Mini-Cog for the diagnosis of Alzheimer’s disease dementia and other dementias within a secondary care setting (Review)

Chan CCH, Fage BA, Burton JK, Smailagic N, Gill SS, Herrmann N, Nikolaou V, Quinn TJ, Noel-Storr AH, Seitz DP

DOI: 10.1002/14651858.CD011414.pub2.

www.cochranelibrary.com
# Table of Contents

HEADER .................................................. 1
ABSTRACT .................................................. 1
PLAIN LANGUAGE SUMMARY .............................. 2
SUMMARY OF FINDINGS FOR THE MAIN COMPARISON .... 4
BACKGROUND .............................................. 6
OBJECTIVES .............................................. 7
METHODS ................................................. 8
RESULTS .................................................. 10
  Figure 1. .................................................. 11
  Figure 2. .................................................. 12
  Figure 3. .................................................. 13
DISCUSSION .............................................. 14
AUTHORS’ CONCLUSIONS .................................. 14
ACKNOWLEDGEMENTS ..................................... 15
REFERENCES .............................................. 15
CHARACTERISTICS OF STUDIES .......................... 19
DATA ....................................................... 27
  Test 1. Mini-Cog in secondary care. .................... 27
ADDITIONAL TABLES ..................................... 27
APPENDICES .............................................. 29
CONTRIBUTIONS OF AUTHORS ........................... 36
DECLARATIONS OF INTEREST ............................ 37
SOURCES OF SUPPORT ................................... 37
DIFFERENCES BETWEEN PROTOCOL AND REVIEW ........ 37
Mini-Cog for the diagnosis of Alzheimer’s disease dementia and other dementias within a secondary care setting

Calvin CH Chan¹, Bruce A Fage², Jennifer K Burton³, Nadja Smialagic⁴, Sudeep S Gill⁵, Nathan Herrmann⁶, Vasilis Nikolaou⁷, Terry J Quinn⁸, Anna H Noel-Storr⁹, Dallas P Seitz¹⁰

¹School of Medicine, Queen’s University, Kingston, Canada. ²Department of Psychiatry, University of Toronto, Toronto, Canada. ³Academic Geriatric Medicine, Institute of Cardiovascular and Medical Sciences, University of Glasgow, Glasgow, UK. ⁴Institute of Public Health, University of Cambridge, Cambridge, UK. ⁵Department of Medicine, Queen’s University, Kingston, Canada. ⁶Hurvitz Brain Sciences Research Program, Sunnybrook Research Institute, Toronto, Canada. ⁷Research In Real Life, Cambridge, UK. ⁸Institute of Cardiovascular and Medical Sciences, University of Glasgow, Glasgow, UK. ⁹Radcliffe Department of Medicine, University of Oxford, Oxford, UK. ¹⁰Department of Psychiatry, Queen’s University, Kingston, Canada

Contact address: Dallas P Seitz, Department of Psychiatry, Queen’s University, 752 King Street West, Kingston, ON, K7L 4X3, Canada. seitzd@providencecare.ca.

Editorial group: Cochrane Dementia and Cognitive Improvement Group.


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

ABSTRACT

Background
The diagnosis of Alzheimer’s disease dementia and other dementias relies on clinical assessment. There is a high prevalence of cognitive disorders, including undiagnosed dementia in secondary care settings. Short cognitive tests can be helpful in identifying those who require further specialist diagnostic assessment; however, there is a lack of consensus around the optimal tools to use in clinical practice. The Mini-Cog is a short cognitive test comprising three-item recall and a clock-drawing test that is used in secondary care settings.

Objectives
The primary objective was to determine the diagnostic accuracy of the Mini-Cog for detecting Alzheimer’s disease dementia and other dementias in secondary care settings. The secondary objectives were to investigate the heterogeneity of test accuracy in the included studies and potential sources of heterogeneity. These potential sources of heterogeneity will include the baseline prevalence of dementia in study samples, thresholds used to determine positive test results, the type of dementia (Alzheimer’s disease dementia or all causes of dementia), and aspects of study design related to study quality.

Search methods
We searched the following sources in September 2012, with an update to 12 March 2019: Cochrane Dementia Group Register of Diagnostic Test Accuracy Studies, MEDLINE (OvidSP), Embase (OvidSP), BIOSIS Previews (Web of Knowledge), Science Citation Index (ISI Web of Knowledge), PsycINFO (OvidSP), and LILACS (BIREME). We made no exclusions with regard to language of Mini-Cog administration or language of publication, using translation services where necessary.

Selection criteria
We included cross-sectional studies and excluded case-control designs, due to the risk of bias. We selected those studies that included the Mini-Cog as an index test to diagnose dementia where dementia diagnosis was confirmed with reference standard clinical assessment.
using standardised dementia diagnostic criteria. We only included studies in secondary care settings (including inpatient and outpatient hospital participants).

Data collection and analysis

We screened all titles and abstracts generated by the electronic database searches. Two review authors independently checked full papers for eligibility and extracted data. We determined quality assessment (risk of bias and applicability) using the QUADAS-2 tool. We extracted data into two-by-two tables to allow calculation of accuracy metrics for individual studies, reporting the sensitivity, specificity, and 95% confidence intervals of these measures, summarising them graphically using forest plots.

Main results

Three studies with a total of 2560 participants fulfilled the inclusion criteria, set in neuropsychology outpatient referrals, outpatients attending a general medicine clinic, and referrals to a memory clinic. Only n = 1415 (55.3%) of participants were included in the analysis to inform evaluation of Mini-Cog test accuracy, due to the selective use of available data by study authors. There were concerns related to high risk of bias with respect to patient selection, and unclear risk of bias and high concerns related to index test conduct and applicability. In all studies, the Mini-Cog was retrospectively derived from historic data sets. No studies included acute general hospital inpatients. The prevalence of dementia ranged from 32.2% to 87.3%. The sensitivities of the Mini-Cog in the individual studies were reported as 0.67 (95% confidence interval (CI) 0.63 to 0.71), 0.60 (95% CI 0.48 to 0.72), and 0.87 (95% CI 0.83 to 0.90). The specificity of the Mini-Cog for each individual study was 0.87 (95% CI 0.81 to 0.92), 0.65 (95% CI 0.57 to 0.73), and 1.00 (95% CI 0.94 to 1.00). We did not perform meta-analysis due to concerns related to risk of bias and heterogeneity.

Authors' conclusions

This review identified only a limited number of diagnostic test accuracy studies using Mini-Cog in secondary care settings. Those identified were at high risk of bias related to patient selection and high concerns related to index test conduct and applicability. The evidence was indirect, as all studies evaluated Mini-Cog differently from the review question, where it was anticipated that studies would conduct Mini-Cog and independently but contemporaneously perform a reference standard assessment to diagnose dementia. The pattern of test accuracy varied across the three studies. Future research should evaluate Mini-Cog as a test in itself, rather than derived from other neuropsychological assessments. There is also a need for evaluation of the feasibility of the Mini-Cog for the diagnosis of dementia to help adequately determine its role in the clinical pathway.

PLAIN LANGUAGE SUMMARY

How accurate is the Mini-Cog in detecting dementia among patients in inpatient and outpatient hospital settings?

Why is recognising dementia important?

Dementia is a common and important condition, and many of those living with dementia have never had the condition diagnosed. Diagnosis provides opportunities for social support, advance care planning and, in specific disease types, treatment with medication. However, incorrectly diagnosing dementia when it is not present (a false-positive result) can be distressing for the individual and their family and lead to a waste of resources in diagnostic tests.

What was the aim of the review?

The aim of this Cochrane Review was to find out how accurate the Mini-Cog test is for detecting dementia among patients in inpatient and outpatient hospital settings. The researchers included three studies to answer this question.

What was studied in the review?

The Mini-Cog is a short test of memory and thinking skills that tests the ability of an individual to remember three specific objects, named at the beginning of a short assessment, repeated at the time and recalled by the individual later. In addition, the individual being assessed is asked to draw a clock face at a specific time. Points are scored based on the ability to recall the three items and the completeness of the clock. The Mini-Cog is a short test that would typically be used to identify if someone was having difficulty with memory and thinking skills who would benefit from referral to a specialist for more detailed assessment.

What are the main results of the review?

Mini-Cog for the diagnosis of Alzheimer's disease dementia and other dementias within a secondary care setting (Review)
The review included data from three relevant studies with a total of 2560 participants. However, the study authors did not use data from many of those participants they assessed, leaving results from only 1415 participants that provide complete and useful information for addressing the review question.

All three studies scored the Mini-Cog results in the way that was recommended by the developers of the tool. There was no clear pattern in the results of what a positive result of a Mini-Cog test meant across the three studies, making it difficult to draw summary conclusions. Using the studies with the highest and lowest Mini-Cog results indicated that if the Mini-Cog were to be used in secondary care in a group of 1000 people, where 640 (64%) have dementia, an estimated 510 to 557 would have a positive Mini-Cog, of which 0 to 126 would be incorrectly classified as having dementia. Of the 443 to 490 people with a result indicating dementia is not present, 83 to 256 would be incorrectly classified as not having dementia.

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

In the included studies, the diagnosis of dementia was made by assessing all patients with a detailed clinical assessment. Detailed clinical assessment is the reference standard to which the Mini-Cog was compared. This is likely to have been a reliable method for determining whether patients actually had dementia. However, there were some problems with how the studies were conducted in terms of the people who were included and how the Mini-Cog was calculated, which could result in the Mini-Cog appearing more accurate than it actually is. We decided that it was not appropriate to group the studies together to describe the average performance of the Mini-Cog, due to the differences among them.

To whom do the results of this review apply?

The studies included in the review were conducted in the USA, Germany, and Brazil. Two studies included those patients referred to specialists evaluating memory and thinking skills, and one study recruited individuals attending a medical outpatient clinic. The percentage of people with a final diagnosis of dementia was between 32% and 87% (an average of 64%).

What are the implications of this review?

The small number of studies identified and variation in how they used the Mini-Cog limit the evidence to make recommendations, and suggest that Mini-Cog may not be the best test to recommend for use in inpatient and outpatient secondary care hospital settings.

How up-to-date is this review?

The review authors searched for and considered studies published up to March 2019.
### Summary of Findings for the Main Comparison

What is the accuracy of the Mini-Cog for the diagnosis of Alzheimer’s disease dementia and other dementias within a secondary care setting?

