

## ADDITIONAL SUMMARY OF FINDINGS *[Explanation]*

| What is the accuracy of the AD-8 test for the detection of dementia when differing thresholds are used to define AD-8 positive cases? | |
|---|---|
| **Population** | Adults whose informant completed an AD-8 questionnaire, with no restrictions on the case mix of recruited participants |
| **Setting** | Primary care, community and secondary care |
| **Index test** | AD-8 informant-based questionnaire |
| **Reference standard** | Clinical diagnosis of dementia made using any recognised classification system |
| **Studies** | We included cross-sectional but not case-control studies |

| Test | Summary accuracy (95% CI) | No. of participants (studies) | Dementia prevalence | Implications, quality and comments |
|---|---|---|---|---|
| **AD-8 cut-off score 1** | **Sensitivity:** 0.90 (0.81 to 0.96)<br>**Specificity:** 0.68 (0.62 to 0.74) | n = 324<br>(1 study) | n = 73<br>(60%) | The majority of studies used the informant cut-off score of 2 to define AD-8 test positivity<br>As the cut-off score is raised, there is an increase in specificity. Sensitivity is similar at a cut-off score of 2 or 3, but falls as the cut-off score to define test positivity rises<br>False positives (individuals wrongly identified as having possible dementia and referred for further investigations) and false negatives (individuals with dementia who are not identified and not referred for specialist assessment and management) can both be associated with harms. If used as a screening test (i.e. an initial assessment, which will be followed by further testing) then a pat- |

Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

| | | | | |
|---|---|---|---|---|
| | | | | tern of higher sensitivity may be preferable<br>The dementia prevalence within the included studies was highly varied (12% to 90%) and may reflect that for each cut-off score, data were included from all settings in which the AD-8 has been used. These analyses need to be interpreted with caution as the properties of the test may vary by setting and the desired balance of sensitivity and specificity will vary by purpose of testing |
| **AD-8 cut-off score 2** | **Sensitivity:** 0.92 (0.86 to 0.96)<br>**Specificity:** 0.64 (0.39 to 0.82)<br>**+LR:** 2.53 (1.38 to 4.64)<br>**-ve LR:** 0.12 (0.07 to 0.21) | n = 3659 (7 studies) | n = 1016 (28% range: 12% to 90%) | |
| **AD-8 cut-off score 3** | **Sensitivity:** 0.91 (0.80 to 0.96)<br>**Specificity:** 0.76 (0.57 to 0.89)<br>**+LR:** 3.86 (2.03 to 7.34)<br>**-ve LR:** 0.12 (0.06 to 0.24) | n = 1060 (5 studies) | n = 396 (37% range: 14% to 90%) | |
| **AD-8 cut-off score 7** | **Sensitivity:** 0.83 (0.69 to 0.92)<br>**Specificity:** 0.90 (0.73 to 0.98) | n = 77 (1 study) | n = 47 (61%) | |

95% CI: 95% confidence interval
+LR: positive likelihood ratio
-LR: negative likelihood ratio

Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

| What is the accuracy of the AD-8 test for detection of dementia in different settings and using different languages of administration? | | | | |
|---|---|---|---|---|
| **Population** | Adults whose informant completed an AD-8 questionnaire, with no restrictions on the case mix of recruited participants | | | |
| **Setting** | Primary care, community and secondary care | | | |
| **Index test** | AD-8 informant-based questionnaire using a cut-off score of 2 to define positivity | | | |
| **Reference standard** | Clinical diagnosis of dementia made using any recognised classification system | | | |
| **Studies** | We included cross-sectional but not case-control studies | | | |
| **Comparative analyses** | | | | |
| **Test** | **No. of participants (studies)** | **Dementia prevalence** | **Findings** | **Implications** |
| **Community versus primary care versus secondary care setting** | Total n = 4045 (9)<br>n = 2937 (4) Community<br>n= 309 (1) Primary care<br>n = 716 (4) Secondary care | Total n = 1107 (27%)<br>Community n = 601 (20%)<br>Primary n = 44 (14%)<br>Secondary care n = 462 (65%) | Using a cut-off score of 2, the AD-8 has a higher relative sensitivity (1.11, 95% CI 1.02 to 1.21) in secondary care compared to community care settings<br>Using a cut-off score of 2, the AD-8 has a lower relative specificity (0.51, 95% CI 0.23 to 1.09) in secondary care compared to community care settings<br>Only one study used the AD-8 in primary care settings; data were at a cut-off score of 3<br>Two studies used AD-8 in community and two studies used AD-8 in secondary care settings, with a cut-off score of 3. The numbers were insufficient to allow formal quantitative synthesis | Dementia prevalence in secondary care is more than double that seen in the community<br>The AD-8 has a higher sensitivity, but poorer specificity, in secondary care settings |

Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

| | | | | |
|---|---|---|---|---|
| **English language** | Total: n = 1040 (5) English language | Total n = 535 (51%) | The test accuracy of English language administration of the AD-8, at a cut-off score of 2, has a higher sensitivity and lower specificity than summary estimates for all studies at a cut-off score of 2. Estimates for specificity were imprecise due to the small number of studies included<br>Sensitivity: 0.95 (95% CI 0.87 to 0.98)<br>Specificity: 0.59 (95% CI 0.21 to 0.89)<br>+LR: 2.30 (95% CI 0.87 to 6.07)<br>-LR: 0.08 (95% CI 0.03 to 0.22)<br>We were unable to assess English language at a cut-off score of 3 due to lack of available data at this cut-off score<br>We were unable to assess the non-English-language group due to the small numbers reporting language of administration clearly | There is a need to improve clarity of reporting around language of test administration to evaluate diagnostic accuracy of non-English-language versions of the AD-8 |

95% CI: 95% confidence interval
+LR: positive likelihood ratio
-LR: negative likelihood ratio

# DISCUSSION

## Summary of main results

This review summarises the test accuracy of the AD-8 for dementia, across the range of settings in which it may be applied. Quantitative synthesis from seven of the identified studies suggests a summary sensitivity of 0.92 and specificity of 0.64 when an AD-8 cut-off score of two is used for the diagnosis of dementia. These results are international and include a population of 3659 participants. Using an AD-8 cut-off score of three to diagnose dementia achieves a similar sensitivity of 0.91, and higher specificity of 0.76, based on data from five studies and 1060 participants. There was significant clinical and methodological heterogeneity between the included studies, regarding their study populations, and dementia prevalence (which varied from 12% to 90%).

We found the usual pattern of compromise between sensitivity and specificity at extremes of scoring (for example, an AD-8 score of one versus an AD-8 score of seven). At the thresholds commonly used in practice - cut-off scores of two or three - differences were less obvious. Our summary estimates of test properties at these scores have relatively large confidence intervals, particularly for specificity; this reflects the heterogeneity in the included studies. We can conclude that at usual threshold scores, the AD-8 favours sensitivity over specificity. A cut-off of two was used most commonly, which is consistent with the original development of the tool (Galvin 2005).

The setting of test administration affected the test accuracy estimates produced. Using a cut-off score of two, the AD-8 has a higher relative sensitivity (1.11), but lower relative specificity (0.51) in secondary care compared to community care settings. This suggests use of the AD-8 in secondary care will produce more false positive results in those referred for specialist assessment, but fewer false negatives than when used in community cohorts. Arguably, in those referred for a specialist assessment and having already had some form of cognitive testing, specificity is important. In this case, our data would suggest that the AD-8 is less suited as a diagnostic aid in a secondary-care memory-clinic setting. The data at a cut-off score of three could not be pooled due to the small number of studies in each setting group; similarly, only one study examined the AD-8 in primary care.

Removal of the study with a low average age (prespecified as less than 65 years) resulted in similar estimates of sensitivity (0.91), but improved specificity (0.73). The excluded study still had a mean age of greater than 60 years.

One study included only adults diagnosed as having delirium at the time of presentation (Jackson 2016). Delirium is more likely to occur in adults with a diagnosis of dementia (Inouye 2014), so we performed a sensitivity analysis removing this from the pooled estimate at a cut-off score of three. Removing this study resulted in lower sensitivity and higher specificity than the summary estimate.

Quantitative analysis comparing the accuracy of different reference standard classifications was not possible due to the range of different approaches used.

Analysis comparing the accuracy of different language versions of the AD-8 was limited due to the unclear description of the version used in two studies, and the lack of data available at a cut-off score of three. The summary test accuracy of the English-language version was comparable with the overall pooled estimate.

## Strengths and weaknesses of the review

### Strengths and weaknesses of included studies

Quality assessment using the QUADAS-2 tool identified risks of bias and applicability concerns across all the included studies; we assessed no study as having low risk on all domains. Specific concerns are discussed in Characteristics of included studies and are graphically presented in Figure 2. These included the potential for selection bias introduced by exclusion criteria around comorbidities or unclear exclusion of participants; and issues around the conduct of index tests and suitability of informants, based on length of relationship, age, etc. There was a lack of data from primary care settings and acute hospital inpatient settings, which may be targets for opportunistic assessment using brief instruments.

### Strengths and weaknesses of review process

This review has been conducted using a rigorous search strategy, developed and managed by an experienced Information Specialist in Cochrane. The search was determined using a structured approach. No restrictions were placed based on date or language, and translation services were used where necessary to facilitate the evaluation and inclusion of non-English-language publications. Our correspondence with study authors was successful; we obtained additional data, clarification around aspects that were unclear in the original publication, and additional papers evaluating the AD-8 (Acknowledgements). Quality assessment was informed by use of dementia-specific QUADAS-2 anchoring statements, which were developed for this suite of reviews by experts in the field (Davis 2013).

Our review question was focused on evaluating the diagnostic accuracy of the AD-8 for dementia, thus we excluded data evaluating the AD-8 for diagnosing mild cognitive impairment from our quantitative analyses. We conducted prespecified sensitivity analyses around setting, studies of low average age, disease-specific populations, and language of administration.

## Comparisons with previous research

The AD-8 is one of two informant-based questionnaires whose test accuracy for dementia has been assessed by the Cochrane Dementia and Cognitive Improvement Group; the other is the Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE). The IQCODE was found to have a similar test accuracy


Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

to the AD-8 in community settings (sensitivity 0.80; specificity 0.84; 10 studies; 2644 individuals) (Quinn 2014). In secondary care settings, data were further divided into those evaluating the IQCODE in specialist memory settings and non-specialist settings (Harrison 2015). In keeping with the current review, the majority of the data in the IQCODE review were from specialist settings and test accuracy estimates showed a higher specificity but comparable sensitivity (sensitivity 0.90; specificity 0.54; six studies; 1352 individuals) to the overall secondary care estimate for the AD-8. The relative lack of AD-8 data from primary care settings is consistent with the review of the IQCODE in primary care, which identified only one relevant study in this setting (Harrison 2014).

The AD-8 was designed to differentiate cognitively normal individuals from those with mild dementia as part of clinical assessment in a primary care setting, derived using data from a longitudinal community cohort study (Galvin 2005). Prevalence of dementia was recognised to affect the test accuracy of the AD-8, even at its inception (Galvin 2005). In the original validation study, the inclusion of 56 individuals with mild to severe dementia raised prevalence from 38% to 53%, improved sensitivity to 0.85 from 0.74 without a change to specificity of 0.86 (Galvin 2005). The pooled estimates reported in this review of all-cause/all-severity dementia have a higher sensitivity, but poorer specificity, than described in the first AD-8 report.

During the production of this review, another author group published a systematic review of AD-8 test accuracy (Chen 2018). Although the inclusion and exclusion criteria were similar, the review did not include some studies that we included in our review. It is not clear if these titles were considered and then rejected, or not considered. For using the AD-8 to differentiate dementia from normal cognition, the summary test accuracy data presented in the review by Chen and colleagues were similar to our results, with high sensitivity (0.91, 95% CI 0.89 to 0.92) and lower specificity (0.78, 95% CI 0.76 to 0.80) (Chen 2018). No assessment of internal or external validity was presented.

We excluded case-control designs, in common with the other reviews in the suite, due to the potential to over-estimate test accuracy (Rutjes 2006; Kohn 2013). The case-control studies excluded from this review have been analysed, and a summary sensitivity of 0.88 and specificity of 0.76 was reported (Anwer 2017). The effect of combining these studies with the studies included in our review has not been evaluated, however meta-regression analysis did not show evidence of a systematic difference in the accuracy reported for the AD-8 between non-case-control and case-control designs (Anwer 2017).

### Applicability of findings to the review question

This review focused on collating all available cross-sectional data on the diagnostic accuracy of the AD-8 to identify dementia. Our search strategy was robust and we believe we have evaluated all the evidence relevant to this question. Summary estimates of AD-8 test accuracy are provided where sufficient appropriate data allowed, and findings are reported with consideration of study quality, risks of bias and applicability associated with study conduct.

## AUTHORS' CONCLUSIONS

### Implications for practice

The AD-8 was developed as a tool to help inform the assessment of an individual, rather than being a diagnostic test for dementia in itself. The data from this review support the use of the AD-8 (informant cut-off score of two) as a tool to help identify those who may benefit from further specialist assessment. The setting where the test is administered affects the diagnostic accuracy, though we could not evaluate estimates of performance in primary care using the available data (albeit this may be the setting where the tool is used most often).

Test accuracy is only one metric to consider when using a test in clinical practice. For a test that is likely to be used predominantly in older adults, consideration must be given to feasibility, acceptability and test burden (Lees 2017). The feasibility of using the AD-8 in the range of settings reported cannot be fully determined from the included studies, as there was a lack of data on the handling of indeterminate results or methods of assessing those whose informant was not present. Lack of an available informant was a common reason for exclusion, and one study excluded those whose informant was considered 'unreliable' due to age or perceived knowledge of the individual (Larner 2015). As such, there are insufficient data on the acceptability and feasibility of AD-8 use in practice to offer any useful comment on these properties.

In the context of a large and increasing variety of available cognitive assessment tools (Harrison 2016), the important clinical question is around which test to use for a particular patient. Our review only goes some way to answering this question. We offer data on a single test, rather than formal comparative analyses with other tests. Techniques to allow comparisons across a network of differing tests with a common reference standard are emerging and may be a useful approach for future reviews (Owen 2018).

There is no ideal value for sensitivity or specificity in clinical practice. The optimal 'trade-off' between sensitivity and specificity will depend on the purpose of testing and the implications of erroneous test results. With a pattern of high sensitivity and lower specificity, the AD-8 is best suited for use as an initial screening or triage tool, with a view to selecting those who need more detailed assessment. In this scenario, the higher 'false positives' will go on to have additional testing and an incorrect label of dementia should be avoided. This form of cognitive screening is the usual purpose for which the AD-8 is employed in clinical practice and our review


Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

data would support this approach. It is interesting that the pattern of high sensitivity and lower specificity seems more pronounced when the AD-8 is used in a secondary care (specialist memory service) setting. In this setting, where there is a high prevalence of cognitive impairment and the population is likely to have already had some form of cognitive assessment before referral, the utility of a high-sensitivity screening tool is questionable.

### Implications for research

In this review, we observed differences in test accuracy when the AD-8 was used in different healthcare settings. It is therefore important that future studies stratify results by the setting of recruitment, and that researchers are aware of the differences in dementia prevalence which may affect test performance. We only identified one study that was conducted in each of the primary care and acute hospital settings. These settings may provide opportunities for the use of brief cognitive tests such as the AD-8, and thus diagnostic accuracy should be established.

This review considered only the use of the AD-8 as a single test in isolation. In practice, it is likely that testing is done sequentially, informed by results. This review is one in a series examining the test accuracy of other dementia tests, including the Informant Questionnaire on Cognitive Decline in the Elderly (Harrison 2015), and Montreal Cognitive Assessment (Davis 2015). It would be helpful to consider the testing pathway and determine which tests have greatest diagnostic utility at which stage. Tests may not be used sequentially and informant assessments such as the AD-8 are often completed at the same time as direct-to-patient cognitive tests. How these results should be combined to give the greatest accuracy is a topic that requires further research.

In applying our assessments of internal and external validity, use of the QUADAS-2 tool was complicated by poor reporting within the included studies. Poor reporting is a recognised issue in test accuracy studies (Davis 2013). Specific guidance for reporting test accuracy with a dementia focus is available, and we would recommend that future primary test accuracy research makes use of this resource (Noel-Storr 2014).

The ultimate goal of interventions to improve rates of dementia diagnosis is to enable individuals to access specialist services, resources and support. Historical modelling suggested early diagnosis could improve the quality of life of those diagnosed, and reduce the need for care-home admission (Banerjee 2009). Longitudinal data from the UK identified evidence of increased rates of dementia diagnosis and reductions in antipsychotic prescribing following introduction of the National Dementia Strategy (Donegan 2017). A challenge is therefore for researchers to identify if those identified as having possible dementia and referred for formal diagnosis derive benefit from the impact of this diagnosis through the care and treatment they subsequently receive. Ultimately, the design of studies looking at assessment tools such as the AD-8 may move away from the classical index test versus gold-standard paradigm to a model that looks at test-treatment-outcome pathways (Takwoingi 2018).

## ACKNOWLEDGEMENTS

We are grateful to Dr Andrew J Larner (Cognitive Function Clinic, Walton Centre for Neurology and Neurosurgery, Liverpool, UK) for sharing data on his AD-8 work and alerting us to other potentially relevant papers.

We are grateful to Robin WM Vernooij (Iberoamerican Cochrane Centre) who assisted with translations.

## REFERENCES

### References to studies included in this review

**Chan 2016** ***{published data only}***
Chan QL, Xu X, Shaik MA, Chong SST, Hui RJY, Chen CLH, et al. Clinical utility of the Informant AD8 as a dementia case finding instrument in Primary healthcare. *Journal of Alzheimer's Disease* 2016;**1**:121–7.

**Chio 2017** ***{published data only}***
Chio OI, Yip PK, Liu YC, Chen LH, Wang PC, Tsai TH, et al. Detection of cognitive impairment using self-rated AD8 and informant-reported AD8. *Journal of the Formosan Medical Association/Taiwan yi zhi.* 2017;**117**:42–7.

**Correia 2011** ***{published data only}***
Correia CC, Lima F, Junqueira F, Campos MS, Bastos O, Petribu K, et al. AD8-Brazil: Cross-cultural validation of the Ascertaining Dementia Interview in Portuguese. *Journal of Alzheimer's Disease* 2011;**27**:177–85.

**Galvin 2006** ***{published data only}***
Galvin JE, Roe CM, Xiong X, Morris JC. Validity and reliability of the AD8 informant interview in dementia. *Neurology* 2006;**67**:1942–8.

**Galvin 2007** ***{published data only}***
Galvin JE, Roe CM, Coats MA, Morris JC. Patient's rating of cogntive ability. *Archives of Neurology* 2007;**64**:725–30.

**Jackson 2016** ***{published data only}***
Jackson TA, MacLullich AMJ, Gladman JRF, Lord JM, Sheehan B. Diagnostic test accuracy of informant-based tools to diagnose dementia in older hospital patients with delirium: a prospective cohort study. *Age and Ageing* 2016; **45**:505–11.


**Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.**

**Larner 2015** *{published data only}*
Larner AJ. AD8 Informant Questionnaire for cognitive impairment: pragmatic diagnostic test accuracy study. *Journal of Geriatric Psychiatry and Neurology* 2015;**28**(3): 198–202.

**Meguro 2015** *{published data only}*
Meguro K, Kasai M, Nakamura K. Reliability and validity of the Japanese version of the AD8. *Nihon Ronen Igakkai zasshi [Japanese Journal of Geriatrics]* 2015;**52**(1):61–70.

**Razavi 2014** *{published data only}*
Razavi M, Tolea MI, Margrett J, Martin P, Oakland A, Tscholl DW, et al. Comparison of 2 informant questionnaire screening tools for dementia and mild cognitive impairment AD8 and IQCODE. *Alzheimer Disease and Associated Disorders* 2014;**28**:156–61.

**Yang 2016** *{published data only}*
Yang L, Yan J, Jin X, Jin Y, Yu W, Xu S, et al. Screening for dementia in older adults: comparison of Mini-Mental State Examination, Mini-Cog, Clock Drawing Test and AD8. *PLOS One* 2016;**12**(11):e0168949.

## References to studies excluded from this review

**Basalo 2017** *{published data only}*
Basalo MMG, Fernandez MC, Quintana MO, Rojas JI, Basalo MJG, Bogliotti E, et al. ALBA Screening Instrument (ASI): a brief screening tool for Lewy Body Dementia. *Archives of Gerontology and Geriatrics* 2017;**70**:67–75.

**Blanco 2016** *{published data only}*
Blanco R, Roman F, Iturry M, Leis A, Russo MJ, Bartoloni L. AD8-Argentina questionnaire to detect dementia in primary health care. *Neurologia Argentina* 2016;**8**:231–6.

**Carnero 2013** *{published data only}*
Carnero Pardo C, de la Vega Cotarelo R, Lopez Alcalde S, Martos Aparicio C, Vilchez Carrillo R, Mora Gavilan E, et al. Assessing the diagnostic accuracy (DA) of the Spanish version of the informant-based AD8 questionnaire. *Neurologia* 2013;**28**:88–94.

**Carpenter 2011** *{published data only}*
Carpenter CR, DesPain B, Keeling TN, Shah M, Rothenberger M. The Six-Item Screener and AD8 for the detection of cognitive impairment in Geriatric Emergency Department patients. *Annals of Emergency Medicine* 2011; **57**:653–61.

**Carpenter 2011ii** *{published data only}*
Carpenter CR, Bassett ER, Fischer GM, Shirshekan J, Galvin JE, Morris JC. Four sensitive screening tools to detect cognitive dysfunction in Geriatric Emergency Department patients: Brief Alzheimer's Screen, Short Blessed Test, Ottawa 3DY, and Caregiver-completed AD8. *Academic Emergency Medicine* 2011;**18**:374–84.

**Chen 2017** *{published data only}*
Chen S-F, Liu M-H, Chen N-C, Horng H-D, Tsao W-L, Chang C-C, et al. Educational effects on Ascertain Dementia 8-item informant questionnaire to detect dementia in the Taiwanese population. *International Psychogeriatrics* 2017;**1**:9.

**Chen 2018** *{published data only}*
Chen H-H, Sun F-J, Yeh T-L, Liu H-E, Huang H-L, Kuo BI, et al. The diagnostic accuracy of the Ascertain Dementia 8 questionnaire for detecting cognitive impairment in primary care in the community, clinics and hospitals: a systematic review and meta-analysis. *Family Practice* 2018; **35**:239–46.

**Chin 2013** *{published data only}*
Chin R, Ng A, Narasimhalu K, Kandiah N. Utility of the AD8 as a self-rating tool for cognitive impairment in an Asian population. *American Journal of Alzheimer's Disease and Other Dementias* 2013;**28**:284–8.

**Dong 2013** *{published data only}*
Dong Y, Pang WS, Lim LBS, Yang Y, Morris JC, Hilal S, et al. The informant AD8 is superior to participant AD8 in detecting cognitive impairment in a Memory Clinic setting. *Journal of Alzheimer's Disease* 2013;**35**:159–68.

**Dong 2014** *{published data only}*
Dong Y, Cheng TS, Tsou KYK, Chan QL, Chen CL. Feasibility and acceptability of the informant AD8 for cognitive screening in primary healthcare: a pilot study. *The Scientific World Journal* 2014;**2014**:302834.

**Dyer 2017** *{published data only}*
Dyer AH, Briggs R, Nabeel S, O'Neill D, Kennelly SP. The Abbreviated Mental Test 4 for cognitive screening of older adults presenting to the Emergency Department. *European Journal of Emergency Medicine* 2017;**6**:417–22.

**Galvin 2005** *{published data only}*
Galvin JE, Roe CM, Powlishta KK, Coats MA, Muich SJ, Grant E, et al. The AD8 a brief informant interview to detect dementia. *Neurology* 2005;**65**:559–64.

**Galvin 2007ii** *{published data only}*
Galvin JE, Roe CM, Morris JC. Evaluation of cognitive impairment in older adults. *Archives of Neurology* 2007;**64**: 718–24.

**Galvin 2010** *{published data only}*
Galvin JE, Fagan AM, Holtzman DM, Mintun MA, Morris JC. Relationship of dementia screening tests with biomarkers of Alzheimer's Disease. *Brain: A Journal of Neurology* 2010;**133**:3290–3300.

**Gottesman 2018** *{published data only}*
Gottesman R, Weiss E, Facchini R, Valdesuso K, Marquez C, Zwerling J. The correlation between reported changes in cognitive functioning and objective performance on cognitive screening in English-and Spanish-speaking patients. *Neurology* 2018;**90**(15 Supplement):P5.174.

**Hsieh 2015** *{published data only}*
Hsieh I, Kuan T, Hsieh P, Chem S, Yen W, Lin Y. Detection of early cognitive impairment using AD8 in a young patient with stroke with cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy syndrome: A case report. *Current Topics in Care* 2015;**29**: 133–7.


**Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.**

**Li 2012** *{published data only}*
Li T, Wang HL, Yang YH, Galvin JE, Morris JC, Yu X. The reliability and validity of Chinese version of AD8 (abstract). *Zhonghua Nei Ke Za Zhi* 2012;**51**:777–80.

**Malmstrom 2009** *{published data only}*
Malmstrom TK, Miller DK, Coats MA, Jackson P, Miller JP, Morris JC. Informant-based dementia screening in a population-based sample of African Americans. *Alzheimer Disease and Associated Disorders* 2009;**23**:117–23.

**Munoz 2010** *{published data only}*
Munoz C, Nunez J, Flores P, Behrens MI, Slachevsky YA. Usefulness of a brief informant interview to detect dementia, translated into Spanish (AD8-Ch) [Utilidad de un cuestionario breve dirigido alinformante para el diagnóstico temprano decasos de demencia: La versión chilena del AD8(AD8–Ch)]. *Review of Medicine Chile* 2010;**138**:1063–65.

**Overton 2013** *{published data only}*
Overton ET, Azad TD, Parker N, Shaw DD, Frain J, et al. The Alzheimer's disease-8 and Montreal Cognitive Assessment as screening tools for neurocognitive impairment in HIV-infected patients. *Journal of Neurovirology* 2013;**19**: 109–16.

**Pardo 2013** *{published data only}*
Carnero Pardo C, de la Vega Cotarelo R, Alcalde SL, Aparicio CM, Carrillo V, Gavilan EM, et al. Assessing the diagnostic accuracy (DA) of the Spanish version of the informant-based AD8 questionnaire. *Neurologia* 2013;**28**: 88–94.

**Ryu 2009** *{published data only}*
Ryu HJ, Kim H, Han S. Validity and reliability of the Korean version of the AD8 informant interview (K-AD8) in dementia. *Alzheimer Disease and Associated Disorders* 2009; **23**:371–6.

**Tew 2015** *{published data only}*
Tew CW, Ng TP, Cheong CY, Yap P. A brief dementia test with subjective and objective measures. *Dementia and Geriatric Cognitive Disorders Extra* 2015;**5**(3):341–9.

**Wojtowicz 2017** *{published data only}*
Wojtowicz A, Larner AJ. Diagnostic test accuracy of cognitive screeners in older people. *Progress in Neurology and Psychiatry.* 2017;**21**(1):17–21.

**Xie 2014** *{published data only}*
Xie Y, Gao Y, Jia J, Wang X, Wang Z, Xie H. Utility of AD8 for cognitive impairment in a Chinese physical examination population: a preliminary study. *The Scientific World Journal* 2014;**2014**:804871.

**Yang 2011** *{published data only}*
Yang Y, Galvin J, Morris JC, Lai C, Chou M, Liu C. Application of AD8 questionnaire to screen very mild dementia in Taiwanese. *American Journal of Alzheimer's Disease and Other Dementias* 2011;**26**:134–8.

## Additional references

**Anwer 2017**
Anwer S. A systematic review and meta-analysis of the performance of the AD-8 for the diagnosis of dementia including synthesis of multiple threshold data from individual studies [Masters thesis]. University of Leicester 2017.

**Bahar-Fuchs 2013**
Bahar-Fuchs A, Clare L, Woods B. Cognitive training and cognitive rehabilitation for mild to moderate Alzheimer's disease and vascular dementia. *Cochrane Database of Systematic Reviews* 2013, Issue 6. DOI: 10.1002/ 14651858.CD003260.pub2

**Banerjee 2009**
Banerjee S, Wittenberg R. Clinical and cost effectiveness of services for early diagnosis and intervention in dementia. *International Journal of Geriatric Psychiatry* 2009;**24**(7): 748–54.

**Birks 2018**
Birks JS, Harvey RJ. Donepezil for dementia due to Alzheimer's disease. *Cochrane Database of Systematic Reviews* 2018, Issue 6. DOI: 10.1002/14651858.CD001190

**Boustani 2003**
Boustani M, Peterson B, Hanson L, Harris R, Lohr KN. Screening for dementia in primary care: a summary of the evidence for the US Preventative Services Task Force. *Annals of Internal Medicine* 2003;**138**(11):927–37.

**Bradford 2009**
Bradford A, Kunik ME, Schulz P, Williams SP, Singh H. Missed and delayed diagnosis of dementia in primary care: prevalence and contributing factors. *Alzheimer Disease and Associated Disorders* 2009;**23**(4):306–14.

**Brodaty 2002**
Brodaty H. The GPCOG: a new screening test for dementia designed for general practice. *Journal of the American Geriatrics Society* 2002;**50**(3):530–4.

**Brunet 2012**
Brunet MD, McCartney M, Heath I, Tomlinson J, Gorodon P, Cosgrove J, et al. There is no evidence base for proposed dementia screening. *BMJ* 2012;**345**:e8588.