<table>
<thead>
<tr>
<th>Population</th>
<th>Adult patients who completed the Mini-Cog, with no restrictions on the case mix of recruited participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Setting</td>
<td>Secondary care settings including outpatient clinics, inpatient general, and specialist hospital settings</td>
</tr>
<tr>
<td>Index test</td>
<td>Mini-Cog includes the Clock Draw Tests and 3-word recall</td>
</tr>
<tr>
<td>Reference standard</td>
<td>Clinical diagnosis of dementia using recognised diagnostic criteria</td>
</tr>
<tr>
<td>Studies</td>
<td>Cross-sectional but not case-control studies were included</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Study</th>
<th>Summary accuracy (95% CI)</th>
<th>No. of included participants</th>
<th>Dementia prevalence</th>
<th>Implications, quality, and comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clionsky 2010</td>
<td>Sensitivity: 0.67 (0.63 to 0.71) Specificity: 0.87 (0.81 to 0.92) Positive PV: 0.94 Negative PV: 0.49</td>
<td>702</td>
<td>73.5%</td>
<td>Only 40% of available records collated from neuropsychology and geriatric psychiatry practice were used to evaluate Mini-Cog accuracy. Sampling frame, inclusion and exclusion criteria were not described. Mini-Cog derived from longer cognitive test battery. Included all-cause dementia and reported dementia subtypes diagnosed. Patient selection - high risk of bias; unclear applicability concerns. Index test - unclear risk of bias; high applicability concerns. Reference standard - low risk of bias; low applicability concerns. Flow and timing - high risk of bias.</td>
</tr>
<tr>
<td>Filho 2009</td>
<td>Sensitivity: 0.60 (0.48 to 0.72) Specificity: 0.65 (0.57 to 0.73) Positive PV: 0.45 Negative PV: 0.78</td>
<td>211</td>
<td>32.2%</td>
<td>Multiple exclusion criteria including relying on participants providing informed consent. Only 69% of assessed participants included in the analysis. Focused on those who had 4 or fewer years of education. Mini-Cog derived from longer cognitive test battery.</td>
</tr>
<tr>
<td>Study</td>
<td>Scoring Method</td>
<td>Sensitivity</td>
<td>Specificity</td>
<td>Positive PV</td>
</tr>
<tr>
<td>-------</td>
<td>----------------</td>
<td>-------------</td>
<td>-------------</td>
<td>-------------</td>
</tr>
<tr>
<td>Milian 2012</td>
<td>as per Borson 2000</td>
<td>0.87 (0.83 to 0.90)</td>
<td>1.00 (0.94 to 1.00)</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Sensitivity: 0.87 (0.83 to 0.90)  
Specificity: 1.00 (0.94 to 1.00)  
Positive PV: 1.00  
Negative PV: 0.52

Included individuals admitted to Specialist Memory Clinic setting, but excluded those with depression and mild cognitive impairment which accounts for very high prevalence of dementia. Mini-Cog derived from longer cognitive test battery. Included all-cause dementia and reported dementia subtypes diagnosed.

Patient selection - high risk of bias; high applicability concerns  
Index test - unclear risk of bias; high applicability concerns  
Reference standard - low risk of bias; low applicability concerns  
Flow and timing - high risk of bias

95% CI: 95% confidence interval  
Positive PV: positive predictive value  
Negative PV: negative predictive value
BACKGROUND

Target condition being diagnosed

Alzheimer’s disease (AD) and related forms of dementia are common among older adults, with a prevalence of 8% in individuals over age 65 years, increasing to a prevalence of approximately 43% in adults aged 85 years and older (Thies 2012). Given the increasing number of older adults in most low and middle income countries, the prevalence of dementia is expected to increase considerably in the coming years (Prince 2016). Currently, an estimated 5.4 million Americans are diagnosed with AD, and this number is expected to increase to 6.7 million by 2025 (Thies 2012).

Alzheimer’s disease and related forms of dementia are currently incurable and result in considerable direct and indirect costs, both in terms of formal health care and lost productivity from both the affected individual and their caregivers (Thies 2012). There are several potential benefits to diagnosing AD and other dementias early in the disease course. Earlier diagnosis of AD allows for individuals with AD to make decisions regarding future planning whilst they retain the capacity to do so (Prorok 2013). A diagnosis of dementia is also necessary to access certain services and supports for individuals and their caregivers, and pharmacological treatments such as cholinesterase inhibitors or memantine may provide temporary symptomatic improvement in cognitive and functional symptoms for individuals with mild to moderate AD (Birks 2015; Birks 2018; McShane 2019; Rolinski 2012). There has been a relative lack of research exploring whether proactive case-finding for dementia is cost-effective and the impact on those referred for specialist assessment (Robinson 2015).

The diagnosis of AD is clinical and based on a history of decline in cognition with deficits in memory and at least one other area of cognitive functioning (e.g. apraxia, agnosia, or executive dysfunction). There must be a decline from a previous level of functioning that results in significant social or occupational impairment (American Psychiatric Association 2000; McKhann 2001). A definitive diagnosis of AD can only be achieved at autopsy, but a clinical diagnosis using standardised criteria is associated with a sensitivity of 81% and a specificity of 70% when compared to autopsy-proven cases (Knopman 2001).

Approximately 50% to 80% of all individuals with dementia are ultimately classified as AD (Blennow 2006; Brunnstrom 2009; Canadian Study of Health and Aging 1994). Whilst individuals with dementias share common characteristics, subtle differences can help to provide a diagnosis in the absence of neuropathological examination. Vascular dementias may occur more abruptly or present with a step-wise decline in cognitive functions over time, and account for approximately 15% to 20% of dementias (Canadian Study of Health and Aging 1994; Feldman 2003; Lobo 2000). Dementia with a mixed vascular and Alzheimer’s disease pathology is present in 10% to 30% of cases (Brunnstrom 2009; Crystal 2000; Feldman 2003). A smaller proportion of dementias are associated with dementia with Lewy bodies, Brunnstrom 2009, or Parkinson’s disease dementia (Aarsland 2005). Individuals experiencing frontotemporal dementia account for a smaller proportion of dementias (4% to 8%) and often present with problems in executive function and changes in behaviour, whilst memory is relatively preserved early in the disease course (Brunnstrom 2009; Greicius 2002).

Index test(s)

The Mini-Cog is a brief cognitive test consisting of two components: a delayed three-word recall and the clock-drawing test (Borson 2000). The Mini-Cog was initially developed in a community setting to provide a relatively brief cognitive screening test that was free of educational and cultural biases (Borson 2000). Different scoring algorithms were tested to determine which combination had the optimal balance of sensitivity and specificity (McCarten 2011; Scanlan 2001). The Mini-Cog takes approximately three to five minutes to complete in routine practice (Borson 2000; Holsinger 2007; Scanlan 2001). The Mini-Cog has been reported to have little potential for bias with education or language (Borson 2000; Borson 2005).

Clinical pathway

Dementia typically begins with subtle cognitive changes and progresses gradually over the course of several years. There is a presumed period when people are asymptomatic although the disease pathology may be progressing (Ritchie 2015). Individuals or their relatives may first notice subtle impairments of short-term memory or other areas of cognitive functioning. Gradually, the severity of cognitive deficits becomes apparent resulting in difficulty completing complex activities of daily living such as management of finances and medications, or operating motor vehicles (Ngjovanan 2001; Péres 2008). The attribution of cognitive symptoms to normal aging may cause delays in the diagnosis and treatment of AD or other types of dementia (Koch 2010). Consequently, there is a need for accurate brief cognitive screening tests to help distinguish between the cognitive changes associated with normal aging and changes that might indicate a dementia.

Cognitive assessment in a secondary care setting occurs in two broad contexts. Either individuals are referred from primary care or community health services for further evaluation of possible memory complaints. Alternatively, individuals may have their cognition evaluated as part of a comprehensive assessment of care needs following an acute, unscheduled hospitalisation. As a consequence, those assessed in secondary care settings would likely have some cognitive complaints or symptoms at the time of evaluation. Previous estimates suggest that 16% of those attending memory clinic had no cognitive impairments (Pusswald 2013), and 62% of hospitalised older adults had no cognitive impairment (Reynish 2000; Robinson 2015).
Secondary care settings that may use the Mini-Cog or other screening tests would include neurology, geriatric medicine, geriatric psychiatry services, or memory clinics. Typically, individuals who are assessed in secondary care settings would receive more detailed neuropsychological testing along with other investigations that are needed in order to confirm a diagnosis of dementia.

**Prior test(s)**

Mini-Cog is recommended for use as an initial screening test for dementia, therefore it is unlikely that individuals will have any formal testing completed prior to the administration of the Mini-Cog. The extent of any prior cognitive assessment may vary depending on source of referral (Fisher 2007).

**Role of index test(s)**

Most older adults with memory complaints will first present to their general practitioner or other primary healthcare provider (e.g. nurses or a nurse practitioner). Primary healthcare providers may then refer an individual to a secondary care setting such as a neurologist, geriatrician, or geriatric psychiatrist. Some countries have also recommended that brief cognitive screening tests be administered to all older adults in order to help screen for undetected or asymptomatic cognitive impairment (Cordell 2013), although routine screening of older adults for dementia is controversial (Martin 2015). We would anticipate that the Mini-Cog would be utilised as a screening test to guide further evaluation of cognitive complaints for individuals in secondary care and not as a diagnostic test in most instances.

**Alternative test(s)**

The Cochrane Dementia and Cognitive Improvement Group (CDCIG) have conducted a series of diagnostic test accuracy reviews of cognitive tests, biomarkers, and informant tools, as a planned programme of reviews (Davis 2013). This review focused only on the test accuracy of Mini-Cog, and alternative tests are not included as they have been examined in separate reviews.

**Rationale**

Cognitive diagnostic tests are required to assess cognition and assist in diagnosing conditions such as mild cognitive impairment and dementia. Comprehensive evaluation, conducted by psychologists or dementia specialists such as general psychiatrists, geriatric psychiatrists, geriatricians, or neurologists, using standardised diagnostic criteria, would be the reference standard for assessing cognition and diagnosing dementia in older adults. However, access to these specialised resources is scarce and expensive, and as such they are not practical for routine use in the evaluation of cognitive complaints (Pimlott 2009; Yaffe 2008). Whilst there are some cognitive tests that can be performed by healthcare providers who are not dementia specialists, many of these tests are time consuming and may not be practical to use as a first-line cognitive screen in secondary care settings. As such, brief but relatively accurate cognitive screening tests are required for healthcare providers in secondary care settings to identify individuals who may require more in-depth evaluation of cognition. In secondary care settings, brief cognitive screening tests may be used to guide subsequent evaluations or to complement more detailed evaluations.

Utilising a standard cognitive screening test like the Mini-Cog also promotes effective communication between healthcare providers. Sensitivity and specificity of such tests vary depending upon the setting in which they are used (Holsinger 2007). Some studies have found that in primary care the majority of older adults with dementia are undiagnosed (Boustani 2005; Sternberg 2000). In addition, many primary care providers have difficulty in accurately diagnosing dementia, and mild dementia in particular is underdiagnosed (van den Dungen 2011). Early diagnosis and treatment of dementia can have clinical benefits for the patient, their community, and the healthcare system (Bennett 2003; Prorok 2013; Thies 2012). Accurate diagnosis of dementia is also important in order to initiate dementia therapeutics including both non-pharmacological treatments and pharmacological treatments such as cholinesterase inhibitors, Birks 2015; Birks 2018; Rolinski 2012, or memantine (McShane 2019). A brief and simple cognitive screening test such as the Mini-Cog that could be used in secondary care settings would allow healthcare professionals or lay people to initially screen older adults for the presence of dementia. Individuals that screen positive for cognitive impairment on the Mini-Cog may then be further investigated for the presence of dementia using additional cognitive tests or other investigations. As Mini-Cog is brief, widely available, easy to administer, and has been reported to have reasonable test accuracy properties (Brodaty 2006; Lin 2013), it may be well suited for use as an initial cognitive screening test in secondary care. The Mini-Cog has been recommended as a suitable cognitive screening test for primary care in some countries (Cordell 2013). The current review examined the diagnostic accuracy of the Mini-Cog in secondary care settings. Separate DTA reviews have been undertaken evaluating the use of Mini-Cog in community and primary care settings (Fage 2015; Seitz 2018).

**OBJECTIVES**

To determine the diagnostic accuracy of the Mini-Cog for detecting Alzheimer’s disease dementia and other dementias in a secondary care setting.

**Secondary objectives**

To investigate the heterogeneity of test accuracy in the included studies and potential sources of heterogeneity. These potential
sources of heterogeneity will include the baseline prevalence of dementia in study samples, thresholds used to determine positive test results, the type of dementia (Alzheimer’s disease dementia or all causes of dementia), and aspects of study design related to study quality.

We identified gaps in the evidence where further research is required.

METHODOLOGY

Criteria for considering studies for this review

Types of studies

We included all cross-sectional studies with well-defined populations that utilised the Mini-Cog as an index cognitive screening test compared to a reference standard to identify dementia. Included studies also utilised the Mini-Cog as a screening test and not for confirmation of diagnosis. We excluded case-control studies and longitudinal designs in which there was a gap of more than four weeks between administration of the index test and reference standard.

Participants

Study participants were sampled from a secondary care setting and may or may not have ultimately been diagnosed with AD or other dementias. Participants may have had cognitive complaints or dementia at baseline, although their cognitive status should not be known to the individual administering the Mini-Cog or the reference standard. We excluded studies on participants with a developmental disability which prevented them from completing the Mini-Cog. Secondary care settings included inpatient and outpatient hospital participants. We excluded studies including participants in either a community or primary care setting, as these are topics of other reviews (Fage 2015; Seitz 2018).

Index tests

Mini-Cog test

The Mini-Cog consists of a three-word recall task and the clock-drawing test. The standard scoring system involves assigning a score of zero to three points on the word recall test for the correct recall of 0, 1, 2, or 3 words, respectively. The clock-drawing test is scored as being either ‘normal’ or ‘abnormal’. A positive test on the Mini-Cog (that is dementia) is assigned if either the delayed word recall score is zero out of three, or if the delayed recall score is either one or two, and the clock-drawing test is abnormal. A score of three on the delayed word recall or one to two on the delayed word recall with a normal clock drawing is a negative test (i.e. is no dementia) (Borson 2000).

Studies had to include the results of the Mini-Cog. Where multiple scoring algorithms were utilised, we explored the differences in results in subgroup analysis. We anticipated that the Mini-Cog would be utilised as a screening test to guide further evaluation of cognitive complaints for individuals in secondary care and not as a diagnostic test in most instances.

Target conditions

Target conditions included any stage of AD or other types of dementia including vascular dementia, dementia with Lewy bodies, Parkinson’s disease dementia, or frontotemporal dementia.

Reference standards

Whilst a definitive diagnosis can only be made postmortem at autopsy, there are clinical criteria for the diagnosis of most forms of dementia. All dementia diagnostic criteria require that an individual has deficits in multiple areas of cognition that results in impairment in daily functioning and is not caused by either the effects of a substance or a general medical condition. We describe potential reference standards for the diagnosis of all-cause dementia or specific types of dementia. All-cause dementia is commonly diagnosed using criteria such as the Diagnostic and Statistical Manual of Mental Disorders IV (DSM-IV) criteria for dementia (American Psychiatric Association 2000), the DSM-5 criteria for major neurocognitive disorder (American Psychiatric Association 2013), or the International Classification of Diseases (ICD) diagnosis of dementia (World Health Organization 2016). The standard clinical diagnostic criteria commonly used for Alzheimer’s disease dementia include the National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer’s Disease and Related Disorders Association (NINCDS-ADRA) criteria for probable or possible dementia (McKhan 1984; McKhan 2011). Diagnostic criteria for other types of dementia include the National Institute of Neurological Disorders and Stroke and Association Internationale pour la Recherche et l’Enseignement en Neurosciences (NINDS-AIREN) criteria for vascular dementia (Roman 1993), standard criteria for dementia with Lewy bodies (McKeith 2005), and for frontotemporal dementia (McKhan 2001). Evaluation often includes laboratory investigations, many of which are useful for excluding alternative diagnoses (Feldman 2008).

Additional procedures to help confirm the diagnosis include specific findings on neuroimaging (either computed tomography (CT) or magnetic resonance imaging (MRI)). These investigations are typically used to confirm the diagnosis rather than rule out the possibility of dementia. Whilst these clinical criteria for dementia are considered the reference standard for the purposes of our review, the sensitivity and specificity of these clinical reference
standards may vary when compared to neuropathological criteria for dementia (Nagy 1998).

**Search methods for identification of studies**

**Electronic searches**

We searched the Cochrane Dementia Group Register of Diagnostic Test Accuracy Studies, MEDLINE (OvidSP), Embase (OvidSP), BIOSIS Citation Index (ISI Web of Science), Web of Science Core Collection Science Citation Index (ISI Web of Science), PsycINFO (OvidSP), and LILACS (BIREME) (Latin American and Caribbean Health Science Information database). See Appendix 1 for the search strategy used and to view the ‘generic’ search that is run regularly for the Cochrane Dementia and Cognitive Improvement Group Register of Diagnostic Test Accuracy Studies. Controlled vocabulary such as MeSH terms and Emtree were used where appropriate. In order to maximise sensitivity and allow inclusion on the basis of population-based sampling to be assessed at testing (see Selection of studies), there was no attempt to restrict studies based on sampling frame or setting in the searches that were developed. We did not use search filters (collections of terms aimed at reducing the number needed to screen) as an overall limiter because those published have not proved sensitive enough (Whiting 2011). We applied no language restriction to the electronic searches, and used translation services as necessary. A single review author with extensive experience in systematic reviews performed the initial searches (ANS). Two review authors independently screened abstracts and titles.

**Searching other resources**

We searched the reference lists of all relevant studies for additional relevant studies. These studies were also used to search the electronic databases to identify additional studies through the use of the related article feature. We asked research groups authoring studies used in the analysis for any unpublished data.

**Data collection and analysis**

**Selection of studies**

Included studies had to:

1. make use of the Mini-Cog as a cognitive diagnostic tool;
2. include patients from a secondary care setting who may or may not have dementia or cognitive complaints. We did not include case-control studies;
3. clearly explain how a diagnosis of dementia was confirmed according to a reference standard such as the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition Text Revision (DSM IV-TR) or NINCDS-ADRDA at the same time or within the same four-week period that the Mini-Cog was administered. Formal neuropsychological evaluation was not required for a diagnosis of dementia. We first selected articles on the basis of abstracts and titles. We retrieved the full texts of those articles deemed potentially eligible, and two review authors independently assessed these for inclusion in the review. Any disagreements were settled by consulting a third review author.

**Data extraction and management**

Two review authors extracted the following data from all included studies.

- Author, journal, and year of publication.
- Scoring algorithm for the Mini-Cog including cut-points used to define a positive screen.
- Method of Mini-Cog administration including who administered and interpreted the test, their training, and whether or not the readers of the Mini-Cog and reference standard were blind (masked) to the results of the other test.
- Reference criteria and method used to confirm diagnosis of AD or other dementias.
- Baseline demographic characteristics of the study population including age, gender, ethnicity, spectrum of presenting symptoms, comorbidity, educational achievement, language, baseline prevalence of dementia, country, apolipoprotein E (ApoE) status, methods of participant recruitment, and sampling procedures.
- Length of time between administration of index test (Mini-Cog) and reference standard.
- The true positives, true negatives, false positives, false negatives, disease prevalence, sensitivity and specificity, and positive and negative likelihood ratios of the index test in defining dementia.
- Version of translation (if applicable).
- Prevalence of dementia in the study population.
- Estimates of test reproducibility (if available).

**Assessment of methodological quality**

We used the Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) criteria to assess data quality (Whiting 2011). The QUADAS-2 criteria contain assessment domains for patient selection, index test, reference test, and flow and timing. Each domain has suggested signalling questions to assist with the ‘Risk of bias’ assessment for each domain (Appendix 2). The potential risk of bias associated with each domain is rated as being at high, low, or uncertain risk of bias. In addition, we performed an assessment of the applicability of the study to the review question for each domain using the guide provided in the QUADAS-2. We used a standardised ‘Risk of bias’ template to extract data on the risk
of bias for each study using the form provided by the UK Support Unit Cochrane Diagnostic Test Accuracy Group (Appendix 3). We summarised the quality assessment results graphically and presented a narrative summary in the Results.

Statistical analysis and data synthesis
We performed statistical analysis as per the Cochrane guidelines for diagnostic test accuracy reviews (Macaskill 2010). Two-by-two tables were constructed separately for the Mini-Cog results for Alzheimer’s disease dementia and all-cause dementia where this information was available.
We entered data from the individual studies into Review Manager 5 (Review Manager 2014). We used reported data on test accuracy and disease prevalence or the true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN), whichever was reported by the individual study. We calculated the sensitivity, specificity, and positive and negative likelihood ratios as well as measures of statistical uncertainty (e.g. 95% confidence intervals). We presented data from each study graphically by plotting sensitivities and specificities on a coupled forest plot.
We did not conduct our prespecified statistical analysis, using bivariate random-effects for meta-analysis due to concerns related to methodological and clinical heterogeneity across the included studies.