**Burn 2018**
Burn AM, Fleming J, Brayne C, Fox C, Bunn F. Dementia case-finding in hospitals: a qualitative study exploring the views of healthcare professionals in English primary care and secondary care. *BMJ Open* 2018;**8**(3):e020521.

**Chodosh 2004**
Chodosh J, Petitti DB, Elliott M, Hays RD, Crooks VC, Reuben DB, et al. Physician recognition of cognitive impairment: evaluating the need for improvement. *Journal of the American Geriatrics Society* 2004;**52**(7):1051–9.

**Cordell 2013**
Cordell CB, Borson S, Boustani M, Chodosh J, Reuben D, Verghese J, et al. Medicare Detection of Cognitive Impairment Workgroup. Alzheimer's Association recommendations for operationalizing the detection of cognitive impairment during Medicare annual wellness visit


**Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.**

in a primary care setting. *Alzheimer's & Dementia* 2013;**9** (2):141–50.

**Davis 2013**

David DJH, Creavin ST, Noel-Storr A, Quinn TJ, Smailagic N, Hyde C, et al. Neuropsychological tests for the diagnosis of Alzheimer's disease dementia and other dementias: a generic protocol for cross-sectional and delayed verification studies.. *Cochrane Database of Systematic Reviews* 2013, Issue 4. DOI: 10.1002/14651858.CD010460

**Davis 2015**

Davis DH, Creavin ST, Yip JL, Noel-Storr AH, Brayne C, Cullum S. Montreal Cognitive Assessment for the diagnosis of Alzheimer's disease and other dementias. *Cochrane Database of Systematic Reviews* 2015, Issue 10. DOI: 10.1002/14651858.CD010775.pub2

**Donegan 2017**

Donegan K, Fox N, Black N, Livingston G, Banerjee S, Burns A. Trends in diagnosis and treatment for people with dementia in the UK from 2005 to 2015: a longitudinal retrospective cohort study. *Lancet Public Health* 2017;**2**: e149–56.

**DSM-IV**

American Psychiatric Association. *Diagnostic and statistical manual of mental disorders (4th Ed)*. Washington DC: American Psychiatric Association, 2000.

**Ferri 2005**

Ferri CP, Prince M, Brayne C, Brodaty H, Fratiglioni L, Ganguli M, et al. Alzheimer's Disease International. Global prevalence of dementia: a Delphi consensus study. *Lancet* 2005;**366**(9503):2112–7.

**Folstein 1975**

Folstein MF, Folstein SE, McHugh PR. "Minimental state": a practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research* 1975;**12**(3): 189–98.

**Glanville 2010**

Glanville JM, Cikalo M, Crawford F, Dozier M, Lowson P. Handsearching for reports of diagnostic test accuracy studies adding to the evidence base. Joint Cochrane and Campbell Colloquium, Madrid, Spain. 2010.

**Greenhalgh 2005**

Greenhalgh T, Peacock R. Effectiveness and efficiency of search methods in systematic reviews of complex evidence. *BMJ* 2005;**331**(7524):1064–5.

**Harrison 2014**

Harrison JK, Fearon P, Noel-Storr AH, McShane R, Stott DJ, Quinn TJ. Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE) for the diagnosis of dementia within a general practice (primary care) setting. *Cochrane Database of Systematic Reviews* 2014, Issue 7. DOI: 10.1002/14651858.CD010771.pub2

**Harrison 2015**

Harrison JK, Fearon P, Noel-Storr AH, McShane R, Stott DJ, Quinn TJ. Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE) for the diagnosis of dementia within a secondary care setting. *Cochrane Database of Systematic Reviews* 2015, Issue 3. DOI: 10.1002/14651858.CD010772.pub2

**Harrison 2016**

Harrison JK, Noel-Storr AH, Demeyere N, Reynish EL, Quinn TJ. Outcome measures in a decade of dementia and mild cognitive impairment trials. *Alzheimer's Research and Therapy* 2016;**8**(1):48.

**Hebert 2013**

Hebert LE, Weuve J, Scherr PA, Evans DA. Alzheimer disease in the United States (2010-2050) estimated using the 2010 census. *Neurology* 2013;**80**(19):1778–83.

**ICD-10**

World Health Organisation. *International Classification of Diseases (ICD) 10th Edition*. WHO, 1992.

**Iliffe 2009**

Iliffe S, Robinson L, Brayne C, Goodman C, Rait G, Manthorpe J, Ashley P, DeNDRoN Primary Care Clinical Studies Group. Primary care and dementia: 1. diagnosis, screening and disclosure. *International Journal of Geriatric Psychiatry* 2009;**24**(9):895–901.

**Iliffe 2010**

Illiffe S, Pealing L. Subjective memory problems. *BMJ* 2010;**340**:c1425.

**Inouye 2014**

Inouye SK, Westendorp RGJ, Saczynski JS. Delirium in elderly people. *The Lancet* 2014;**383**(9920):911–22.

**Jackson 2017**

Jackson TA, Gladman JRF, Harwood RH, MacLullich AMJ, Sampson EL, Sheehan B, Davis DHJ. Challenges and opportunities in understanding dementia and delirium in the acute hospital. *PLoS Medicine* 2017;**14**(3):e1002247.

**Jorm 2004**

Jorm AF. The Informant Questionnaire on cognitive decline in the elderly (IQCODE): a review. *International Psychogeriatrics* 2004;**16**(3):275–93.

**Kohn 2013**

Kohn MA, Carpenter CR, Newman TB. Understanding the direction of bias in studies of diagnostic test accuracy. *Academic Emergency Medicine* 2013;**20**(11):1194–1206.

**Lang 2017**

Lang L, Clifford A, Wei L, Zhang D, Leung D, Augustine G, et al. Prevalence and determinants of undetected dementia in the community: a systematic literature review and a meta-analysis. *BMJ Open* 2017;**7**(2):e011146.

**Lees 2017**

Lees RA, Hendry Ba K, Broomfield N, Stott D, Larner AJ, Quinn TJ. Cognitive assessment in stroke: feasibility and test accuracy using differing approaches to scoring of incomplete items. *International Journal of Geriatric Psychiatry* 2017;**32**(10):1072–78.


**Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.**

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

**Matthews 2013**
Matthews FE, Arthur A, Barnes LE, Bond J, Jagger C, Robinson L, et al. A two-decade comparison of prevalence of dementia in individuals aged 65 and older from three geographical areas of England: results of the Cognitive Function and Ageing Study I and II. *Lancet* 2013;**382** (9902):1405–12.

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

**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.

**McKhann 2001**
McKhann GM, Albert MS, Grossman M, Miller B, Dickson D, Trojanowski JQ: Work Group on Frontotemporal dementia and Pick's disease. Clinical and pathological diagnosis of frontotemporal dementia: report of the Work Group on Frontotemporal Dementia and Pick's Disease. *Archives of Neurology* 2001;**58**(11):1803–9.

**McKhann 2011**
McKhann GM, Knopman DS, Chertkow H, Hyman BT, Jack CR Jr, Kawas CH, et al. The diagnosis of dementia due to Alzheimer's disease:recommendations from the National Institute on Aging - Alzheimer's Association workgroup on diagnostic guidelines for Alzheimer's disease. *Alzheimer's & Dementia* 2011;**7**(3):263–9.

**McShane 2006**
McShane R, Areosa Sastre A, Minakaran N. Memantine for dementia. *Cochrane Database of Systematic Reviews* 2006, Issue 2. DOI: 10.1002/14651858.CD003154

**Menon 2011**
Menon R, Larner AJ. Use of cognitive screening instruments in primary care: the impact of national dementia directives. *Family Practice* 2011;**28**(3):272–6.

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

**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**:364–73.

**Office for National Statistics 2018**
Office for National Statistics. Overview of the UK population: November 2017. Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/articles/overviewoftheukpopulation/november2018 2018.

**Owen 2018**
Owen RK, Cooper NJ, Quinn TJ, Lees RA, Sutton AJ. Network meta-analysis of diagnostic test accuracy studies identifies and ranks the optimal diagnostic tests and thresholds for health care policy and decision-making. *Journal of Clinical Epidemiology* 2018;**99**:64–74.

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

**Quinn 2014**
Quinn TJ, Fearon P, Noel-Starr A, Young C, McShane R, Stott DJ. Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE) for the diagnosis of dementia within community dwelling populations. *Cochrane Database of Systematic Reviews* 2014, Issue 4. DOI: 10.1002/14651858.CD010079.pub2

**RevMan 2012 [Computer program]**
The Nordic Cochrane Centre, The Cochrane Collaboration. Review Manager (RevMan). Version 5.2. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2012.

**Robinson 2015**
Robinson L, Tang E, Taylor JP. Dementia: timely diagnosis and early intervention. *BMJ* 2015;**350**:h3029.

**Román 1993**
Román GC, Tatemichi TK, Erkinjutti T, Cummings JL, Masdeu JC, Garcia JH, et al. Vascular dementia: diagnostic criteria for research studies. Report of the NINDS-AIREN international workshop. *Neurology* 1993;**43**(2):250–60.

**Rutjes 2006**
Rutjes AWS, Reitsma JB, Di Nisio M, Smidt N, van Rijn JC, Bossuyt PMM. Evidence of bias and variation in diagnostic accuracy studies. *Canadian Medical Association Journal* 2006;**174**(4):469–476.

**SAS release 9.4 [Computer program]**
SAS Institute. SAS release 9.4. SAS Institute, 2013.

**Takwoingi 2018**
Takwoingi Y, Quinn TJ. Review of diagnostic test accuracy studies in older people. *Age and Ageing* 2018;**47**:349–55.

**Valcour 2000**
Valcour VG, Masaki KH, Curb JD, Blanchette PL. The detection of dementia in the primary care setting. *Archives of Internal Medicine* 2000;**160**(19):2964–8.

**van Enst 2014**
van Enst WA, Ochodo E, Scholten RJ, Hooft L, Leeflang MM. Investigation of publication bias in meta-analyses of


**Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.**

diagnostic test accuracy: a meta-epidemiological study. *BMC Medical Research Methodology* 2014;**14**:70.
* *Indicates the major publication for the study*


**Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.**

## CHARACTERISTICS OF STUDIES

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

**Chan 2016**

| **Study characteristics** | | | |
|---|---|---|---|
| Patient sampling | Patient-informant dyads were recruited in two phases.<br>Phase I (screening): "participants were recruited if they were aged 60 and above, had provided informed consent, and had a suitable informant with frequent interactions for at least 10 hours."<br>Phase II (validation): "participants were recruited if they had completed phase 1, provided informed consent, and had no physical disabilities that affected their ability to perform the cognitive assessments." Participants underwent additional evaluations<br>Most exclusions were due to patient choice; no other exclusion criteria were specified | | |
| Patient characteristics and setting | Patients at government-subsidised primary healthcare settings aged over 60 years | | |
| Index tests | AD-8 Chinese and Malay translations | | |
| Target condition and reference standard(s) | DSM-IV criteria and CDR global scores used to determine presence of dementia and severity | | |
| Flow and timing | There probably is a gap between Phase I and Phase II, but the duration is not specified or explained | | |
| Comparative | | | |
| Notes | All patients who made it to Phase II were included, but more than 90% of the patients found eligible for Phase I were lost | | |
| **Methodological quality** | | | |
| **Item** | **Authors' judgement** | **Risk of bias** | **Applicability concerns** |
| **DOMAIN 1: Patient Selection** | | | |
| Was a consecutive or random sample of patients enrolled? | Unclear | | |
| Was a case-control design avoided? | Unclear | | |
| Did the study avoid inappropriate exclusions? | Unclear | | |
| | | **Unclear** | **Low** |
| **DOMAIN 2: Index Test All tests** | | | |


**Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.**

**Chan 2016** (*Continued*)

| | | | |
|---|---|---|---|
| Were the index test results interpreted without knowledge of the results of the reference standard? | Yes | | |
| If a threshold was used, was it pre-specified? | 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? | Unclear | | |
| Did all patients receive the same reference standard? | Unclear | | |
| Were all patients included in the analysis? | No | | |
| | | **High** | |

**Chio 2017**

| | |
|---|---|
| **Study characteristics** | |
| Patient sampling | Patients were recruited from two neurology outpatient clinics. Patients with subjective or objective memory complaints were referred by physicians<br>Exclusions:<br>1. diagnosis of psychological disorder other than dementia;<br>2. inability to communicate verbally.<br>Participants who scored CDR of higher than 2 were excluded from further analysis |


**Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.**

**Chio 2017** (*Continued*)

| | | | |
|---|---|---|---|
| Patient characteristics and setting | Patients recruited from neurology outpatient clinics<br>Physician referrals if they had subjective or objective memory complaints | | |
| Index tests | AD-8. It is unclear whether the test was translated or conducted in English | | |
| Target condition and reference standard(s) | CDR<br>Early cognitive impairment | | |
| Flow and timing | The timing between questionnaire and data collection varies between participants and is not described in detail | | |
| Comparative | | | |
| Notes | | | |
| **Methodological quality** | | | |
| **Item** | **Authors' judgement** | **Risk of bias** | **Applicability concerns** |
| **DOMAIN 1: Patient Selection** | | | |
| Was a consecutive or random sample of patients enrolled? | Unclear | | |
| Was a case-control design avoided? | Unclear | | |
| 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 | | |
| If a threshold was used, was it pre-specified? | No | | |
| | | **Unclear** | **Unclear** |
| **DOMAIN 3: Reference Standard** | | | |
| Is the reference standards likely to correctly classify the target condition? | Yes | | |


**Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.**

**Chio 2017** (*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? | Unclear | | |
| Did all patients receive the same reference standard? | Yes | | |
| Were all patients included in the analysis? | No | | |
| | | **Unclear** | |

**Correia 2011**

| | |
|---|---|
| **Study characteristics** | |
| Patient sampling | "The elderly were identified and invited to participate in the survey"<br>Sample size was pre-calculated based on total population and dementia prevalence in that population (aimed to recruit 111 participants which included a 10% surplus)<br>Study does not state how these specific participants were selected from a population of 18,904 elderly<br>However, dementia prevalence was found to be 13.8%, so not vastly over the stated dementia prevalence of the whole target population (7.1%)<br>The only stated exclusion criterion is not having a relative who has known the individual for over 10 years; it is unclear if there were others |
| Patient characteristics and setting | Community-dwelling elderly aged 65 years and over in a city in Brazil |
| Index tests | AD-8, Portuguese translation |
| Target condition and reference standard(s) | DSM-IV<br>Carried out independently and blinded to AD-8 results<br>Combination of patient and informant assessment |
| Flow and timing | Clinical assessment and independent informant questionnaire performed on same day |
| Comparative | |


**Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.**

| | | | |
|---|---|---|---|
| Notes | | | |
| **Methodological quality** | | | |
| **Item** | **Authors' judgement** | **Risk of bias** | **Applicability concerns** |
| **DOMAIN 1: Patient Selection** | | | |
| Was a consecutive or random sample of patients enrolled? | Unclear | | |
| Was a case-control design avoided? | Yes | | |
| Did the study avoid inappropriate exclusions? | Unclear | | |
| | | **Unclear** | **Low** |
| **DOMAIN 2: Index Test All tests** | | | |
| Were the index test results interpreted without knowledge of the results of the reference standard? | Yes | | |
| If a threshold was used, was it pre-specified? | Yes | | |
| | | **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 | | |


**Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.**

| | | | |
|---|---|---|---|
| Did all patients receive the same reference standard? | Yes | | |
| Were all patients included in the analysis? | Yes | | |
| | | **Low** | |

**Galvin 2006**

| **Study characteristics** | | | |
|---|---|---|---|
| Patient sampling | Consecutive sample of patients referred to memory clinic | | |
| Patient characteristics and setting | Memory clinic referrals | | |
| Index tests | AD-8 | | |
| Target condition and reference standard(s) | Clinical diagnosis made using various diagnostic criteria including DSM | | |
| Flow and timing | Clinical assessment and informant assessment within one week of each other | | |
| Comparative | | | |
| Notes | | | |
| **Methodological quality** | | | |
| **Item** | **Authors' judgement** | **Risk of bias** | **Applicability concerns** |
| **DOMAIN 1: Patient Selection** | | | |
| Was a consecutive or random sample of patients enrolled? | Yes | | |
| Was a case-control design avoided? | Yes | | |
| Did the study avoid inappropriate exclusions? | Yes | | |
| | | **Low** | **Low** |
| **DOMAIN 2: Index Test All tests** | | | |


**Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.**

**Galvin 2006** (*Continued*)

| | | | |
|---|---|---|---|
| Were the index test results interpreted without knowledge of the results of the reference standard? | Yes | | |
| If a threshold was used, was it pre-specified? | 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? | No | | |
| | | **Unclear** | |

**Galvin 2007**

| | |
|---|---|
| **Study characteristics** | |
| Patient sampling | Consecutive patient-informant dyads enrolled in a longitudinal study |
| Patient characteristics and setting | Community-dwelling volunteers |
| Index tests | AD-8 |


**Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.**

**Galvin 2007** (*Continued*)

| | | | |
|---|---|---|---|
| Target condition and reference standard(s) | CDR used to determine presence or absence of dementia and to stage its severity. Authors classified CDR 0.5 as mild dementia | | |
| Flow and timing | Timing of clinical assessment and informant questionnaire not clear | | |
| Comparative | | | |
| Notes | | | |
| **Methodological quality** | | | |
| **Item** | **Authors' judgement** | **Risk of bias** | **Applicability concerns** |
| **DOMAIN 1: Patient Selection** | | | |
| Was a consecutive or random sample of patients enrolled? | Yes | | |
| Was a case-control design avoided? | Yes | | |
| Did the study avoid inappropriate exclusions? | Yes | | |
| | | **Low** | **Low** |
| **DOMAIN 2: Index Test All tests** | | | |
| Were the index test results interpreted without knowledge of the results of the reference standard? | Yes | | |
| If a threshold was used, was it pre-specified? | 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 | | |


**Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.**

**Galvin 2007** (*Continued*)

| | | | |
|---|---|---|---|
| | | **Low** | **Low** |
| **DOMAIN 4: Flow and Timing** | | | |
| Was there an appropriate interval between index test and reference standard? | Unclear | | |
| Did all patients receive the same reference standard? | Yes | | |
| Were all patients included in the analysis? | No | | |
| | | **Unclear** | |

**Jackson 2016**

| | | | |
|---|---|---|---|
| **Study characteristics** | | | |
| Patient sampling | Consecutive, consenting older adult admissions to a UK hospital with delirium | | |
| Patient characteristics and setting | All patients had delirium at time of recruitment, so results may not be applicable to an unselected cohort | | |
| Index tests | AD-8 | | |
| Target condition and reference standard(s) | Diagnostic interview based on DSM-IV performed at three months post-delirium, to assess pre-delirium dementia | | |
| Flow and timing | Of 228 with delirium, only 125 contributed data. Diagnostic interview was performed at least three months following questionnaire administration, in context of acute illness there may have been important cognitive change in this time | | |
| Comparative | | | |
| Notes | Main reason for non-contribution data was lack of an informant | | |
| **Methodological quality** | | | |
| **Item** | **Authors' judgement** | **Risk of bias** | **Applicability concerns** |
| **DOMAIN 1: Patient Selection** | | | |
| Was a consecutive or random sample of patients enrolled? | Yes | | |


**Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.**

| | | | |
|---|---|---|---|
| Was a case-control design avoided? | Yes | | |
| Did the study avoid inappropriate exclusions? | Yes | | |
| | | **Low** | **High** |
| **DOMAIN 2: Index Test All tests** | | | |
| Were the index test results interpreted without knowledge of the results of the reference standard? | Yes | | |
| If a threshold was used, was it pre-specified? | Yes | | |
| | | **Low** | **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** |
| **DOMAIN 4: Flow and Timing** | | | |
| Was there an appropriate interval between index test and reference standard? | Unclear | | |
| Did all patients receive the same reference standard? | Yes | | |
| Were all patients included in the analysis? | No | | |
| | | **High** | |


**Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.**

**Larner 2015**

| **Study characteristics** | | | |
|---|---|---|---|
| Patient sampling | Consecutive patient-informant dyads recruited prospectively<br>Exclusions: (1) patients without an informant; (2) patients with an unreliable informant, which consisted of informants with insufficient knowledge, dementia already diagnosed, language barriers and very young informant | | |
| Patient characteristics and setting | Referrals to the Cognitive Function Clinic at a regional neuroscience centre | | |
| Index tests | AD-8 | | |
| Target condition and reference standard(s) | Cognitive impairment<br>DSM IV was used, as well as looking at patient history for cognitive symptoms and functional performance, neuroradiological examination, neuropsychological assessment<br>Diagnosis was performed at the discretion of an experienced clinician | | |
| Flow and timing | Clinical and informant assessment performed on the same day. Large numbers attended alone (i.e. no informant) and so were not included in analysis | | |
| Comparative | | | |
| Notes | | | |
| **Methodological quality** | | | |
| **Item** | **Authors' judgement** | **Risk of bias** | **Applicability concerns** |
| **DOMAIN 1: Patient Selection** | | | |
| Was a consecutive or random sample of patients enrolled? | Yes | | |
| Was a case-control design avoided? | Yes | | |
| Did the study avoid inappropriate exclusions? | Unclear | | |
| | | **Low** | **Low** |
| **DOMAIN 2: Index Test All tests** | | | |
| Were the index test results interpreted without knowledge of the results of the reference standard? | Yes | | |


**Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.**

**Larner 2015** (*Continued*)

| | | | |
|---|---|---|---|
| If a threshold was used, was it pre-specified? | Yes | | |
| | | **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 | | |
| | | **Unclear** | |

**Meguro 2015**

| | |
|---|---|
| **Study characteristics** | |
| Patient sampling | Community residents volunteered |
| Patient characteristics and setting | Community residents aged 75 years and older in Northern Japan |
| Index tests | AD-8 (Japanese translation) |
| Target condition and reference standard(s) | CDR |
| Flow and timing | Insufficient information to draw a flow chart |


**Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.**

**Meguro 2015** (*Continued*)

| | | | |
|---|---|---|---|
| Comparative | | | |
| Notes | | | |
| **Methodological quality** | | | |
| **Item** | **Authors' judgement** | **Risk of bias** | **Applicability concerns** |
| **DOMAIN 1: Patient Selection** | | | |
| Was a consecutive or random sample of patients enrolled? | Unclear | | |
| Was a case-control design avoided? | Unclear | | |
| Did the study avoid inappropriate exclusions? | Unclear | | |
| | | **Unclear** | **Low** |
| **DOMAIN 2: Index Test All tests** | | | |
| Were the index test results interpreted without knowledge of the results of the reference standard? | Unclear | | |
| If a threshold was used, was it pre-specified? | Unclear | | |
| | | **Unclear** | **Unclear** |
| **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** | | | |


**Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.**

**Meguro 2015** (*Continued*)

| | | | |
|---|---|---|---|
| Was there an appropriate interval between index test and reference standard? | Unclear | | |
| Did all patients receive the same reference standard? | Unclear | | |
| Were all patients included in the analysis? | Unclear | | |
| | | **Unclear** | |

**Razavi 2014**

| **Study characteristics** | | | |
|---|---|---|---|
| Patient sampling | Consecutive sample from memory clinic plus patients from a neurology clinic. Sample was 'enriched' with cognitively normal control patients<br>33 patients with comorbid medical conditions that could potentially affect cognition but were not dementias were excluded | | |
| Patient characteristics and setting | Mixed population from various referral pathways | | |
| Index tests | AD-8 | | |
| Target condition and reference standard(s) | The study states that it uses "published criteria" but does not further operationalise | | |
| Flow and timing | 33 patients with comorbid medical conditions that could potentially affect cognition but were not dementias were excluded. The timing of clinical and informant assessment is not clear | | |
| Comparative | | | |
| Notes | | | |
| **Methodological quality** | | | |
| **Item** | **Authors' judgement** | **Risk of bias** | **Applicability concerns** |
| **DOMAIN 1: Patient Selection** | | | |
| Was a consecutive or random sample of patients enrolled? | Yes | | |
| Was a case-control design avoided? | Unclear | | |


**Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.**

| | | | |
|---|---|---|---|
| Did the study avoid inappropriate exclusions? | No | | |
| | | **High** | **High** |
| **DOMAIN 2: Index Test All tests** | | | |
| Were the index test results interpreted without knowledge of the results of the reference standard? | Yes | | |
| If a threshold was used, was it pre-specified? | Yes | | |
| | | **Low** | **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** | **Low** |
| **DOMAIN 4: Flow and Timing** | | | |
| Was there an appropriate interval between index test and reference standard? | Unclear | | |
| Did all patients receive the same reference standard? | Yes | | |
| Were all patients included in the analysis? | No | | |
| | | **High** | |


**Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.**

**Yang 2016**

| **Study characteristics** | | | |
|---|---|---|---|
| Patient sampling | A cross-sectional survey was administered in 4 communities across 12 counties in the Zhejiang province using multi-stage stratified random cluster sampling<br>12 administrative districts were divided into 4 type districts, based on economic levels. One district was systematically chosen from each of the 4 district types<br>Inclusion criteria: (1) subjects aged 65 years or older; (2) subjects could complete CDT; (3) subjects/informants could sign consent<br>Exclusion criteria: subjects who had any of the following conditions: (1) acute severe disease such as acute myocardial infarction, acute cerebral infarction and other severe heart, brain, kidney and liver disease; (2) severe deafness or aphasia or subjects who did not consent | | |
| Patient characteristics and setting | Patients were aged 65 years and older living at home or in an institution in the Zhejiang province | | |
| Index tests | AD-8, unclear which language used | | |
| Target condition and reference standard(s) | Dementia<br>NIA-AA criteria | | |
| Flow and timing | Timing of assessments is not described. | | |
| Comparative | | | |
| Notes | | | |
| **Methodological quality** | | | |
| **Item** | **Authors' judgement** | **Risk of bias** | **Applicability concerns** |
| **DOMAIN 1: Patient Selection** | | | |
| Was a consecutive or random sample of patients enrolled? | Unclear | | |
| Was a case-control design avoided? | Unclear | | |
| 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 | | |


**Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.**

| | | | |
|---|---|---|---|
| If a threshold was used, was it pre-specified? | Yes | | |
| | | **Unclear** | **Unclear** |
| **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? | Unclear | | |
| Did all patients receive the same reference standard? | Yes | | |
| Were all patients included in the analysis? | Yes | | |
| | | **Unclear** | |

CDR: Clinical Dementia Rating scale
DSM-IV: Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition
NIA-AA: National Institute on Aging and Alzheimer's Association

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

| Study | Reason for exclusion |
|---|---|
| Basalo 2017 | Inappropriate index test |
| Blanco 2016 | Case-control study |


**Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.**

(*Continued*)

| | |
|---|---|
| Carnero 2013 | Case-control study |
| Carpenter 2011 | No dementia diagnosis reference standard |
| Carpenter 2011ii | No dementia diagnosis reference standard |
| Chen 2017 | AD-8 completed by patients |
| Chen 2018 | No original data |
| Chin 2013 | AD-8 completed by patients |
| Dong 2013 | Case-control study |
| Dong 2014 | No dementia diagnosis reference standard |
| Dyer 2017 | No dementia diagnosis reference standard |
| Galvin 2005 | Scale development paper, data incorporated in larger dataset included in analysis |
| Galvin 2007ii | Same dataset as included paper |
| Galvin 2010 | No dementia diagnosis reference standard |
| Gottesman 2018 | No dementia diagnosis reference standard |
| Hsieh 2015 | Case report |
| Li 2012 | Case-control study |
| Malmstrom 2009 | The study subgroup which could have been eligible did not include any participants with a cognitive diagnosis compatible with our prespecified criteria for dementia |
| Munoz 2010 | Case-control study |
| Overton 2013 | Case-control study |
| Pardo 2013 | Case-control study (also no useable DTA data) |
| Ryu 2009 | Case-control study |
| Tew 2015 | Case-control study |
| Wojtowicz 2017 | Same dataset as included paper |
| Xie 2014 | No dementia diagnosis reference standard (also AD-8 completed by patients) |


**Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.**

(*Continued*)

| | |
|---|---|
| Yang 2011 | Case-control study |

DTA: diagnostic test accuracy


**Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.**

## DATA

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

### Tests. Data tables by test

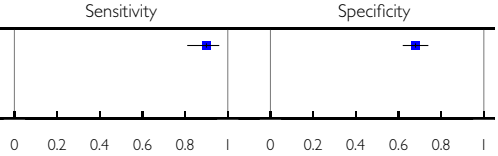
| Test | No. of studies | No. of participants |
|---|---|---|
| 1 AD-8 informant cutpoint 1 | 1 | 324 |
| 2 AD-8 informant cutpoint 2 | 7 | 3659 |
| 3 AD-8 informant cutpoint 3 | 5 | 1060 |
| 4 AD-8 informant cutpoint 7 | 1 | 77 |
| 5 Sensitivity analysis removing low average age | 6 | 3447 |
| 6 Sensitivity analysis removing disease-specific study | 4 | 983 |
| 7 English language studies - cutpoint 2 | 4 | 963 |
| 8 English language studies - cutpoint 3 | 3 | 642 |
| 9 Sensitivity analysis removing CDR 0.5 - cutpoint 2 | 6 | 3335 |
| 10 Sensitivity analysis removing CDR 0.5 - cutpoint 3 | 4 | 736 |

### Test 1. AD-8 informant cutpoint 1.

Review: AD-8 for detection of dementia across a variety of healthcare settings

Test: 1 AD-8 informant cutpoint 1

| Study | TP | FP | FN | TN | Sensitivity | Specificity |
|---|---|---|---|---|---|---|
| Galvin 2007 | 66 | 80 | 7 | 171 | 0.90 [ 0.81, 0.96 ] | 0.68 [ 0.62, 0.74 ] |