Investigations of heterogeneity
The potential sources of heterogeneity included baseline prevalence of cognitive impairment in the target population, the cut-points used to determine a positive test result, the reference standard used to diagnose dementia, the type of dementia (Alzheimer’s disease dementia or all-cause dementia), the severity of dementia in the study sample, and aspects related to study quality. We have presented narrative results where data were available to describe the between-study heterogeneity.

Sensitivity analyses
We planned to conduct sensitivity analyses to investigate study quality and the impact on overall diagnostic accuracy of Mini-Cog. We did not perform this analysis given the methodological concerns with all three of the identified studies. We also did not evaluate the impact of individual studies on summary outcome measures, as it was not considered appropriate to calculate a summary measure from the data in the three studies due to their heterogeneity.

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

RESULTS
Results of the search
The results of the literature search are summarised in Figure 1. An initial review of the electronic databases in September 2012 identified 108 articles. We updated this search in January 2013, adding an additional 106 articles, and completed a second update in February 2015 that identified another 34 potentially relevant citations. The same search strategy was employed for this review that was used in separate reviews of the Mini-Cog in the community setting and the primary care setting (Fage 2015; Seitz 2018). We performed a final update search for this review only in March 2019, which identified 324 articles. After removal of duplicates, two review authors independently reviewed a total of 468 abstracts and citations to determine those that were potentially eligible. We reviewed a total of 62 full-text articles for eligibility, of which 59 were excluded. Reasons for exclusions were: a lack of a reference standard (N = 20), incorrect setting (N = 14), duplicate publications (N = 12), failure to include the Mini-Cog as an index test (N = 5), wrong study design (N = 4), not using Mini-Cog to diagnose dementia (N = 3), and lack of sufficient data to be included in the review (N = 1). We included three studies in the review.
Figure 1. Study flow diagram.

N = 1688 records identified through database searching (4 searches: September 2012, July 2013, February 2015, March 2019) -> N = 2 additional records identified through other sources

N = 456 records after duplicates removed and first assessment based on title and abstract screening by Cochrane Information Specialist

N = 408 records screened -> N = 406 records excluded

N = 59 full-text articles excluded, with reasons:
- Lack of reference standard = 20
- Setting other than secondary care = 14
- Duplicate publication/data = 12
- Did not use Mini-Cog as index test = 5
- Wrong study design = 4
- Not diagnosing dementia = 3
- Insufficient information available = 1

N = 62 full-text articles assessed for eligibility

N = 3 studies included in qualitative synthesis

Mini-Cog for the diagnosis of Alzheimer's disease dementia and other dementias within a secondary care setting (Review)
Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.
Methodological quality of included studies

The results of the QUADAS-2 assessment for the three included studies are presented graphically in Figure 2.

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

We considered all studies at high risk of bias with respect to patient selection. This was as a result of non-consecutive samples or exclusion of records available for inclusion. There were high applicability concerns with respect to patient selection from two of the studies (Filho 2009; Milian 2012). These were as a result of excluding those patients with common and important conditions, such as depression or mild cognitive impairment, the exclusion of those with sensory impairments, and the need for some participants to provide informed consent to participate. The inclusion/exclusion criteria for Clionsky 2010 were not reported in the paper, rendering assessment of applicability as at unclear risk.

An additional feature common to all the included studies was the way the index test was performed and evaluated. All studies retrospectively derived the Mini-Cog using responses to the three-word recall and clock-drawing test that were collected as part of a larger and longer battery of neuropsychological tests. The accuracy of the Mini-Cog may have been affected when the result of the Mini-Cog stemmed from more comprehensive testing compared to when the component tests were administered by themselves. As such we rated all studies as at unclear risk of bias for the index test domain and as at high concern regarding the applicability of the results, given that this is not how short cognitive index tests would typically be performed.

In all cases we considered that the reference standard used was...
likely to correctly identify the target condition, and the risk of bias associated with this was low and applicability concerns were low. However, there are risks of incorporation bias if the assessments used to derive the Mini-Cog were known to those making the reference standard diagnosis. We assessed two studies as high risk in the flow and timing domain due to use of multiple reference standards, Clionsky 2010, and selective use of data (Clionsky 2010; Filho 2009). We considered one study as at low risk of bias for the flow and timing domain (Milian 2012).

Findings

We included three papers in the final review (Clionsky 2010; Filho 2009; Milian 2012). The study population included in the review was 1415 participants. However, this represents only 55.2% of the total population (n = 2560) reported in the included studies. Data from 1145 participants were not available for use (n = 1050 excluded from Clionsky 2010, as only one of their five data sets were included in the analysis for Mini-Cog test accuracy; and n = 95 excluded from Filho 2009 due to incomplete data or having more than four years of formal education, as the study only included those considered to have low levels of education). Clionsky 2010 reported the development and validation of a new cognitive test; this process was based on using historic clinical data sets, and only one of these data sets reported Mini-Cog test accuracy data, thus the majority of participants in the study do not contribute data to this review. Clionsky 2010 included participants referred to neuropsychology services for assessment; Filho 2009 recruited a selection of participants attending for general medical treatment on an outpatient basis; and Milian 2012 included individuals referred to a memory clinic. None of the studies evaluated Mini-Cog among general hospital inpatients.

Key features of the studies are summarised in the Characteristics of included studies table. Additional features of these studies are also reported in Table 1. One study did not report dementia subtype, using a binary classification of dementia versus no dementia (Filho 2009). The other studies reported dementia subtype in more detail, however there was some overlap in the categorisation of subtypes, making direct comparison more challenging (Table 1). Alzheimer’s disease was the most common subtype in 57.3%, Clionsky 2010, and 49.1%, Milian 2012, of participants. The prevalence of dementia in the study samples varied from 32.2% (95% confidence interval (CI) 26.3 to 38.8%), Filho 2009, to 87.3% (95% CI 84.1 to 89.9%), Milian 2012, although important exclusions apply to the study populations that impact these figures (e.g. exclusion of those with mild cognitive impairment, depression or any other mental health diagnoses or the need for participants to provide informed consent). The descriptive variables of ethnicity, comorbidity, spectrum of presenting symptoms, and ApoE status were not reported in any of the included studies.

All studies utilised the original scoring system (Borson 2000). Filho 2009 reported multiple thresholds to define Mini-Cog test positivity and the impact of varying the threshold on sensitivity and specificity. Data from the original scoring system are those used to provide comparability between results. Meta-analysis was planned to determine a summary pooled estimate for the diagnostic test accuracy of Mini-Cog in identifying dementia in secondary care settings. However, due to the methodological limitations and heterogeneity of included studies, we did not perform quantitative synthesis. The extracted data, including sensitivity, specificity, and likelihood ratios for each study, are summarised in Summary of findings. A forest plot is presented in Figure 3. The sensitivities of the Mini-Cog in the individual studies were reported as 0.67 (95% CI 0.63 to 0.71) (Clionsky 2010), 0.60 (95% CI 0.48 to 0.72) (Filho 2009), and 0.87 (95% CI 0.83 to 0.90) (Milian 2012). The specificity of the Mini-Cog for each individual study was 0.87 (95% CI 0.81 to 0.92) (Clionsky 2010), 0.65 (95% CI 0.57 to 0.73) (Filho 2009), and 1.00 (95% CI 0.94 to 1.00) (Milian 2012). Positive predictive values were 0.94 (Clionsky 2010), 0.45 (Filho 2009), and 1.00 (Milian 2012). Negative predictive values were 0.49 (Clionsky 2010), 0.78 (Filho 2009), and 0.52 (Milian 2012).

Figure 3. Forest plot of individual study results using Mini-Cog in secondary care for the diagnosis of dementia.

<table>
<thead>
<tr>
<th>Study</th>
<th>TP</th>
<th>FP</th>
<th>FN</th>
<th>TN</th>
<th>Sensitivity (95% CI)</th>
<th>Specificity (95% CI)</th>
<th>Sensitivity (95% CI)</th>
<th>Specificity (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clionsky 2010</td>
<td>346</td>
<td>24</td>
<td>170</td>
<td>192</td>
<td>0.67 [0.63, 0.71]</td>
<td>0.87 [0.81, 0.92]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Filho 2009</td>
<td>41</td>
<td>6</td>
<td>27</td>
<td>63</td>
<td>0.60 [0.48, 0.72]</td>
<td>0.65 [0.57, 0.73]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Milian 2012</td>
<td>380</td>
<td>0</td>
<td>56</td>
<td>64</td>
<td>0.87 [0.83, 0.90]</td>
<td>1.00 [1.00, 1.00]</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
One study reported test accuracy by dementia subtype, classifying these as: all dementia, Alzheimer's dementia, and non-Alzheimer's dementia (Milian 2012). Sensitivity was reported to be higher for Alzheimer's dementia (0.91 versus 0.87) and lower for non-Alzheimer's dementia (0.83 versus 0.87) compared to all dementias, although the authors do not report formal statistical comparisons for these results (Milian 2012).

We did not perform planned subgroup analyses to investigate potential sources of heterogeneity and sensitivity analyses due to the methodological and clinical heterogeneity. Study characteristics regarding prevalence of dementia, cut-points to determine a positive test, the reference standard used to diagnose dementia, and the type of dementia are reported in Table 1. Severity of dementia was not reported in any of the included studies. Study quality is described above and in the Characteristics of included studies table.

**Discussion**

**Summary of main results**

We found only three studies that evaluated the test accuracy of the Mini-Cog in secondary care settings compared to a reference standard assessment using recognised dementia diagnostic criteria. Only 55.3% of available patient data was used to evaluate the test accuracy of Mini-Cog. Our 'Risk of bias' assessment identified concerns regarding both the internal and external validity of the included studies. In all studies the Mini-Cog was retrospectively derived from historic data sets. No studies included acute general hospital inpatients. The prevalence of dementia ranged from 32.2% to 87.3%. The sensitivities of the Mini-Cog in the individual studies were reported as 0.67 (95% CI 0.63 to 0.71), 0.60 (95% CI 0.48 to 0.72), and 0.87 (95% CI 0.83 to 0.90). The specificity of the Mini-Cog for each individual study was 0.87 (95% CI 0.81 to 0.92), 0.65 (95% CI 0.57 to 0.73), and 1.00 (95% CI 0.94 to 1.00). Positive predictive values were 0.94, 0.45, and 1.00. Negative predictive values were 0.49, 0.78, and 0.52. We did not perform meta-analysis due to the concerns regarding risk of bias and heterogeneity.

**Strengths and weaknesses of the review**

We conducted the review following the published protocol (Chan 2014), with only limited differences (Differences between protocol and review). The search strategy was robust and conducted by an Information Specialist, using an approach common across the Mini-Cog reviews. Study quality was formally evaluated using the QUADAS-2 methodology (Appendix 2) and utilising anchoring statements common across the Cochrane Dementia and Cognitive Improvement Group suite of diagnostic test accuracy reviews (Appendix 3).

Limitations of the review primarily reflect the lack of eligible studies for inclusion and heterogeneity of populations and methods. These precluded the conduct of the prespecified analyses, including evaluating dementia subtype and producing a summary of test estimate Mini-Cog test accuracy in secondary care and more detailed evaluation of the effect of heterogeneity on study findings. We recognise that had the identified studies been more methodologically and clinically consistent, specific methods have been developed to allow meta-analysis of diagnostic test accuracy studies when there are small numbers of studies identified for inclusion (Takwoingi 2017). Finally, in common with other diagnostic test accuracy reviews in dementia, the reliance on an imperfect reference standard is an important limitation to acknowledge.

**Applicability of findings to the review question**

This review sought to evaluate the test accuracy of the Mini-Cog for detecting dementia in secondary care. None of the included studies evaluated the Mini-Cog as a traditional test accuracy approach would anticipate: where the Mini-Cog would be conducted and this would be followed by an independent and contemporaneous reference standard assessment to diagnose dementia. All three studies retrospectively derived Mini-Cog from more lengthy neuropsychological tests, which had the potential to introduce bias. There is also potential for incorporation bias within the reference standard, as the components of these tests may have helped inform reference standard assessment. Finally, the three populations studied are non-consecutive with exclusions which affect the prevalence of dementia and applicability of results to current clinical practice. As such, evaluation of the included studies make it difficult to answer the question underpinning this review.

**Authors' conclusions**

**Implications for practice**

The available studies have not considered use of Mini-Cog as a stand-alone cognitive assessment compared to reference standard diagnosis of dementia. The methodological limitations and heterogeneity, coupled with the small number of included studies, means there is limited information to draw firm conclusions. The results are inconsistent with respect to the pattern of test accuracy. The currently available evidence thus does not support recommending Mini-Cog as a short cognitive test for use in secondary care settings. This finding is in-keeping with the evidence regarding use of Mini-Cog in primary care, Seitz 2018, and community settings, Fage 2015. A range of brief cognitive tests are available and used in clinical practice (Velayudhan 2014), although with a lack of empirical head-to-head data comparing their test accuracy. Other cognitive tests, such as the Montreal Cognitive Assessment, Davis 2015, or Mini-Mental State Examination, Creavin 2016.
appear to have higher levels of sensitivity 0.94 (no 95% confidence interval (CI) reported) and 0.85 (95% CI 0.74 to 0.92) respectively, which may be desirable in a screening test. However, the greater burden of testing associated with longer instruments needs to be taken into account.

Implications for research

Future research should specifically evaluate the Mini-Cog as an index test, rather than as part of a wider cognitive test battery, or when authors are attempting to create a new diagnostic instrument. Use of multiple different instruments and creation of new tools is common across the dementia and cognitive impairment literature, and potentially represents a barrier to research due to lack of standardisation and heterogeneity of assessment (Harrison 2016).

Furthermore, the included studies did not address any feasibility questions around use of Mini-Cog. Some studies specifically excluded those with sensory impairment or hand movement limitations. Mini-Cog is a short cognitive test and thus may show promise with respect to feasibility in practice. The ability of hospitalised adult populations to complete diagnostic assessments is a critical part of determining the applicability and usefulness of a testing strategy in clinical practice and one that has been commonly overlooked in the dementia diagnostic test accuracy literature in secondary care (Elliott 2019; Harrison 2015; Lees 2017). Authors conducting diagnostic test accuracy research in the dementia field should utilise reporting guidance to help improve transparency and allow a more complete critical evaluation of the methodology employed (Noel-Storr 2014).

Finally, the aim of evaluating the test accuracy of tools to identify dementia is to improve diagnostic pathways to achieve better outcomes for individuals using health and care services. Further research is needed to evaluate the role of diagnostic strategies in changing individual and population-level outcomes to ensure that use of instruments is appropriate.

Acknowledgements

We would like to thank peer reviewers Tom Dening and Andrew Larner and consumer reviewer Cathie Hofstetter for their comments and feedback.

References

References to studies included in this review

Clionsky 2010  {published data only}

Filho 2009  {published data only}

Milian 2012  {published data only}

References to studies excluded from this review

Borson 2000  {published data only}

Borson 2003  {published data only}

Borson 2005  {published data only}

Borson 2006  {published data only}

Carnero-Pardo 2013  {published data only}

Carnero-Pardo 2019  {published data only}

Chen 2011 *(published data only)*

Dash 2004 *(published data only)*

Del Ser 2000 *(published data only)*

Dougherty 2010 *(published data only)*

Fuchs 2012 *(published data only)*

Geschke 2019 *(published data only)*

Holsinger 2012 *(published data only)*

McCarten 2012 *(published data only)*

Milian 2013 *(published data only)*

Sonnett 2012 *(published data only)*

Steenland 2008 *(published data only)*

Wilber 2005 *(published data only)*

Wright 2011 *(published data only)*

Yang 2016 *(published data only)*

Yang 2018 *(published data only)*

Additional references

Aarsland 2005

American Psychiatric Association 2000

American Psychiatric Association 2013

Bennett 2003

Birks 2015

Birks 2018

Blennow 2006
patients with dementia in primary care: a systematic review. 
BMC Family Practice 2010;11:52.

Lees 2017

Lin 2013

Lobo 2000

Macaskill 2010

Martin 2015

McCarten 2011

McKeith 2005

McKhann 1984

McKhann 2001

McKhann 2011

McShane 2019

Nagy 1998

Njegovan 2001

Noel-Storr 2014

Pimlott 2009

Prince 2016

Prorok 2013

Pusswald 2013

Péret 2008
Mini-Cog for the diagnosis of Alzheimer's disease dementia and other dementias within a secondary care setting (Review)

Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.
## Characteristics of included studies  
*ordered by study ID*

### Clionsky 2010

<table>
<thead>
<tr>
<th>Study characteristics</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Patient sampling</strong></td>
<td>The study had access to 1752 patient records collated from 5 data sources from neuropsychology and geriatric psychiatry practice. Sampling frame for record collection is not described. 702 records collected between 2005 and 2008 were used to generate data evaluating Mini-Cog test accuracy.</td>
</tr>
<tr>
<td><strong>Patient characteristics and setting</strong></td>
<td>Records used in the analysis were collected in a neuropsychology setting where individuals were referred by a physician or agency in the community. No inclusion or exclusion criteria are listed.</td>
</tr>
<tr>
<td><strong>Index tests</strong></td>
<td>Mini-Cog was retrospectively derived from results of the Mini-Mental State Examination (which contains 3-item recall) and Clock Draw Test, so the index test was not conducted contemporaneously as expected in the review question. Mini-Cog was scored according to original criteria in Borson 2000.</td>
</tr>
<tr>
<td><strong>Target condition and reference standard(s)</strong></td>
<td>Clinical diagnosis of dementia was determined based on DSM-IV criteria by 1 of 6 licensed psychologists. Patients were evaluated based on their age- and education-adjusted neuropsychological test scores, medical and psychiatric history, and interview with a family informant.</td>
</tr>
<tr>
<td><strong>Flow and timing</strong></td>
<td>Study authors made use of 5 data sets to perform the retrospective analyses reported in the paper. Those in neuropsychology received a different reference standard assessment to those assessed in geriatric psychiatry. Reason for use of only 1 set of individuals assessed in neuropsychology for calculation of test accuracy results not provided in the paper.</td>
</tr>
</tbody>
</table>

### Methodological quality

<table>
<thead>
<tr>
<th>Item</th>
<th>Authors’ judgement</th>
<th>Risk of bias</th>
<th>Applicability concerns</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>DOMAIN 1: Patient Selection</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Was a consecutive or random sample of patients enrolled?</td>
<td>No</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Was a case-control design avoided?</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Did the study avoid inappropriate exclusions?</td>
<td>Unclear</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| | High | Unclear |

---

Mini-Cog for the diagnosis of Alzheimer’s disease dementia and other dementias within a secondary care setting (Review)  
Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.
**DOMAIN 2: Index Test All tests**

<table>
<thead>
<tr>
<th>Question</th>
<th>Clionsky 2010</th>
</tr>
</thead>
<tbody>
<tr>
<td>Were the index test results interpreted without knowledge of the results of the reference standard?</td>
<td>Unclear</td>
</tr>
<tr>
<td>If a threshold was used, was it pre-specified?</td>
<td>Yes</td>
</tr>
</tbody>
</table>

**DOMAIN 3: Reference Standard**

<table>
<thead>
<tr>
<th>Question</th>
<th>Clionsky 2010</th>
</tr>
</thead>
<tbody>
<tr>
<td>Is the reference standards likely to correctly classify the target condition?</td>
<td>Yes</td>
</tr>
<tr>
<td>Were the reference standard results interpreted without knowledge of the results of the index tests?</td>
<td>Yes</td>
</tr>
</tbody>
</table>

**DOMAIN 4: Flow and Timing**

<table>
<thead>
<tr>
<th>Question</th>
<th>Clionsky 2010</th>
</tr>
</thead>
<tbody>
<tr>
<td>Was there an appropriate interval between index test and reference standard?</td>
<td>Yes</td>
</tr>
<tr>
<td>Did all patients receive the same reference standard?</td>
<td>No</td>
</tr>
<tr>
<td>Were all patients included in the analysis?</td>
<td>No</td>
</tr>
</tbody>
</table>

**Filho 2009**

**Study characteristics**

| Patient sampling | Convenience sample of 306 recruited individuals, 65 years of age or older, seeking general medical treatment as outpatients at Internal Medicine Clinic of the Policlinica Piquet Carneiro at Rio de Janeiro State University Hospital. Sampling limited by the number of consenting individuals and availability of research team to assess individuals each day. Occasionally, patients returned next day to finish their testing. Final sample restricted only to those 211 individuals who had complete |
Patient characteristics and setting

Inclusion criteria included preserved hearing and comprehension to fully participate in the study and sign an informed consent form. Exclusions were reports of a serious uncorrected visual or auditory deficiency; being at an advanced stage of cognitive disturbance or having any mental illness that could compromise understanding of and performance on the test procedures; having a native language other than Portuguese; and difficulty in hand movement due to rheumatic or neurological diseases.