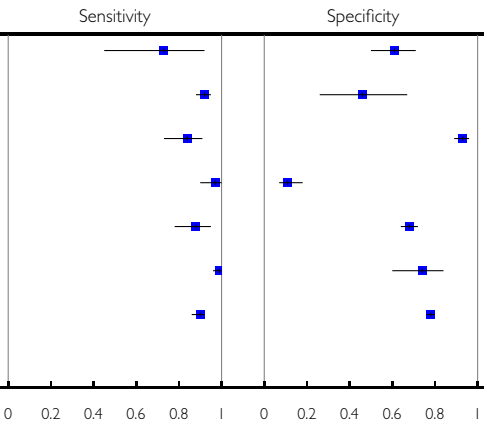
Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

## Test 2. AD-8 informant cutpoint 2.

Review: AD-8 for detection of dementia across a variety of healthcare settings

Test: 2 AD-8 informant cutpoint 2

| Study | TP | FP | FN | TN | Sensitivity | Specificity |
|---|---|---|---|---|---|---|
| Correia 2011 | 11 | 37 | 4 | 57 | 0.73 [ 0.45, 0.92 ] | 0.61 [ 0.50, 0.71 ] |
| Galvin 2006 | 200 | 13 | 17 | 11 | 0.92 [ 0.88, 0.95 ] | 0.46 [ 0.26, 0.67 ] |
| Galvin 2007 | 61 | 18 | 12 | 233 | 0.84 [ 0.73, 0.91 ] | 0.93 [ 0.89, 0.96 ] |
| Larner 2015 | 67 | 127 | 2 | 16 | 0.97 [ 0.90, 1.00 ] | 0.11 [ 0.07, 0.18 ] |
| Meguro 2015 | 61 | 161 | 8 | 342 | 0.88 [ 0.78, 0.95 ] | 0.68 [ 0.64, 0.72 ] |
| Razavi 2014 | 128 | 15 | 1 | 42 | 0.99 [ 0.96, 1.00 ] | 0.74 [ 0.60, 0.84 ] |
| Yang 2016 | 398 | 339 | 46 | 1232 | 0.90 [ 0.86, 0.92 ] | 0.78 [ 0.76, 0.80 ] |

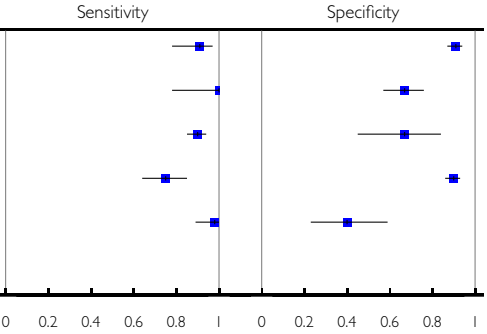



## Test 3. AD-8 informant cutpoint 3.

Review: AD-8 for detection of dementia across a variety of healthcare settings

Test: 3 AD-8 informant cutpoint 3

| Study | TP | FP | FN | TN | Sensitivity | Specificity |
|---|---|---|---|---|---|---|
| Chan 2016 | 40 | 24 | 4 | 241 | 0.91 [ 0.78, 0.97 ] | 0.91 [ 0.87, 0.94 ] |
| Correia 2011 | 15 | 31 | 0 | 63 | 1.00 [ 0.78, 1.00 ] | 0.67 [ 0.57, 0.76 ] |
| Galvin 2006 | 195 | 8 | 22 | 16 | 0.90 [ 0.85, 0.94 ] | 0.67 [ 0.45, 0.84 ] |
| Galvin 2007 | 55 | 25 | 18 | 226 | 0.75 [ 0.64, 0.85 ] | 0.90 [ 0.86, 0.93 ] |
| Jackson 2016 | 46 | 18 | 1 | 12 | 0.98 [ 0.89, 1.00 ] | 0.40 [ 0.23, 0.59 ] |





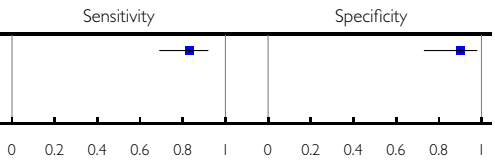
**Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.**

## Test 4. AD-8 informant cutpoint 7.

Review: AD-8 for detection of dementia across a variety of healthcare settings

Test: 4 AD-8 informant cutpoint 7

| Study | TP | FP | FN | TN | Sensitivity | Specificity |
|---|---|---|---|---|---|---|
| Jackson 2016 | 39 | 3 | 8 | 27 | 0.83 [ 0.69, 0.92 ] | 0.90 [ 0.73, 0.98 ] |

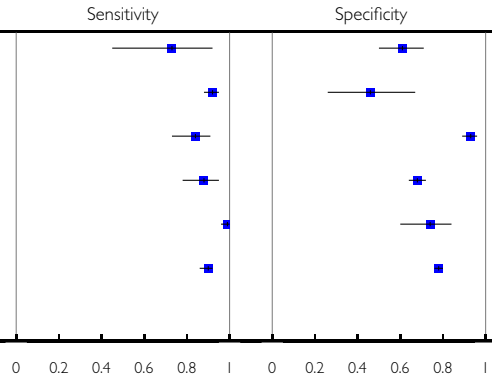



## Test 5. Sensitivity analysis removing low average age.

Review: AD-8 for detection of dementia across a variety of healthcare settings

Test: 5 Sensitivity analysis removing low average age

| Study | TP | FP | FN | TN | Sensitivity | Specificity |
|---|---|---|---|---|---|---|
| Correia 2011 | 11 | 37 | 4 | 57 | 0.73 [ 0.45, 0.92 ] | 0.61 [ 0.50, 0.71 ] |
| Galvin 2006 | 200 | 13 | 17 | 11 | 0.92 [ 0.88, 0.95 ] | 0.46 [ 0.26, 0.67 ] |
| Galvin 2007 | 61 | 18 | 12 | 233 | 0.84 [ 0.73, 0.91 ] | 0.93 [ 0.89, 0.96 ] |
| Meguro 2015 | 61 | 161 | 8 | 342 | 0.88 [ 0.78, 0.95 ] | 0.68 [ 0.64, 0.72 ] |
| Razavi 2014 | 128 | 15 | 1 | 42 | 0.99 [ 0.96, 1.00 ] | 0.74 [ 0.60, 0.84 ] |
| Yang 2016 | 398 | 339 | 46 | 1232 | 0.90 [ 0.86, 0.92 ] | 0.78 [ 0.76, 0.80 ] |





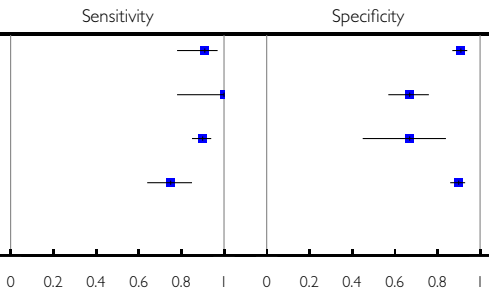
Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

## Test 6. Sensitivity analysis removing disease-specific study.

Review: AD-8 for detection of dementia across a variety of healthcare settings

Test: 6 Sensitivity analysis removing disease-specific study

| Study | TP | FP | FN | TN | Sensitivity | Specificity |
|---|---|---|---|---|---|---|
| Chan 2016 | 40 | 24 | 4 | 241 | 0.91 [ 0.78, 0.97 ] | 0.91 [ 0.87, 0.94 ] |
| Correia 2011 | 15 | 31 | 0 | 63 | 1.00 [ 0.78, 1.00 ] | 0.67 [ 0.57, 0.76 ] |
| Galvin 2006 | 195 | 8 | 22 | 16 | 0.90 [ 0.85, 0.94 ] | 0.67 [ 0.45, 0.84 ] |
| Galvin 2007 | 55 | 25 | 18 | 226 | 0.75 [ 0.64, 0.85 ] | 0.90 [ 0.86, 0.93 ] |

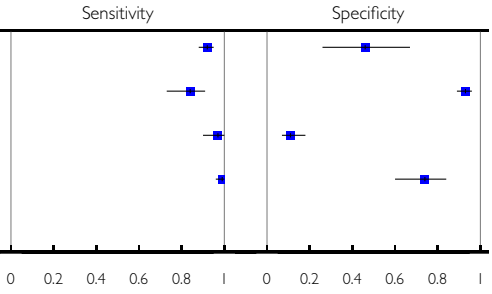



## Test 7. English language studies - cutpoint 2.

Review: AD-8 for detection of dementia across a variety of healthcare settings

Test: 7 English language studies - cutpoint 2

| Study | TP | FP | FN | TN | Sensitivity | Specificity |
|---|---|---|---|---|---|---|
| Galvin 2006 | 200 | 13 | 17 | 11 | 0.92 [ 0.88, 0.95 ] | 0.46 [ 0.26, 0.67 ] |
| Galvin 2007 | 61 | 18 | 12 | 233 | 0.84 [ 0.73, 0.91 ] | 0.93 [ 0.89, 0.96 ] |
| Larner 2015 | 67 | 127 | 2 | 16 | 0.97 [ 0.90, 1.00 ] | 0.11 [ 0.07, 0.18 ] |
| Razavi 2014 | 128 | 15 | 1 | 42 | 0.99 [ 0.96, 1.00 ] | 0.74 [ 0.60, 0.84 ] |





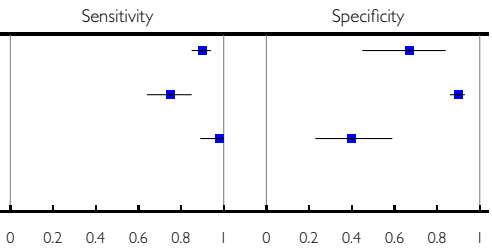
Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

### Test 8. English language studies - cutpoint 3.

Review: AD-8 for detection of dementia across a variety of healthcare settings

Test: 8 English language studies - cutpoint 3

| Study | TP | FP | FN | TN | Sensitivity | Specificity |
|---|---|---|---|---|---|---|
| Galvin 2006 | 195 | 8 | 22 | 16 | 0.90 [ 0.85, 0.94 ] | 0.67 [ 0.45, 0.84 ] |
| Galvin 2007 | 55 | 25 | 18 | 226 | 0.75 [ 0.64, 0.85 ] | 0.90 [ 0.86, 0.93 ] |
| Jackson 2016 | 46 | 18 | 1 | 12 | 0.98 [ 0.89, 1.00 ] | 0.40 [ 0.23, 0.59 ] |

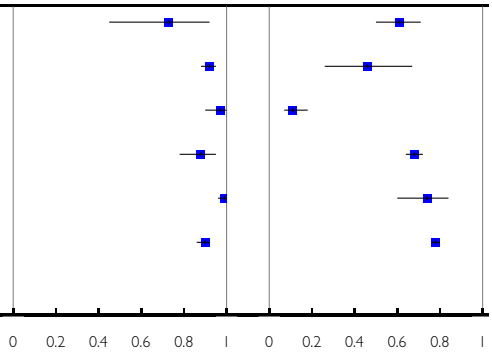



### Test 9. Sensitivity analysis removing CDR 0.5 - cutpoint 2.

Review: AD-8 for detection of dementia across a variety of healthcare settings

Test: 9 Sensitivity analysis removing CDR 0.5 - cutpoint 2

| Study | TP | FP | FN | TN | Sensitivity | Specificity |
|---|---|---|---|---|---|---|
| Correia 2011 | 11 | 37 | 4 | 57 | 0.73 [ 0.45, 0.92 ] | 0.61 [ 0.50, 0.71 ] |
| Galvin 2006 | 200 | 13 | 17 | 11 | 0.92 [ 0.88, 0.95 ] | 0.46 [ 0.26, 0.67 ] |
| Larner 2015 | 67 | 127 | 2 | 16 | 0.97 [ 0.90, 1.00 ] | 0.11 [ 0.07, 0.18 ] |
| Meguro 2015 | 61 | 161 | 8 | 342 | 0.88 [ 0.78, 0.95 ] | 0.68 [ 0.64, 0.72 ] |
| Razavi 2014 | 128 | 15 | 1 | 42 | 0.99 [ 0.96, 1.00 ] | 0.74 [ 0.60, 0.84 ] |
| Yang 2016 | 398 | 339 | 46 | 1232 | 0.90 [ 0.86, 0.92 ] | 0.78 [ 0.76, 0.80 ] |





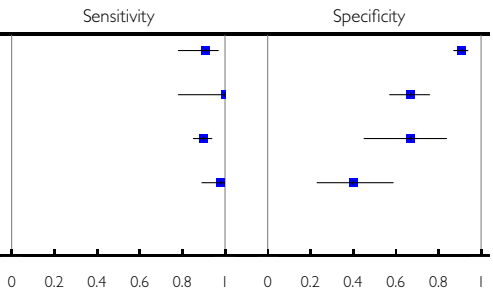
Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

### Test 10. Sensitivity analysis removing CDR 0.5 - cutpoint 3.

Review: AD-8 for detection of dementia across a variety of healthcare settings

Test: 10 Sensitivity analysis removing CDR 0.5 - cutpoint 3

| Study | TP | FP | FN | TN | Sensitivity | Specificity |
|---|---|---|---|---|---|---|
| Chan 2016 | 40 | 24 | 4 | 241 | 0.91 [ 0.78, 0.97 ] | 0.91 [ 0.87, 0.94 ] |
| Correia 2011 | 15 | 31 | 0 | 63 | 1.00 [ 0.78, 1.00 ] | 0.67 [ 0.57, 0.76 ] |
| Galvin 2006 | 195 | 8 | 22 | 16 | 0.90 [ 0.85, 0.94 ] | 0.67 [ 0.45, 0.84 ] |
| Jackson 2016 | 46 | 18 | 1 | 12 | 0.98 [ 0.89, 1.00 ] | 0.40 [ 0.23, 0.59 ] |




## ADDITIONAL TABLES

**Table 1. Summary of test accuracy at study level by dementia prevalence**

| Study | Included participants (n) | Dementia n (%) | AD-8 Threshold | Sensitivity (%) | Specificity (%) |
|---|---|---|---|---|---|
| Galvin 2006 | 241 | 217 (90) | 2 | 92 | 46 |
| Galvin 2006 | 241 | 217 (90) | 3 | 89 | 67 |
| Razavi 2014 | 186 | 129 (69) | 2 | 99 | 74 |
| Jackson 2016 | 77 | 47 (61) | 3 | 98 | 40 |
| Larner 2015 | 212 | 69 (33) | 2 | 97 | 11 |
| Galvin 2007 | 324 | 73 (23) | 2 | 84 | 93 |
| Galvin 2007 | 324 | 73 (23) | 3 | 75 | 90 |
| Yang 2016 | 2015 | 444 (22) | 2 | 90 | 78 |
| Correia 2011 | 109 | 15 (14) | 2 | 73 | 61 |
| Correia 2011 | 109 | 15 (14) | 3 | 100 | 67 |
| Chan 2016 | 309 | 44 (14) | 3 | 91 | 91 |
| Meguro 2015 | 572 | 69 (12) | 2 | 88 | 68 |


Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

**Table 1. Summary of test accuracy at study level by dementia prevalence** (*Continued*)

Where multiple thresholds were reported in the primary paper, we present the data for both a cut-off score of 2 and 3
The total number of participants is adjusted to reflect the data used in quantitative synthesis

# APPENDICES

## Appendix 1. WHO International Classification of Disease - Dementia

F00 - F09 ORGANIC, INCLUDING SYMPTOMATIC, MENTAL DISORDERS
DEMENTIA
G1. Evidence of each of the following.
(1) A decline in memory, which is most evident in the learning of new information, although in more severe cases, the recall of previously learned information may be also affected. The impairment applies to both verbal and non-verbal material. The decline should be objectively verified by obtaining a reliable history from an informant, supplemented, if possible, by neuropsychological tests or quantified cognitive assessments. The severity of the decline, with mild impairment as the threshold for diagnosis, should be assessed as follows.
Mild: a degree of memory loss sufficient to interfere with everyday activities, though not so severe as to be incompatible with independent living. The main function affected is the learning of new material. For example, the individual has difficulty in registering, storing and recalling elements in daily living, such as where belongings have been put, social arrangements, or information recently imparted by family members.
Moderate: a degree of memory loss which represents a serious handicap to independent living. Only highly learned or very familiar material is retained. New information is retained only occasionally and very briefly. The individual is unable to recall basic information about where he lives, what he has recently been doing, or the names of familiar persons.
Severe: a degree of memory loss characterized by the complete inability to retain new information. Only fragments of previously learned information remain. The subject fails to recognise even close relatives.
(2) A decline in other cognitive abilities characterized by deterioration in judgement and thinking, such as planning and organizing, and in the general processing of information. Evidence for this should be obtained when possible from interviewing an informant, supplemented, if possible, by neuropsychological tests or quantified objective assessments. Deterioration from a previously higher level of performance should be established. The severity of the decline, with mild impairment as the threshold for diagnosis, should be assessed as follows.
Mild: the decline in cognitive abilities causes impaired performance in daily living, but not to a degree making the individual dependent on others. More complicated daily tasks or recreational activities cannot be undertaken.
Moderate: the decline in cognitive abilities makes the individual unable to function without the assistance of another in daily living, including shopping and handling money. Within the home, only simple chores are preserved. Activities are increasingly restricted and poorly sustained.
Severe: the decline is characterised by an absence, or virtual absence, of intelligible ideation. The overall severity of the dementia is best expressed as the level of decline in memory or other cognitive abilities, whichever is the more severe (e.g. mild decline in memory and moderate decline in cognitive abilities indicate a dementia of moderate severity).
G2. Preserved awareness of the environment during a period of time long enough to enable the unequivocal demonstration of G1. When there are superimposed episodes of delirium the diagnosis of dementia should be deferred.
G3. A decline in emotional control or motivation, or a change in social behaviour, manifest as at least one of the following:

1. emotional lability;
2. irritability;
3. apathy;
4. coarsening of social behaviour.


**Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.**

G4. For a confident clinical diagnosis, G1 should have been present for at least six months; if the period since the manifest onset is shorter, the diagnosis can only be tentative.
Comments: the diagnosis is further supported by evidence of damage to other higher cortical functions, such as aphasia, agnosia, apraxia.
Judgement about independent living or the development of dependence (upon others) need to take account of the cultural expectation and context.
Dementia is specified here as having a minimum duration of six months to avoid confusion with reversible states with identical behavioural syndromes, such as traumatic subdural haemorrhage (S06.5), normal pressure hydrocephalus (G91.2) and diffuse or focal brain injury (S06.2 and S06.3).
A fifth character may be used to indicate the presence of additional symptoms, in the categories F00-F03 (F00 Dementia in Alzheimer's disease; F01 Vascular dementia; F02 Dementia in diseases classified elsewhere; and F03 Unspecified dementia), as follows:
.x0 without additional symptoms
.x1 with other symptoms, predominantly delusional
.x2 with other symptoms, predominantly hallucinatory
.x3 with other symptoms, predominantly depressive
.x4 with other mixed symptoms
A sixth character may be used to indicate the severity of the dementia:
.xx0 mild
.xx1 moderate
.xx2 severe
As mentioned above the overall severity of the dementia depends on the level of memory or intellectual impairment, whichever is the more severe.
F00 DEMENTIA IN ALZHEIMER'S DISEASE
A. The general criteria for dementia (G1 to G4) must be met.
B. There is no evidence from the history, physical examination or special investigations for any other possible cause of dementia (e.g. cerebrovascular disease, Parkinson's disease, Huntington's disease, normal pressure hydrocephalus), a systemic disorder (e.g. hypothyroidism, vit. B12 or folic acid deficiency, hypercalcaemia), or alcohol- or drug-abuse.
Comments: The diagnosis is confirmed by post mortem evidence of neurofibrillary tangles and neuritic plaques in excess of those found in normal ageing of the brain.
The following features support the diagnosis, but are not necessary elements: Involvement of cortical functions as evidenced by aphasia, agnosia or apraxia; decrease of motivation and drive, leading to apathy and lack of spontaneity; irritability and disinhibition of social behaviour; evidence from special investigations that there is cerebral atrophy, particularly if this can be shown to be increasing over time. In severe cases there may be Parkinson-like extrapyramidal changes, logoclonia, and epileptic fits.
Specification of features for possible subtypes. Because of the possibility that subtypes exist, it is recommended that the following characteristics be ascertained as a basis for a further classification: age at onset; rate of progression; the configuration of the clinical features, particularly the relative prominence (or lack) of temporal, parietal or frontal lobe signs; any neuropathological or neurochemical abnormalities, and their pattern.
The division of AD into subtypes can at present be accomplished in two ways: first by taking only the age of onset and labelling AD as either early or late, with an approximate cut-off point at 65 years; or secondly, by assessing how well the individual conforms to one of the two putative syndromes, early or late onset type. It should be noted that it is unlikely that a sharp distinction exists between early and late onset type. Early onset type may occur in late life, just as late onset type may occasionally have an onset under the age of 65.
The following criteria may be used to differentiate F00.0 from F00.1, but it should be remembered that the status of this subdivision is still controversial.
F00.0 Dementia in Alzheimer's disease with early onset
1. The criteria for dementia in Alzheimer's disease (F00) must be met, and the age at onset being under 65 years.
2. In addition, at least one of the following requirements must be met:
(a) evidence of a relatively rapid onset and progression;
(b) in addition to memory impairment, there is aphasia (amnesic or sensory), agraphia, alexia, acalculia, or apraxia (indicating the presence of temporal, parietal and/or frontal lobe involvement).
F00.1 Dementia in Alzheimer's disease with late onset
1. The criteria for dementia in Alzheimer's disease (F00) must be met and the age at onset must be 65 or more.
2. In addition, at least one of the following requirements must be met:


**Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.**

(a) evidence of a very slow, gradual onset and progression (the rate of the latter may be known only retrospectively after a course of 3 years or more);
(b) predominance of memory impairment G1.1, over intellectual impairment G1.2 (see general criteria for dementia).
F00.2 Dementia in Alzheimer's disease, atypical or mixed type
Use this term and code for dementias that have important atypical features or that fulfil criteria for both early and late onset type of Alzheimer's disease. Mixed Alzheimer's and vascular dementia is also included here.
F00.9 Dementia in Alzheimer's disease, unspecified
F01 VASCULAR DEMENTIA
G1. The general criteria for dementia (G1 to G4) must be met.
G2. Unequal distribution of deficits in higher cognitive functions, with some affected and others relatively spared. Thus memory may be quite markedly affected while thinking, reasoning and information processing may show only mild decline.
G3. There is clinical evidence of focal brain damage, manifest as at least one of the following:
(1) unilateral spastic weakness of the limbs;
(2) unilaterally increased tendon reflexes;
(3) an extensor plantar response;
(4) pseudobulbar palsy.
G4. There is evidence from the history, examination, or tests, of a significant cerebrovascular disease, which may reasonably be judged to be etiologically related to the dementia (e.g. a history of stroke; evidence of cerebral infarction).
The following criteria may be used to differentiate subtypes of vascular dementia, but it should be remembered that the usefulness of this subdivision may not be generally accepted.
F01.0 Vascular dementia of acute onset
A. The general criteria for vascular dementia (F01) must be met.
B. The dementia develops rapidly (i.e. usually within one month, but within no longer than three months) after a succession of strokes, or (rarely) after a single large infarction.
F01.1 Multi-infarct dementia
A. The general criteria for vascular dementia (F01) must be met.
B. The onset of the dementia is gradual (i.e. within three to six months), following a number of minor ischaemic episodes.
Comments: It is presumed that there is an accumulation of infarcts in the cerebral parenchyma. Between the ischaemic episodes there may be periods of actual clinical improvement.
F01.2 Subcortical vascular dementia
A. The general criteria for vascular dementia (F01) must be met.
B. A history of hypertension.
C. Evidence from clinical examination and special investigations of vascular disease located in the deep white matter of the cerebral hemispheres, with preservation of the cerebral cortex.
F01.3 Mixed cortical and subcortical vascular dementia
Mixed cortical and subcortical components of the vascular dementia may be suspected from the clinical features, the results of investigations (including autopsy), or both.
F01.8 Other vascular dementia
F01.9 Vascular dementia, unspecified
F02 DEMENTIA IN OTHER DISEASES CLASSIFIED ELSEWHERE
F02.0 Dementia in Pick's disease
A. The general criteria for dementia (G1 to G4) must be met.
B. Slow onset with steady deterioration.
C. Predominance of frontal lobe involvement evidenced by two or more of the following:
(1) emotional blunting;
(2) coarsening of social behaviour;
(3) disinhibition;
(4) apathy or restlessness;
(5) aphasia.
D. Relative preservation, in the early stages, of memory and parietal lobe functions.
F02.1 Dementia in Creutzfeldt-Jakob disease
A. The general criteria for dementia (G1 to G4) must be met.
B. Very rapid progression of the dementia, with disintegration of virtually all higher cerebral functions.


**Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.**

C. The emergence, usually after or simultaneously with the dementia, of one or more of the following types of neurological symptoms and signs:
(1) pyramidal symptoms;
(2) extrapyramidal symptoms;
(3) cerebellar symptoms;
(4) aphasia;
(5) visual impairment.
Comments: An akinetic and mute state is the typical terminal stage. An amyotrophic variant may be seen, where the neurological signs precede the onset of the dementia. A characteristic electroencephalogram (periodic spikes against a slow and low voltage background), if present in association with the above clinical signs, will increase the probability of the diagnosis. However, the diagnosis can be confirmed only by neuropathological examination (neuronal loss, astrocytosis, and spongiform changes). Because of the risk of infection, this should be carried out only under special protective conditions.
F02.2 Dementia in Huntington's disease
A. The general criteria for dementia (G1 to G4) must be met.
B. Subcortical functions are affected first and dominate the picture of dementia throughout; manifest as slowness of thinking or movement and personality alteration with apathy or depression.
C. Presence of involuntary choreiform movements, typically of the face, hands or shoulders, or in the gait. The patient may attempt to conceal them by converting them into a voluntary action.
D. A history of Huntington's disease in one parent or a sibling; or a family history which suggests the disorder.
E. The absence of clinical features otherwise accounting for the abnormal movements.
Comments: In addition to involuntary choreiform movements there may be development of extrapyramidal rigidity or spasticity with pyramidal signs.
F02.3 Dementia in Parkinson's disease
A. The general criteria for dementia (G1 to G4) must be met.
B. Diagnosis of Parkinson's disease.
C. Absence of cognitive impairment attributable to anti-parkinsonian medication.
D. There is no evidence from the history, physical examination or special investigations for any other possible cause of dementia, including other forms of brain disease, damage or dysfunction (e.g. cerebrovascular disease, HIV disease, Huntington's disease, normal pressure hydrocephalus), a systemic disorder (e.g. hypothyroidism, vit. B12 or folic acid deficiency, hypercalcaemia), or alcohol or drug abuse.
If criteria are also fulfilled for dementia in Alzheimer's disease with late onset (F00.1), this category F00.1 should be used in combination with Parkinson's disease G20.
F02.4 Dementia in human immunodeficiency (HIV) disease
A. The general criteria for dementia (G1 to G4) must be met.
B. Diagnosis of HIV infection.
C. There is no evidence from the history, physical examination or special investigations for any other possible cause of dementia, including other forms of brain disease, damage or dysfunction (e.g. Alzheimer's disease, cerebrovascular disease, Parkinson's disease, Huntington's disease, normal pressure hydrocephalus), a systemic disorder (e.g. hypothyroidism, vit. B12 or folic acid deficiency, hypercalcaemia), or alcohol or drug abuse.
F02.8 Dementia in other specified diseases classified elsewhere
Dementia can occur as a manifestation or consequence of a variety of cerebral and somatic conditions. To specify the etiology, the ICD-10 code for the underlying condition should be added.
F03 UNSPECIFIED DEMENTIA
This category should be used when the general criteria for dementia are met, but when it is not possible to identify one of the specific types (F00.0-F02.9).

**Diagnostic and Statistical Manual of Mental Disorders, 4th Edition, Text Revision**

**Dementia Codes**

Dementia of the Alzheimer's Type, with early onset
294.10 Without behavioural disturbance
294.11 With behavioural disturbance
Dementia of the Alzheimer's Type, with late onset
294.10 Without behavioural disturbance
294.11 With behavioural disturbance


**Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.**

Vascular dementia
290.40 Uncomplicated
290.41 With delirium
290.42 With delusions
290.43 With depressed mood
Dementia due to HIV disease
294.10 Without behavioural disturbance
294.11 With behavioural disturbance
Dementia due to head trauma
294.10 Without behavioural disturbance
294.11 With behavioural disturbance
Dementia due to Parkinson's disease
294.10 Without behavioural disturbance
294.11 With behavioural disturbance
Dementia due to Huntington's disease
294.10 Without behavioural disturbance
294.11 With behavioural disturbance
Dementia due to Pick's disease
294.10 Without behavioural disturbance
294.11 With behavioural disturbance
Dementia due to Creutzfeldt-Jakob Disease
294.10 Without behavioural disturbance
294.11 With behavioural disturbance
Dementia due to... [indicate other general medical condition]
294.10 Without behavioural disturbance
294.11 With behavioural disturbance
294.8 Dementia NOS

## Appendix 2. AD8 Informant Questionnaire

| Remember, "Yes, a change" indicates that there has been a change in the last several years caused by cognitive (thinking and memory) problems | **YES**<br>**A change** | **NO**<br>**No change** | **N/A**<br>**Don't know** |
|---|---|---|---|
| **1.** Problems with judgment (e.g., problems making decisions, bad financial decisions, problems with thinking) | | | |
| **2.** Less interest in hobbies/activities | | | |
| **3.** Repeats the same things over and over (questions, stories, or statements) | | | |


Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

(*Continued*)

| | | | |
|---|---|---|---|
| **4.** Trouble learning how to use a tool, appliance, or gadget (e.g., VCR, computer, microwave, remote control) | | | |
| **5.** Forgets correct month or year | | | |
| **6.** Trouble handling complicated financial affairs (e.g. balancing checkbook, income taxes, paying bills) | | | |
| **7.** Trouble remembering appointments | | | |
| **8.** Daily problems with thinking and/or memory | | | |
| **TOTAL AD-8 SCORE** | | | |

## Appendix 3. AD-8 administration and scoring guidelines

*A spontaneous self-correction is allowed for all responses without counting as an error.*

The questions are given to the respondent on a clipboard for self-administration or can be read aloud to the respondent either in person or over the phone. It is preferable to administer the AD-8 to an informant, if available. If an informant is not available, the AD-8 may be administered to the patient.