Index tests

Multiple thresholds for defining a positive result using Mini-Cog are reported in the paper. These include the methodology described by Borson 2000. Mini-Cog was not collected at the time of patient assessment but derived retrospectively, so the index test was not conducted contemporaneously as expected in the review question.

Target condition and reference standard(s)

The diagnosis of dementia was made based on the formal criteria of DSM-IV, as agreed upon between geriatrician and neuropsychologist. This included clinical impression and neuropsychological evaluation, including some components of the index tests.

Flow and timing

Neuropsychological assessments could be administered on different days. Patients who were lost in follow-up and those who did not finish their evaluation were excluded. Those with more than 4 years of education were excluded from the analysis despite being assessed.

Comparative

Notes

Methodological quality

<table>
<thead>
<tr>
<th>Item</th>
<th>Authors’ judgement</th>
<th>Risk of bias</th>
<th>Applicability concerns</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>DOMAIN 1: Patient Selection</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Was a consecutive or random sample of patients enrolled?</td>
<td>No</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Was a case-control design avoided?</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Did the study avoid inappropriate exclusions?</td>
<td>No</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>High</td>
<td></td>
</tr>
<tr>
<td><strong>DOMAIN 2: Index Test All tests</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Were the index test results interpreted without knowledge of the results of the reference standard?</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
If a threshold was used, was it pre-specified? | 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
---|---

**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
---|---

**Milian 2012**

**Study characteristics**

Patient sampling
---
Participants selected retrospectively from all admitted patients to the Memory Clinic of the Department of Psychiatry and Psychotherapy of the University Hospital of Tübingen between 2004 to 2009. Sampling frame is not described, but non-consecutive sample included based on exclusion criteria

Patient characteristics and setting
---
Sample composed of older adults with normal or corrected-to-normal visual acuity, “sufficient hearing ability”, and a wide education range. Those described as having “severe handicap affecting the ability to perform the required tasks”, mild cognitive impairment, or a depressive episode were excluded from analysis in the study. Also, patients with underlying neurological and psychiatric disorder unrelated to the diagnosis of dementia were excluded. Study was set in an inpatient memory clinic in Germany of individuals referred for evaluation of cognition
Study participants completed the clock-drawing test and Mini-Mental State Examination (which contains 3-item recall) as part of their assessment. This was done without knowledge of reference standard diagnosis. However, the full index test (Mini-Cog) was retrospectively derived, so the index test was not conducted contemporaneously as expected in the review question. Mini-Cog results were scored in accordance with original scoring guidance from tool authors (Borson 2000).

Clinical diagnosis of dementia based on DSM-IV criteria, the ICD-10 of Mental and Behavioural Disorders, and NINCDS-ADRDA. Index test results (from Mini-Cog) were not calculated at the time of reference standard diagnosis. Other neuropsychological tests results were available and informed reference standard diagnosis

Information to inform the index test and reference standard was collected contemporaneously as part of a diagnostic assessment. The items used to derive the index test were collected prior to completion of a reference standard diagnosis

<table>
<thead>
<tr>
<th>Index tests</th>
<th>Study participants completed the clock-drawing test and Mini-Mental State Examination (which contains 3-item recall) as part of their assessment. This was done without knowledge of reference standard diagnosis. However, the full index test (Mini-Cog) was retrospectively derived, so the index test was not conducted contemporaneously as expected in the review question. Mini-Cog results were scored in accordance with original scoring guidance from tool authors (Borson 2000).</th>
</tr>
</thead>
<tbody>
<tr>
<td>Target condition and reference standard(s)</td>
<td>Clinical diagnosis of dementia based on DSM-IV criteria, the ICD-10 of Mental and Behavioural Disorders, and NINCDS-ADRDA. Index test results (from Mini-Cog) were not calculated at the time of reference standard diagnosis. Other neuropsychological tests results were available and informed reference standard diagnosis</td>
</tr>
<tr>
<td>Flow and timing</td>
<td>Information to inform the index test and reference standard was collected contemporaneously as part of a diagnostic assessment. The items used to derive the index test were collected prior to completion of a reference standard diagnosis</td>
</tr>
</tbody>
</table>

**DOMAIN 1: Patient Selection**

<table>
<thead>
<tr>
<th>Item</th>
<th>Authors’ judgement</th>
<th>Risk of bias</th>
<th>Applicability concerns</th>
</tr>
</thead>
<tbody>
<tr>
<td>Was a consecutive or random sample of patients enrolled?</td>
<td>No</td>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td>Was a case-control design avoided?</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Did the study avoid inappropriate exclusions?</td>
<td>No</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**DOMAIN 2: Index Test All tests**

<table>
<thead>
<tr>
<th>Item</th>
<th>Authors’ judgement</th>
<th>Risk of bias</th>
<th>Applicability concerns</th>
</tr>
</thead>
<tbody>
<tr>
<td>Were the index test results interpreted without knowledge of the results of the reference standard?</td>
<td>Unclear</td>
<td></td>
<td></td>
</tr>
<tr>
<td>If a threshold was used, was it pre-specified?</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

**Notes**

**Methodological quality**

---
### DOMAIN 3: Reference Standard

<table>
<thead>
<tr>
<th>Question</th>
<th>Yes</th>
<th>Unclear</th>
<th>Low</th>
<th>Low</th>
</tr>
</thead>
<tbody>
<tr>
<td>Is the reference standards likely to correctly classify the target condition?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Were the reference standard results interpreted without knowledge of the results of the index tests?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### DOMAIN 4: Flow and Timing

<table>
<thead>
<tr>
<th>Question</th>
<th>Yes</th>
<th>Yes</th>
<th>Yes</th>
<th>Low</th>
</tr>
</thead>
<tbody>
<tr>
<td>Was there an appropriate interval between index test and reference standard?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Did all patients receive the same reference standard?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Were all patients included in the analysis?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

DSM-IV - Diagnostic and Statistical Manual of Mental Disorders, fourth edition; ICD-10 - International Classification of Diseases, Tenth Revision; NINCDS-ADRDA - National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer's Disease and Related Disorders Association

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

<table>
<thead>
<tr>
<th>Study</th>
<th>Reason for exclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Borson 2000</td>
<td>Wrong setting - not in secondary care</td>
</tr>
<tr>
<td>Borson 2003</td>
<td>Wrong setting - not in secondary care</td>
</tr>
<tr>
<td>Borson 2005</td>
<td>Wrong setting - not in secondary care</td>
</tr>
<tr>
<td>Borson 2006</td>
<td>Wrong setting - not in secondary care</td>
</tr>
<tr>
<td>Carnero-Pardo 2013</td>
<td>Wrong setting - not in secondary care</td>
</tr>
<tr>
<td>Study</td>
<td>Issue</td>
</tr>
<tr>
<td>------------------</td>
<td>----------------------------------------------------------------------</td>
</tr>
<tr>
<td>Carnero-Pardo 2019</td>
<td>Wrong outcome measure - not diagnosing dementia</td>
</tr>
<tr>
<td>Chen 2011</td>
<td>Participants did not receive gold standard evaluation using standardised diagnostic criteria</td>
</tr>
<tr>
<td>Dash 2004</td>
<td>Participants did not receive gold standard evaluation using standardised diagnostic criteria</td>
</tr>
<tr>
<td>Del Ser 2000</td>
<td>Study did not use Mini-Cog as index test.</td>
</tr>
<tr>
<td>Dougherty 2010</td>
<td>Wrong study design</td>
</tr>
<tr>
<td>Fuchs 2012</td>
<td>Wrong setting - not in secondary care</td>
</tr>
<tr>
<td>Geschke 2019</td>
<td>Participants did not receive gold standard evaluation using standardised diagnostic criteria</td>
</tr>
<tr>
<td>Holsinger 2012</td>
<td>Wrong setting - not in secondary care</td>
</tr>
<tr>
<td>McCarten 2012</td>
<td>Wrong setting - not in secondary care</td>
</tr>
<tr>
<td>Milian 2013</td>
<td>Duplicate publication/data</td>
</tr>
<tr>
<td>Sonnett 2012</td>
<td>Participants did not receive gold standard evaluation using standardised diagnostic criteria</td>
</tr>
<tr>
<td>Steenland 2008</td>
<td>Wrong study design</td>
</tr>
<tr>
<td>Wilber 2005</td>
<td>Participants did not receive gold standard evaluation using standardised diagnostic criteria</td>
</tr>
<tr>
<td>Wright 2011</td>
<td>Study did not use Mini-Cog as index test.</td>
</tr>
<tr>
<td>Yang 2016</td>
<td>Wrong setting - not in secondary care</td>
</tr>
<tr>
<td>Yang 2018</td>
<td>Wrong setting - not in secondary care</td>
</tr>
</tbody>
</table>
D A T A
Presented below are all the data for all of the tests entered into the review.

Tests. Data tables by test

<table>
<thead>
<tr>
<th>Test</th>
<th>No. of studies</th>
<th>No. of participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Mini-Cog in secondary care</td>
<td>3</td>
<td>1415</td>
</tr>
</tbody>
</table>

Test 1. Mini-Cog in secondary care.

Review: Mini-Cog for the diagnosis of Alzheimer’s disease dementia and other dementias within a secondary care setting

<table>
<thead>
<tr>
<th>Study</th>
<th>TP</th>
<th>FP</th>
<th>FN</th>
<th>TN</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clionsky 2010</td>
<td>346</td>
<td>24</td>
<td>170</td>
<td>162</td>
<td>0.67 [0.63, 0.71]</td>
<td>0.87 [0.81, 0.92]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Filho 2009</td>
<td>41</td>
<td>50</td>
<td>27</td>
<td>93</td>
<td>0.60 [0.48, 0.72]</td>
<td>0.65 [0.57, 0.73]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Milian 2012</td>
<td>380</td>
<td>0</td>
<td>58</td>
<td>64</td>
<td>0.87 [0.83, 0.90]</td>
<td>1.00 [0.94, 1.00]</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

A D D I T I O N A L T A B L E S

Table 1. Overview of study characteristics

<table>
<thead>
<tr>
<th>Study ID</th>
<th>Country</th>
<th>Study participants (N)</th>
<th>Mean age in years (SD)</th>
<th>Female gender %</th>
<th>Level of education</th>
<th>Mini-Cog scoring</th>
<th>Reference standard</th>
<th>Dementia prevalence N (%)</th>
<th>Dementia subtype N (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clionsky 2010</td>
<td>USA</td>
<td>1752 records available to authors; 702</td>
<td>78.2 (7.2)</td>
<td>61.0</td>
<td>12.8 years (+/-3.1)</td>
<td>As per Borson 2000</td>
<td>DSM-IV</td>
<td>516 (73.5; 95% CI)</td>
<td>Alzheimer’s disease 402 (57.6%)</td>
</tr>
</tbody>
</table>
Table 1. Overview of study characteristics (Continued)

<table>
<thead>
<tr>
<th>Study</th>
<th>Country</th>
<th>Participants</th>
<th>Cognitive impairment</th>
<th>Mini-Cog</th>
<th>DSM-IV Classification</th>
<th>Dementia Subtypes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Filho 2009</td>
<td>Brazil</td>
<td>306; 211 (69.0%) included in analysis</td>
<td>Dementia 74.0 (5.8) No dementia 72.0 (5.0)</td>
<td>72.5</td>
<td>Whole included sample ≤ 4 years</td>
<td>Multi-thresholds reported including scoring as per Borson 2000</td>
</tr>
<tr>
<td>Milian 2012</td>
<td>Germany</td>
<td>502; all included in analysis</td>
<td>Dementia 75.0 (8.5) No cognitive impairment 73.1 (5.5)</td>
<td>61.4</td>
<td>≤ 8 years: 47.7%; 9 to 11 years: 26.3%; ≥ 12 years: 26.0%</td>
<td>As per Borson 2000</td>
</tr>
</tbody>
</table>

(Continued)
Table 1. Overview of study characteristics (Continued)

<table>
<thead>
<tr>
<th>Source and platform</th>
<th>Search strategy</th>
<th>Hits retrieved</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cochrane Dementia and Cognitive Improvement Group DTA register (see bottom of this table for more information in the search narrative) [Date of most recent search 12 March 2019]</td>
<td>1. &quot;mini-Cog&quot; [all fields] 2. minicog [all fields] 3. (MCE and (cognit* OR dement* OR screen* OR Alzheimer*)) [all fields] 3. or/1-3</td>
<td>Sept 2012: 452 Jul 2013: 34 Feb 2015: 7 Mar 2019: 7</td>
</tr>
<tr>
<td>MEDLINE In-Process and Other Non-Indexed Citations and MEDLINE 1950 to present (OvidSP) [Date of most recent search 12 March 2019]</td>
<td>1. &quot;mini-Cog&quot;.ti,ab. 2. minicog.ti,ab. 3. (MCE and (cognit* OR dement* OR screen* OR Alzheimer*)).ti,ab 3. or/1-3</td>
<td>Sept 2012: 91 Jul 2013: 12 Feb 2015: 31 Mar 2019: 67</td>
</tr>
<tr>
<td>Embase 1974 to 11 March 2019 (OvidSP) [Date of most recent search 12 March 2019]</td>
<td>1. &quot;mini-cog*&quot;.mp. 2. minicog*.mp. 3. (MCE and (cognit* OR dement* OR screen* OR Alzheimer*)).ti,ab 3. or/1-3</td>
<td>Sept 2012: 96 Jul 2013: 37 Feb 2015: 80 Mar 2019: 205</td>
</tr>
</tbody>
</table>

95% CI - 95% confidence interval; DSM-IV - Diagnostic and Statistical Manual of Mental Disorders, fourth edition; ICD-10 - International Classification of Diseases, Tenth Revision; N - number; NINCDS-ADRDA - National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer’s Disease and Related Disorders Association; SD - standard deviation

APPENDICES

Appendix 1. Sources searched and search strategies

<table>
<thead>
<tr>
<th>Source and platform</th>
<th>Search strategy</th>
<th>Hits retrieved</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cochrane Dementia and Cognitive Improvement Group DTA register (see bottom of this table for more information in the search narrative) [Date of most recent search 12 March 2019]</td>
<td>1. &quot;mini-Cog&quot; [all fields] 2. minicog [all fields] 3. (MCE and (cognit* OR dement* OR screen* OR Alzheimer*)) [all fields] 3. or/1-3</td>
<td>Sept 2012: 452 Jul 2013: 34 Feb 2015: 7 Mar 2019: 7</td>
</tr>
<tr>
<td>MEDLINE In-Process and Other Non-Indexed Citations and MEDLINE 1950 to present (OvidSP) [Date of most recent search 12 March 2019]</td>
<td>1. &quot;mini-Cog&quot;.ti,ab. 2. minicog.ti,ab. 3. (MCE and (cognit* OR dement* OR screen* OR Alzheimer*)).ti,ab 3. or/1-3</td>
<td>Sept 2012: 91 Jul 2013: 12 Feb 2015: 31 Mar 2019: 67</td>
</tr>
<tr>
<td>Embase 1974 to 11 March 2019 (OvidSP) [Date of most recent search 12 March 2019]</td>
<td>1. &quot;mini-cog*&quot;.mp. 2. minicog*.mp. 3. (MCE and (cognit* OR dement* OR screen* OR Alzheimer*)).ti,ab 3. or/1-3</td>
<td>Sept 2012: 96 Jul 2013: 37 Feb 2015: 80 Mar 2019: 205</td>
</tr>
</tbody>
</table>
### Mini-Cog for the diagnosis of Alzheimer’s disease dementia and other dementias within a secondary care setting (Review)

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

**Search narrative:** The searches focus on a single concept - the index test. However, in order to ensure additional sensitivity, a broad search for neuropsychological and cognitive tests is run to populate the Cochrane Dementia and Cognitive Improvement Group’s DTA register. This search is run every six months in MEDLINE and Embase (OvidSP). The most recent search was run in February 2019. The MEDLINE search can be seen below:

1. “word recall”.ti,ab.
2. (“7-minute screen” or “seven-minute screen”).ti,ab.
3. (“6 item cognitive impairment test” or “six-item cognitive impairment test”).ti,ab.
4. “6 CIT”.ti,ab.
5. “AB cognitive screen”.ti,ab.
8. AD8.ti,ab.
11. “brief alzheimer* screen”.ti,ab.
12. “brief cognitive scale”.ti,ab.
15. “community screening interview for dementia”.ti,ab.
18. “cognitive capacity screening examination”.ti,ab.
21. (“Dem Tect” or DemTect).ti,ab.
22. “object memory evaluation”.ti,ab.
23. “IQCODE”.ti,ab.
24. “mattis dementia rating scale”.ti,ab.
25. “memory impairment screen”.ti,ab.

<table>
<thead>
<tr>
<th>Source</th>
<th>Topic</th>
<th>Date of most recent search</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Total before automated de-duplication</strong></td>
<td></td>
<td></td>
<td>1658</td>
</tr>
<tr>
<td><strong>Total after automated de-duplication and first assessment</strong></td>
<td></td>
<td></td>
<td>468</td>
</tr>
</tbody>
</table>
27. “mini-cog”.ti,ab.
29. “mmse”.ti,ab.
30. “modified mini-mental state exam”.ti,ab.
31. “3MS”.ti,ab.
32. “neurobehavioral cognitive status exam*”.ti,ab.
33. “cognistat”.ti,ab.
34. “quick cognitive screening test”.ti,ab.
35. “QCST”.ti,ab.
36. “rapid dementia screening test”.ti,ab.
37. “RDST”.ti,ab.
38. “repeatably for the assessment of neuropsychological status”.ti,ab.
40. “rowland universal dementia assessment scale”.ti,ab.
41. “rudas”.ti,ab.
42. “self-administered gerocognitive exam*”.ti,ab.
43. (“self-administered” and “SAGE”).ti,ab.
44. “self-administered computerized screening test for dementia”.ti,ab.
45. “short and sweet screening instrument”.ti,ab.
46. “sassi”.ti,ab.
47. “short cognitive performance test”.ti,ab.
49. (“six item screener” or “6-item screener”).ti,ab.
50. “short memory questionnaire”.ti,ab.
51. (“short memory questionnaire” and “SMQ”).ti,ab.
52. “short orientation memory concentration test”.ti,ab.
53. “s-omc”.ti,ab.
54. “short blessed test”.ti,ab.
55. “short portable mental status questionnaire”.ti,ab.
56. “spmsq”.ti,ab.
57. “short test of mental status”.ti,ab.
58. “telephone interview of cognitive status modified”.ti,ab.
59. “tics-m”.ti,ab.
60. “trail making test”.ti,ab.
61. “verbal fluency categories”.ti,ab.
63. “general practitioner assessment of cognition”.ti,ab.
64. “GPCOG”.ti,ab.
67. “time and change test”.ti,ab.
68. “modified world test”.ti,ab.
69. “symptoms of dementia screener”.ti,ab.
70. “dementia questionnaire”.ti,ab.
71. “7MS”.ti,ab.
72. (“concord informant dementia scale” or CIDS).ti,ab.
73. (SAPH or “dementia screening and perceived harm”).ti,ab.
74. or/1-73
75. exp Dementia/
76. Delirium, Dementia, Amnestic, Cognitive Disorders/
77. dement*.ti,ab.
78. alzheimer*.ti,ab.
79. AD.ti,ab.
80. ("lewy body" or DLB or LBD or FTD or FTLD or "frontotemporal lobar degeneration" or "frontotemporal dement*").ti,ab.
81. "cognit* impair*".ti,ab.
82. (cognit* adj4 (disorder* or declin* or fail* or function* or degenerat* or deteriorat*)).ti,ab.
83. (memory adj3 (complain* or declin* or function* or disorder*)).ti,ab.
84. or/75-83
85. exp "sensitivity and specificity"/
86. "reproducibility of results"/
87. (predict* adj3 (dement* or AD or alzheimer*)).ti,ab.
88. (identif* adj3 (dement* or AD or alzheimer*)).ti,ab.
89. (discriminat* adj3 (dement* or AD or alzheimer*)).ti,ab.
90. (distinguish* adj3 (dement* or AD or alzheimer*)).ti,ab.
91. (differenti* adj3 (dement* or AD or alzheimer*)).ti,ab.
92. diagnos*.ti.
93. di.fs.
94. sensitivit*.ab.
95. specificit*.ab.
96. (ROC or "receiver operat*").ab.
97. Area under curve/
98. ("Area under curve" or AUC).ab.
99. (detect* adj3 (dement* or AD or alzheimer*)).ti,ab.
100. sROC.ab.
101. accura*.ti,ab.
102. (likelihood adj3 (ratio* or function*)).ab.
103. (conver* adj3 (dement* or AD or alzheimer*)).ti,ab.
104. ((true or false) adj3 (positive* or negative*)).ab.
105. ((positive* or negative* or false or true) adj3 rate*).ti,ab.
106. or/85-105
107. exp dementia/di
108. Cognition Disorders/di [Diagnosis]
109. Memory Disorders/di
110. or/107-109
111. *Neuropsychological Tests/
112. *Questionnaires/
113. Geriatric Assessment/mt
114. *Geriatric Assessment/
115. Neuropsychological Tests/mt, st
116. "neuropsychological test*".ti,ab.
117. (neuropsychological adj (assess* or evaluat* or test*)).ti,ab.
118. (neuropsychological adj (assess* or evaluat* or test* or exam* or battery)).ti,ab.
119. Self report/
120. self-assessment/ or diagnostic self evaluation/
121. Mass Screening/
122. early diagnosis/
123. or/111-122
Appendix 2. Assessment of methodological quality table QUADAS-2 tool