When administered to an informant, specifically ask the respondent to rate change in the patient.

When administered to the patient, specifically ask the patient to rate changes in his/her ability for each of the items, ***without*** attributing causality.

If read aloud to the respondent, it is important for the clinician to carefully read the phrase as worded and give emphasis to note changes due to cognitive problems (not physical problems).

There should be a one second delay between individual items.

No timeframe for change is required.

The final score is a sum of the number items marked "Yes, A change".

## Appendix 4. Commonly used cognitive assessments/screening tools

| TEST | Cochrane DTA review in process |
|---|---|
| Mini-mental state examination (MMSE) | YES |
| GPcog | YES |
| Minicog | YES |


**Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.**

(*Continued*)

| | |
|---|---|
| Memory Impairment Screen (MIS) | Still available |
| Abbreviated mental testing | Still available |
| Clock drawing tests (CDT) | Still available |
| Montreal Cognitive Assessment (MoCA) | YES |
| IQCODE (informant interview) | YES |

For each test, the planned review will encompass diagnostic test accuracy in community; primary and secondary care settings.

## Appendix 5. Sources searched and search strategies

| Source | Search strategy | Hits retrieved |
|---|---|---|
| ALOIS DTA<br>[Date of most recent search: 7 June 2018] | AD8 OR AD-8 OR "AD 8" OR "Alzheimer's Disease eight question screen"<br>See Appendix 6 for the searches run to populate the ALOIS register | May 2014: 15<br>May 2017: 3<br>June 2018: 0 |
| MEDLINE In-process and other non-indexed citations and MEDLINE 1950-present (Ovid SP)<br>[Date of most recent search: 7 June 2018] | 1. AD8.ti,ab.<br>2. "informant questionnaire on cognitive decline".ti,ab.<br>3. "Alzheimer's Disease eight question screen".ti,ab.<br>4. "AD 8".ti,ab.<br>5. ("informant* questionnair*" adj3 (dement* or screening)).ti,ab<br>6. ("screening test*" adj2 (dement* or alzheimer*)).ti,ab.<br>7. or/1-6 | May 2014: 300<br>May 2017: 54<br>June 2018: 48 |
| EMBASE<br>1974-2018 June 6 (Ovid SP)<br>[Date of most recent search: 7 June 2018] | 1. AD8.ti,ab.<br>2. "informant questionnaire on cognitive decline".ti,ab.<br>3. "Alzheimer's Disease eight question screen".ti,ab.<br>4. "AD 8".ti,ab.<br>5. ("informant* questionnair*" adj3 (dement* or screening)).ti,ab<br>6. ("screening test*" adj2 (dement* or alzheimer*)).ti,ab.<br>7. or/1-6 | May 2014: 487<br>May 2017: 75<br>June 2018: 58 |


**Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.**

(*Continued*)

| | | |
|---|---|---|
| PSYCINFO<br>1806-June week 1 2018 (Ovid SP)<br>[Date of most recent search: 7 June 2018] | 1. AD8.ti,ab.<br>2. "informant questionnaire on cognitive decline".ti,ab.<br>3. "Alzheimer's Disease eight question screen".ti,ab.<br>4. "AD 8".ti,ab.<br>5. ("informant* questionnair*" adj3 (dement* or screening)).ti,ab<br>6. ("screening test*" adj2 (dement* or alzheimer*)).ti,ab.<br>7. or/1-6 | May 2014: 159<br>May 2017: 66<br>June 2018: 31 |
| CINAHL (EBSCOhost)<br>[Date of most recent search: 7 June 2018] | S13 S1 OR S12<br>S12 S10 AND S11<br>S11 TX informer* OR informant*<br>S10 S5 AND S9<br>S9 S6 OR S7 OR S8<br>S8 (MM "Alzheimer's Disease/DI")<br>S7 (MM "Dementia") OR (MM "Delirium, Dementia, Amnestic, Cognitive Disorders")<br>S6 TX dement* OR alzheimer* OR cognit*<br>S5 S2 OR S3 OR S4<br>S4 TX (cognit* OR memory) n3 screen*<br>S3 TX "neuropsychological test*" OR "neurocog* test*" OR "cog* test*"<br>S2 TX "neuropsychological assess*" OR "neurocog* assess*" OR "cog* assess*"<br>S1 TX AD8 | May 2014: 219<br>May 2017: 49<br>June 2018: 29 |
| Biosis previews 1926 to present (ISI Web of Science)<br>[Date of most recent search: 7 June 2018] | (AD8 or "informant questionnaire on cognitive decline" or "Alzheimer's Disease eight question screen" or "AD 8" or ("informant* questionnair*" and dement*) or ("screening test*" and dement*) or ("informant* questionnair*" and alzheimer*) or ("screening test*" and alzheimer*)) | May 2014: 710<br>May 2017: 104<br>June 2018: 93 |
| Web of Science Core Collection (1945-present) (ISI Web of Science)<br>[Date of most recent search: 7 June 2018] | (AD8 or "informant questionnaire on cognitive decline" or "Alzheimer's Disease eight question screen" or "AD 8" or ("informant* questionnair*" and dement*) or ("screening test*" and dement*) or ("informant* questionnair*" and alzheimer*) or ("screening test*" and alzheimer*)) | May 2014: 869<br>May 2017: 188<br>June 2018: 101 |
| LILACS (BIREME)<br>[Date of most recent search: 7 June 2018] | (AD8 or "informant questionnaire on cognitive decline" or "Alzheimer's Disease eight question screen" or "AD 8" | May 2014: 1<br>May 2017: 1<br>June 2018: 1 |


**Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.**

(*Continued*)

| | | |
|---|---|---|
| MEDION database (Meta-analyses van Diagnostisch Onderzoek www.mediondatabase.nl); DARE (Database of Abstracts of Reviews of Effects) and HTA Database (Health Technology Assessments Database), both *The Cochrane Library*; ARIF database (Aggressive Research Intelligence Facility; www.arif.bham.ac.uk). [Date of most recent search: 7 June 2018] | (AD8 or "informant questionnaire on cognitive decline" or "Alzheimer's Disease eight question screen" or "AD 8" | June 2018: 0 |
| TOTAL before de-duplication | | May 2014: 2760<br>May 2017: 540<br>June 2018: 361<br>TOTAL: 3661 |
| TOTAL after de-duplication and first-assessment by CDCIG information specialist based on titles and abstracts | | May 2014: 447<br>May 2017: 162<br>June 2018: 5<br>TOTAL: 614 |

## Appendix 6. Search strategy (MEDLINE Ovid SP) run for specialised register (ALOIS)

Search narrative: The searches detailed above are very simple, essentially single concept strategies based on the index test (AD-8). This is a sensitive approach to take. More complex and developed searches are run each month for the dementia group.
Every month the following strategy is run in Medline (via Ovid SP). The results are screened based on a reading of title and abstract. The full texts (where there is one) are then obtained and a few key details about each study are extracted including Index test/s and details of population and setting. For this review it was expected that most studies would be identified through a search of multiple sources based on one concept (the index test in question). However, we felt it was worth also searching ALOIS for any studies which had evaluated the accuracy of AD-8 but had not referred to it in the bibliographic details of the reference.


**Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.**

| | |
|---|---|
| MEDLINE In-process and other non-indexed citations and MEDLINE 1950-present (Ovid SP) | 1. "word recall".ti,ab.<br>2. "7-minute screen".ti,ab.<br>3. "6 item cognitive impairment test".ti,ab.<br>4. "6 CIT".ti,ab.<br>5. "AB cognitive screen".ti,ab.<br>6. "abbreviated mental test".ti,ab.<br>7. "ADAS-cog".ti,ab.<br>8. AD8.ti,ab.<br>9. "inform* interview".ti,ab.<br>10. "animal fluency test".ti,ab.<br>11. "brief alzheimer* screen".ti,ab.<br>12. "brief cognitive scale".ti,ab.<br>13. "clinical dementia rating scale".ti,ab.<br>14. "clinical dementia test".ti,ab.<br>15. "community screening interview for dementia".ti,ab.<br>16. "cognitive abilities screening instrument".ti,ab.<br>17. "cognitive assessment screening test".ti,ab.<br>18. "cognitive capacity screening examination".ti,ab.<br>19. "clock drawing test".ti,ab.<br>20. "deterioration cognitive observee".ti,ab.<br>21. "Dem Tect".ti,ab.<br>22. "fuld object memory evaluation".ti,ab.<br>23. "IQCODE".ti,ab.<br>24. "mattis dementia rating scale".ti,ab.<br>25. "memory impairment screen".ti,ab.<br>26. "minnesota cognitive acuity screen".ti,ab.<br>27. "mini-cog".ti,ab.<br>28. "mini-mental state exam*".ti,ab.<br>29. "mmse".ti,ab.<br>30. "modified mini-mental state exam".ti,ab.<br>31. "3MS".ti,ab.<br>32. "neurobehavioural cognitive status exam*".ti,ab.<br>33. "cognistat".ti,ab.<br>34. "quick cognitive screening test".ti,ab.<br>35. "QCST".ti,ab.<br>36. "rapid dementia screening test".ti,ab.<br>37. "RDST".ti,ab.<br>38. "repeatable battery for the assessment of neuropsychological status".ti,ab<br>39. "RBANS".ti,ab.<br>40. "rowland universal dementia assessment scale".ti,ab.<br>41. "rudas".ti,ab.<br>42. "self-administered gerocognitive exam*".ti,ab.<br>43. ("self-administered" and "SAGE").ti,ab.<br>44. "self-administered computerized screening test for dementia".ti,ab<br>45. "short and sweet screening instrument".ti,ab.<br>46. "sassi".ti,ab.<br>47. "short cognitive performance test".ti,ab. |


**Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.**

48. “syndrome kurztest”.ti,ab.
49. “six item screener”.ti,ab.
50. “short memory questionnaire”.ti,ab.
51. (“short memory questionnaire” and “SMQ”).ti,ab.
52. “short orientation memory concentration test”.ti,ab.
53. “s-omc”.ti,ab.
54. “short blessed test”.ti,ab.
55. “short portable mental status questionnaire”.ti,ab.
56. “spmsq”.ti,ab.
57. “short test of mental status”.ti,ab.
58. “telephone interview of cognitive status modified”.ti,ab
59. “tics-m”.ti,ab.
60. “trail making test”.ti,ab.
61. “verbal fluency categories”.ti,ab.
62. “WORLD test”.ti,ab.
63. “general practitioner assessment of cognition”.ti,ab.
64. “GPCOG”.ti,ab.
65. “Hopkins verbal learning test”.ti,ab.
66. “HVLT”.ti,ab.
67. “time and change test”.ti,ab.
68. “modified world test”.ti,ab.
69. “symptoms of dementia screener”.ti,ab.
70. “dementia questionnaire”.ti,ab.
71. “7MS”.ti,ab.
72. (“concord informant dementia scale” or CIDS).ti,ab.
73. (SAPH or “dementia screening and perceived harm*”).ti,ab
74. or/1-73
75. exp Dementia/
76. Delirium, Dementia, Amnestic, Cognitive Disorders/ OR Cognition Disorders/ OR Memory Disorders/
77. dement*.ti,ab.
78. alzheimer*.ti,ab.
79. AD.ti,ab.
80. (“lewy bod*” or DLB or LBD).ti,ab.
81. “cognit* impair*”.ti,ab.
82. (cognit* adj4 (disorder* or declin* or fail* or function*)).ti,ab
83. (memory adj3 (complain* or declin* or function*)).ti,ab.
84. or/75-83
85. exp “sensitivity and specificity”/
86. “reproducibility of results”/
87. (predict* adj3 (dement* or AD or alzheimer*)).ti,ab.
88. (identif* adj3 (dement* or AD or alzheimer*)).ti,ab.
89. (discriminat* adj3 (dement* or AD or alzheimer*)).ti,ab.
90. (distinguish* adj3 (dement* or AD or alzheimer*)).ti,ab.
91. (differenti* adj3 (dement* or AD or alzheimer*)).ti,ab.
92. diagnos*.ti.
93. di.fs.
94. sensitivit*.ab.


**Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.**

| | |
|---|---|
| | 95. specificit*.ab.<br>96. (ROC or "receiver operat*").ab.<br>97. Area under curve/<br>98. ("Area under curve" or AUC).ab.<br>99. (detect* adj3 (dement* or AD or alzheimer*)).ti,ab.<br>100. sROC.ab.<br>101. accura*.ti,ab.<br>102. (likelihood adj3 (ratio* or function*)).ab.<br>103. (conver* adj3 (dement* or AD or alzheimer*)).ti,ab.<br>104. ((true or false) adj3 (positive* or negative*)).ab.<br>105. ((positive* or negative* or false or true) adj3 rate*).ti,ab<br>106. or/85-105<br>107. exp dementia/di OR Delirium, Dementia, Amnestic, Cognitive Disorders/di<br>108. Cognition Disorders/di<br>109. Memory Disorders/di<br>110. or/107-109<br>111. *Neuropsychological Tests/<br>112. *Questionnaires/<br>113. Geriatric Assessment/mt<br>114. *Geriatric Assessment/<br>115. Neuropsychological Tests/mt, st<br>116. "neuropsychological test*".ti,ab.<br>117. (neuropsychological adj (assess* or evaluat* or test*)).ti,ab<br>118. (neuropsychological adj (assess* or evaluat* or test* or exam* or battery)).ti,ab<br>119. Self report/<br>120. self-assessment/ or diagnostic self evaluation/<br>121. Mass Screening/<br>122. early diagnosis/<br>123. or/111-122<br>124. 74 or 123<br>125. 110 and 124<br>126. 74 or 123<br>127. 84 and 106 and 126<br>128. 74 and 106<br>129. 125 or 127 or 128<br>130. (animals not (humans and animals)).sh.<br>131. 129 not 130<br>The concepts for this are:<br>**A** Specific neuropsychological tests<br>**B** General terms (both free text and MeSH) for tests/testing/screening<br>**C** Outcome: dementia diagnosis (unfocused MeSH with diagnostic sub-headings)<br>**D** Condition of interest: Dementia (general dementia terms both free text and MeSH - exploded and unfocused)<br>**E** Methodological filter: not used to limit all search |


**Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.**

| | |
|---|---|
| | The concept combinations are:<br>1. (A OR B) AND C<br>2. (A OR B) AND D AND E<br>3. A AND E |

## Appendix 7. Example translation proforma

| | |
|---|---|
| Study ID | |
| Population and setting | **Which of the following does the study best fall into?**<br>a) Asymptomatic community<br>b) Symptomatic community<br>c) Asymptomatic primary<br>d) Symptomatic primary<br>e) Secondary care (memory focus)<br>f) Secondary care (non-memory focus)<br>g) Mixed<br>h) None of the above<br>If h) please describe the setting population: |
| Study design: Participants | **a) How were they sampled?**<br>a. Randomly (cluster sampling)<br>b. Consecutively<br>c. Not reported<br>d. Other (please describe)<br>**b) Was a case control design avoided?**<br>a. Yes<br>b. No<br>c. Not sure (please elaborate if possible)<br>**c) Please describe any exclusion criteria stated in the paper** |
| Participant characteristics | Of the participants included in the study:<br>a) **What was their mean age?** Please say Not reported if this is not reported in the paper:<br>b) **What percentage of the participants were female?** Please say Not reported if this is not reported in the paper or easy to calculate from the information given:<br>c) **What was the average level of education of the participants?** Please just state whatever information is in the paper regarding the educational level of the participants<br>**What was the mean AD-8 score?** Please say Not reported if this is not reported in the paper or easy to calculate from the information given<br>d) **If the study concerns MCI "progression" how was MCI defined?** Give classification criteria eg Peterson, say not reported if this is not described<br>e) **How many participants were included in the study?** |


Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

(*Continued*)

| | |
|---|---|
| Target condition + reference standard | [Target condition should be dementia. It may be a form of dementia such as Alzheimer's. The reference standard is likely to be DSM-III or DSM-IV: just note down whatever diagnostic criteria they appeared to use to diagnose dementia.]<br>**a) What reference standard has been used?**<br>**b) Did all the participants have the reference standard?**<br>a. Yes<br>b. No. Please describe how those who did have the reference standard had been selected and state how many had it (if this is not reported please say Not reported)<br>**Please provide a brief description/translation of how the reference standard was carried out and by whom:** |
| | [If it's not AD-8 then it's out!]<br>**Was this a translated version of the AD-8 that was used?** If the AD-8 used was translated or modified in any way, is there any evidence that the version used had been validated, or was this study the validation of that version?<br>**Who conducted the Index test?**<br>a) Trained specialist<br>b) Trained non-specialist<br>**Please provide a brief description/translation of how the Index test was carried out:** |
| Duration of time between index test and reference standard | [please put Not reported if this information is not provided - it quite often isn't]<br>a)**What was the time period between the conduct of the tests?**<br>b) **Are there any participants who appear to be unaccounted for?** (you can Unsure if you are not sure about this) |
| Results | **Is the sensitivity and specificity of the AD-8 against the reference standard reported or extractable do you think?**<br>If yes, what cut point on the AD-8 has been used (there may be several cut points used - please list them all. It might be best to give a fairly detailed translation of the results section because we will need to ascertain which score on the AD-8 indicates normality)<br>**If you are able to extract the data, please do so**. Ideally we would like the four numbers for each cutpoint reported: True Positive, True Negative, False Positive, False Negative |


**Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.**

## Appendix 8. Assessment of methodological quality table QUADAS-2 tool

| DOMAIN | PATIENT SELECTION | INDEX TEST | REFERENCE STANDARD | FLOW AND TIMING |
|---|---|---|---|---|
| Description | Describe methods of patient selection: Describe included patients (prior testing, presentation, intended use of index test and setting): | Describe the index test and how it was conducted and interpreted: | Describe the reference standard and how it was conducted and interpreted: | Describe any patients who did not receive the index test(s) and/or reference standard or who were excluded from the 2x2 table (refer to flow diagram): Describe the time interval and any interventions between index test(s) and reference standard: |
| Signalling questions (yes/no/unclear) | Was a consecutive or random sample of patients enrolled? | Were the index test results interpreted without knowledge of the results of the reference standard? | Is the reference standard likely to correctly classify the target condition? | Was there an appropriate interval between index test(s) and reference standard? |
| | Was a case-control design avoided? | If a threshold was used, was it pre-specified? | Were the reference standard results interpreted without knowledge of the results of the index test? | Did all patients receive a reference standard? |
| | Did the study avoid inappropriate exclusions? | | | Did all patients receive the same reference standard? |
| | | | | Were all patients included in the analysis? |
| Risk of bias: High/low/unclear | Could the selection of patients have introduced bias? | Could the conduct or interpretation of the index test have introduced bias? | Could the reference standard, its conduct, or its interpretation have introduced bias? | Could the patient flow have introduced bias? |
| Concerns regarding applicability: High/low/unclear | Are there concerns that the included patients do not match the review question? | Are there concerns that the index test, its conduct, or interpretation differ from the review question? | Are there concerns that the target condition as defined by the reference standard does not match the review question? | |


Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

## Appendix 9. Anchoring statements for quality assessment of AD-8 diagnostic studies

We provide some core anchoring statements for quality assessment of diagnostic test accuracy reviews of AD-8 in dementia. These statements are designed for use with the QUADAS-2 tool and were derived during a two-day, multidisciplinary focus group (Davis 2013).

During the focus group and the piloting/validation of this guidance, it was clear that certain issues were key to assessing quality, while other issues were important to record but less important for assessing overall quality. To assist, we describe a system wherein certain items can dominate. For these dominant items, if scored "high risk" then that section of the QUADAS-2 results table is likely to be scored as high risk of bias regardless of other scores. For example, in dementia diagnostic test accuracy studies, ensuring that clinicians performing dementia assessment are blinded to results of index test is fundamental. If this blinding was not present then the item on reference standard should be scored "high risk of bias", regardless of the other contributory elements.

We have detailed how QUADAS-2 has been operationalised for use with dementia reference standard studies below. In these descriptors dominant items are labelled as "high risk of bias for total section regardless of other items".

In assessing individual items, the score of unclear should only be given if there is genuine uncertainty. In these situations review authors will contact the relevant study teams for additional information.

#### Anchoring statements to assist with assessment for risk of bias

#### Selection

**Was a case-control or similar design avoided?**

Designs similar to case-control that may introduce bias are those designs where the study team deliberately increase or decrease the proportion with the target condition. For example, a population study may be enriched with extra dementia patients from a secondary care setting. Such studies will be automatically labelled high risk of bias and this will be assessed as a potential source of heterogeneity. If case-control is used then grading will be high risk of bias for total section regardless of other items (in fact case-control studies will not be included in this review).

**Was the sampling method appropriate?**

Where sampling is used, the designs least likely to cause bias are consecutive sampling or random sampling. Sampling that is based on volunteers or selecting participants from a clinic or research resource is prone to bias.

**Are exclusion criteria described and appropriate?**

The study will be automatically graded as unclear if exclusions are not detailed (pending contact with study authors). Where exclusions are detailed, the study will be graded as low risk of bias if exclusions are felt to be appropriate by the review authors. Certain exclusions common to many studies of dementia are: medical instability; terminal disease; alcohol/substance misuse; concomitant psychiatric diagnosis; other neurodegenerative condition.

Post hoc exclusions will be labelled high risk of bias for total section regardless of other items.

#### Index Test

**Was AD-8 assessment performed without knowledge of clinical dementia diagnosis?**

Terms such as "blinded" or "independently and without knowledge of" are sufficient and full details of the blinding procedure are not required. This item may be scored as low risk of bias if explicitly described or if there is a clear temporal pattern to order of testing that precludes the need for formal blinding i.e. all AD-8 assessments performed before dementia assessment.

If there is no attempt at blinding grading will be high risk of bias for total section regardless of other items.


Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

**Were AD-8 thresholds prespecified?**

For scales there is often a reference point (in units or categories) above which participants are classified as "test positive"; this may be referred to as threshold; clinical cut-off or dichotomisation point. A study is classified high risk of bias if the authors define the optimal cut-off post-hoc based on their own study data. Certain papers may use an alternative methodology for analysis that does not use thresholds and these papers should be classified as low risk of bias.

**Were sufficient data on AD-8 application given for the test to be repeated in an independent study?**

Particular points of interest for AD-8 include method of administration (for example, self-completed questionnaire versus direct questioning interview); nature of informant; language of assessment. If a novel form of AD-8 is used, details of the scale should be included or a reference given to an appropriate descriptive text. Where AD-8 is used in a novel manner, for example, a translated questionnaire, there should be evidence of validation work.

#### Reference Standard

**Is the assessment used for clinical diagnosis of dementia acceptable?**

Commonly used international criteria to assist with clinical diagnosis of dementia include those detailed in DSM-IV and ICD-10. Criteria specific to dementia subtypes include but are not limited to NINCDS-ADRDA criteria for Alzheimer's dementia; McKeith criteria for Lewy Body dementia; Lund criteria for frontotemporal dementias; and the NINDS-AIREN criteria for vascular dementia. We included use of the Clinical Dementia Rating (CDR) assessment as an accepetd method of dementai diagnosis. Where the criteria used for assessment are not familiar to the review authors or the Cochrane Dementia and Cognitive Improvement Group this item should be classified as high risk of bias.

**Was clinical assessment for dementia performed without knowledge of AD-8?**

Terms such as "blinded" or "independent" are sufficient and full details of the blinding procedure are not required. This may be scored as low risk of bias if explicitly described or if there is a clear temporal pattern to order of testing, i.e. all dementia assessments performed before AD-8 testing.

Informant rating scales and direct cognitive tests present certain problems. It is accepted that informant interview and cognitive testing is a usual component of clinical assessment for dementia, however, specific use of the scale under review in the clinical dementia assessment should be scored as high risk of bias. We have prespecified that dementia diagnosis that explicitly uses AD-8 will be classified as high risk of bias for total section regardless of other items.

**Were sufficient data on dementia assessment method given for the assessment to be repeated in an independent study?**

The criteria used for clinical assessment are discussed in another item. Particular points of interest for dementia assessment include the background of the assessor, training/expertise of the assessor; additional information available to inform diagnosis (neuroimaging; neuropsychological testing).

#### Flow

**Was there an appropriate interval between AD-8 and clinical dementia assessment?**

For a cross-sectional study design, there is potential for change between assessments. The ideal would be same day assessment but this is not always feasible. We have set an arbitrary maximum interval of one month between tests, although this may be revised depending on the test and the stability of the condition of interest.

For a study looking at delayed verification there is no agreement on how long the interval should be between index test and first/last assessment for dementia. An interval of less than six months is unlikely to be sufficient time for progression.


**Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.**

**Did all get the same assessment for dementia regardless of AD-8 result?**

There may be scenarios where only those who score "test positive" on AD-8 have a more detailed assessment. Where dementia assessment (or other reference standard) differs depending on the AD-8 result this should be classified as high risk of bias.

**Were all who received AD-8 assessment included in the final analysis?**

If the study has drop outs these should be accounted for; a maximum proportion of drop outs to remain low risk of bias has been specified as 20%.

**Were missing AD-8 results or un-interpretable AD-8 results reported?**

Where missing results are reported if there is substantial attrition (we have set an arbitrary value of 50% missing data) this should be scored as high risk of bias for total section regardless of other items.

#### Applicability

**Were those included representative of the general population of interest?**

Those included should match the intended population as described in the review question. If not already specified in the review inclusion criteria, setting will be particularly important - the review authors should consider population in terms of symptoms; pre-testing; potential disease prevalence. Studies that use very selected groups or subgroups will be classified as poor applicability.

**Was AD-8 performed consistently and in a manner similar to its use in clinical practice?**

AD-8 studies will be judged against the original description of its use.

**Was clinical diagnosis of dementia (or other reference standard) made in a manner similar to current clinical practice?**

For many reviews, inclusion criteria and assessment for risk of bias will already have assessed the dementia diagnosis. For certain reviews an applicability statement relating to reference standard may not be applicable. There is the possibility that a form of dementia assessment, although valid, may diagnose a far larger proportion with disease than would be seen in usual clinical practice. In this instance the item should be rated poor applicability.

## CONTRIBUTIONS OF AUTHORS

KH performed title selection, data extraction, 'Risk of bias' assessments, and provided comments on the final manuscript.

CG performed title selection, data extraction, 'Risk of bias' assessments, and provided comments on the final manuscript.

RM assisted with title and protocol development, comments on the final manuscript, and supervision.

ANS assisted with creation of the search strategy, initial title searching, and provided comments on the final manuscript.

DJS assisted with title and protocol development, comments on the final manuscript, and supervision.

SA assisted with data extraction, 'Risk of bias' assessments, analyses, and provided comments on the final manuscript.

AJS assisted with analyses, and provided comments on the final manuscript.

JKB assisted with data synthesis and drafting the manuscript.

TJQ drafted the protocol, assisted with data synthesis, drafted the manuscript and supervised.


**Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.**

## DECLARATIONS OF INTEREST

KH: none known

CG: none known

RM: none known

ANS: none known

DJS: none known

SA: none known

AJS: none known

JKB: none known

TJQ: none known

## SOURCES OF SUPPORT


### Internal sources

- No sources of support supplied

### External sources

- Terry Quinn, UK.

TQ is supported by the joint Chief Scientist Office and Stroke Association Senior Clinical Lectureship

- Jennifer Burton, UK.

JKB is supported by an NHS Education for Scotland, Scottish Clinical Research Excellence Development Scheme Clinical Lectureship

- NIHR, UK.

This review was supported by the National Institute for Health Research (NIHR), via Cochrane Infrastructure funding 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

The title of the review has been changed from "AD-8 for diagnosis of dementia across a variety of healthcare settings" to "AD-8 for detection of dementia across a variety of healthcare settings". This change is consistent with the use of AD-8 described in the review - for research and clinical practice - as a tool to identify those who may benefit from diagnostic assessment, rather than as a single diagnostic test.

The secondary objectives text has been altered between the protocol and the review, to be consistent with the other test accuracy reviews of informant tools. The review specifies in the secondary objectives that the diagnostic accuracy has been described at various predefined threshold scores. The review does not describe the accuracy of AD-8 for early detection of cognitive problems, where a later diagnosis of dementia is made (delayed verification diagnosis) as no studies using this methodology were identified in the search.

The modest number of papers limited the sensitivity and subgroup analyses that we could perform. For example, we had planned on having summary analyses for each individual healthcare setting as primary analysis, with an analysis across all healthcare settings as a secondary analysis. Numbers of papers only allowed a summary analysis across all settings at a cut-off score of two.

We had planned on assessing reporting quality using STARD. We have performed this but have not incorporated these data into the review. We note that reporting guidance for all test accuracy and for studies with a dementia outcome has evolved from the original iteration of STARD.



**Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.**