<table>
<thead>
<tr>
<th>Domain</th>
<th>Patient selection</th>
<th>Index test</th>
<th>Reference standard</th>
<th>Flow and timing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Description</td>
<td>Describe methods of patient selection: describe included patients (prior testing, presentation, intended use of index test and setting)</td>
<td>Describe the index test and how it was conducted and interpreted</td>
<td>Describe the reference standard and how it was conducted and interpreted</td>
<td>Describe any patients who did not receive the index test(s) and/or reference standard or who were excluded from the 2 x 2 table (refer to flow diagram): describe the time interval and any interventions between index test(s) and reference standard</td>
</tr>
<tr>
<td>Signalling questions (yes, no, unclear)</td>
<td>Was a consecutive or random sample of patients enrolled?</td>
<td>Were the index test results interpreted without knowledge of the results of the reference standard? If a threshold was used, was it prespecified?</td>
<td>Is the reference standard likely to correctly classify the target condition? Were the reference standard results interpreted without knowledge of the results of the index test?</td>
<td>Was there an appropriate interval between index test(s) and reference standard? Did all patients receive the same reference standard? Were all patients included in the analysis?</td>
</tr>
<tr>
<td>Risk of bias (high, low, unclear)</td>
<td>Could the selection of patients have introduced bias?</td>
<td>Could the conduct or interpretation of the index test have introduced bias?</td>
<td>Could the reference standard, its conduct, or its interpretation have introduced bias?</td>
<td>Could the patient flow have introduced bias?</td>
</tr>
<tr>
<td>Concerns regarding applicability (high, low, unclear)</td>
<td>Are there concerns that the included patients do not match the review question?</td>
<td>Are there concerns that the index test, its conduct, or its interpretation differ from the review question?</td>
<td>Are there concerns that the target condition as defined by the reference standard does not match the review question?</td>
<td>-</td>
</tr>
</tbody>
</table>
Appendix 3. Anchoring statements to assist with assessment of risk of bias

Domain 1: patient selection

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

Was a consecutive or random sample of patients enrolled?
Where sampling is used, the methods least likely to cause bias are consecutive sampling or random sampling, which should be stated and/or described. Non-random sampling or sampling based on volunteers is more likely to be at high risk of bias.

Was a case-control design avoided?
Case-control study designs have a high risk of bias, but are sometimes the only studies available especially if the index test is expensive and/or invasive. Nested case-control designs (systematically selected from a defined population cohort) are less prone to bias, but they will still narrow the spectrum of patients that receive the index test. Study designs (both cohort and case-control) that may also increase bias are those designs where the study team deliberately increase or decrease the proportion of participants with the target condition, for example a population study may be enriched with extra dementia participants from a secondary care setting.

Did the study avoid inappropriate exclusions?
We will automatically grade the study as unclear if exclusions are not detailed (pending contact with study authors). Where exclusions are detailed, the study will be graded as 'low risk' if the review authors consider the exclusions to be appropriate. Certain exclusions common to many studies of dementia are: medical instability; terminal disease; alcohol/substance misuse; concomitant psychiatric diagnosis; other neurodegenerative condition. However, if 'difficult to diagnose' groups are excluded this may introduce bias, so exclusion criteria must be justified. For a community sample we would expect relatively few exclusions. We will label post hoc exclusions 'high risk' of bias.

Applicability: are there concerns that the included patients do not match the review question? (high, low, unclear)
The included patients 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. We will classify studies that use very selected participants or subgroups as having low applicability, unless they are intended to represent a defined target population, for example people with memory problems referred to a specialist and investigated by lumbar puncture.

Domain 2: index test

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

Were the index test results interpreted without knowledge of the reference standard?
Terms such as 'blinded' or 'independently and without knowledge of' are sufficient; full details of the blinding procedure are not required. This item may be scored as 'low risk' if explicitly described or if there is a clear temporal pattern to the order of testing that precludes the need for formal blinding, that is all (neuropsychological test) assessments were performed before the dementia assessment. As most neuropsychological tests are administered by a third party, knowledge of dementia diagnosis may influence their ratings; tests that are self-administered, for example using a computerised version, may have less risk of bias.
Were the index test thresholds prespecified?
For neuropsychological scales there is usually a threshold above which participants are classified as 'test positive'; this may be referred to as threshold, clinical cut-off, or dichotomisation point. Different thresholds are used in different populations. A study is classified as at higher 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 not applicable.

Were sufficient data on (neuropsychological test) application given for the test to be repeated in an independent study?
Particular points of interest include method of administration (e.g. self-completed questionnaire versus direct questioning interview); nature of informant; language of assessment. If a novel form of the index test is used, for example a translated questionnaire, details of the scale should be included and a reference given to an appropriate descriptive text, and there should be evidence of validation.

Applicability: are there concerns that the index test, its conduct, or its interpretation differ from the review question? (high, low, unclear)
Variations in the length, structure, language, and/or administration of the index test may all affect applicability if they vary from those specified in the review question.

Domain 3: reference standard

Risk of bias: could the reference standard, its conduct, or its interpretation have introduced bias? (high, low, unclear)
Is the reference standard likely to correctly classify the target condition?
Commonly used international criteria to assist with clinical diagnosis of dementia include those detailed in the Diagnostic and Statistical Manual of Mental Disorders IV (DSM-IV) and 10th Revision of the International Classification of Diseases (ICD-10). Criteria specific to dementia subtypes include but are not limited to National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer's Disease and Related Disorders Association (NINCDS-ADRDA) criteria for Alzheimer's dementia; McKeith criteria for Lewy body dementia; Lund criteria for frontotemporal dementias; and the National Institute of Neurological Disorders and Stroke and Association Internationale pour la Recherché et l'Enseignement en Neurosciences (NINDS-AIREN) criteria for vascular dementia. Where the criteria used for assessment are not familiar to the review authors and the Cochrane Dementia and Cognitive Improvement Group, this item should be classified as 'high risk of bias'.

Were the reference standard results interpreted without knowledge of the results of the index test?
Terms such as 'blinded' or 'independent' are sufficient; full details of the blinding procedure are not required. This may be scored as 'low risk' if explicitly described or if there is a clear temporal pattern to order of testing, that is all dementia assessments performed before (neuropsychological test) 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.

Was sufficient information on the method of dementia assessment given for the assessment to be repeated in an independent study?
Particular points of interest for dementia assessment include the training and expertise of the assessor; whether additional information was available to inform the diagnosis (e.g. neuroimaging, other neuropsychological test results); and whether this was available for all participants. High risk of bias if method of dementia assessment not described.
Applicability: are there concerns that the target condition as defined by the reference standard does not match the review question? (high, low, unclear)

It is possible that some methods of dementia assessment, although valid, may diagnose a far smaller or larger proportion of participants with disease than in usual clinical practice. For example, the current reference standard for vascular dementia may underdiagnose compared to usual clinical practice. In this instance the item should be rated as having poor applicability.

Domain 4: patient flow and timing

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

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

For a cross-sectional study design, there is potential for the participant to change between assessments, however dementia is a slowly progressive disease that is not reversible. The ideal scenario would be a same-day assessment, but longer periods of time (e.g. several weeks or months) are unlikely to lead to a high risk of bias. For delayed-verification studies, the index and reference tests are necessarily separated in time given the nature of the condition.

Did all participants receive the same reference standard?

There may be scenarios where participants who score ‘test positive’ on the index test have a more detailed assessment for the target condition. Where dementia assessment (or reference standard) differs between participants, this should be classified as high risk of bias.

Were all participants included in the final analysis?

Attrition will vary with study design. Delayed-verification studies will have higher attrition than cross-sectional studies due to mortality, and it is likely to be greater in participants with the target condition. Dropouts (and missing data) should be accounted for. Attrition that is higher than expected (compared to other similar studies) should be treated as high risk of bias. We have defined a cut-off of greater than 20% attrition as being high risk, but this will be highly dependent on the length of follow-up in individual studies.

CONTRIBUTIONS OF AUTHORS

CCHC assisted with protocol development, extraction of data, contributing to drafts of the manuscript.

BAF assisted with protocol development, extraction of data, contributing to drafts of the review.

JKB performed title selection, data extraction, ‘Risk of bias’ assessment, and completed the first draft of the manuscript.

NS assisted with protocol development, assisting with the database searches, and providing comments on drafts of the manuscript.

SSG assisted with protocol development and review of drafts of the manuscript.

NH assisted with protocol development and reviews of drafts of the manuscript.

VN assisted with protocol development, reviews of drafts of the manuscript, and methodological input into presentation of statistical results.

TJQ performed title selection, data extraction, ‘Risk of bias’ assessment, and provided supervision and comments on the final manuscript.

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

DPS assisted with protocol development, data extraction, ‘Risk of bias assessment, supervision, and comments on drafts of the manuscript and the final manuscript.
DECLARATIONS OF INTEREST

Calvin CH Chan: none known
Bruce A Fage: none known
Jennifer K Burton: none known
Nadja Smailagic: none known
Sudeep S Gill: none known
Nathan Herrmann: received consultancy payment from Pfizer
Vasilis Nikolaou: none known
Terry J Quinn: none known
Anna H Noel-Storr: none known
Dallas P Seitz: participated as a site investigator on a trial sponsored by Hoffmann La Roche

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 Kirsty 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 (NHS), or the Department of Health.

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

The protocol stated in 'Selection of studies' that studies had to: “Report estimates of test reproducibility, if completed within the study”. In response to peer reviewers’ comments, this text was moved to the data extraction section as it was not a mandatory item to determine study eligibility.

We updated statistical analysis text, as some studies did not provide numbers of true and false positives and negatives, but gave summary test accuracy data which was entered into Review Manager 5.

We did not conduct planned quantitative analyses, sensitivity analyses, and investigations of heterogeneity due to the methodological heterogeneity identified which made pooling of data inappropriate.