

Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE) for the diagnosis of dementia within a secondary care setting (Review)

Harrison JK, Fearon P, Noel-Storr AH, McShane R, Stott DJ, Quinn TJ



**THE COCHRANE
COLLABORATION®**

This is a reprint of a Cochrane review, prepared and maintained by The Cochrane Collaboration and published in *The Cochrane Library* 2015, Issue 3

<http://www.thecochranelibrary.com>

WILEY

TABLE OF CONTENTS

HEADER	1
ABSTRACT	1
PLAIN LANGUAGE SUMMARY	2
BACKGROUND	3
OBJECTIVES	6
METHODS	6
RESULTS	10
Figure 1.	11
Figure 2.	13
Figure 3.	15
Figure 4.	17
Figure 5.	19
Figure 6.	21
DISCUSSION	29
AUTHORS' CONCLUSIONS	31
ACKNOWLEDGEMENTS	32
REFERENCES	32
CHARACTERISTICS OF STUDIES	39
DATA	61
Test 1. All studies IQCODE 3.3 or closest.	62
Test 2. All 16-item IQCODE.	63
Test 3. All 26-item IQCODE.	63
Test 4. IQCODE 3.3 Threshold.	64
Test 5. IQCODE 3.4 Threshold.	64
Test 6. IQCODE 3.5 Threshold.	65
Test 7. IQCODE 3.6 Threshold.	65
Test 8. IQCODE >3.6 Threshold.	66
Test 9. 16-item IQCODE 3.3 Threshold.	66
Test 10. 16-item IQCODE 3.4 Threshold.	67
Test 11. 16-item IQCODE 3.5 Threshold.	67
Test 12. 16-item IQCODE 3.6 Threshold.	68
Test 13. 26-item IQCODE 3.3 Threshold.	68
Test 14. 26-item IQCODE 3.4 Threshold.	69
Test 15. 26-item IQCODE 3.6 Threshold.	69
Test 16. Sensitivity analysis removing Goncalves.	70
Test 17. Sensitivity analysis removing low average age.	71
ADDITIONAL TABLES	71
APPENDICES	72
CONTRIBUTIONS OF AUTHORS	93
DECLARATIONS OF INTEREST	94
SOURCES OF SUPPORT	94
DIFFERENCES BETWEEN PROTOCOL AND REVIEW	94

[Diagnostic Test Accuracy Review]

Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE) for the diagnosis of dementia within a secondary care setting

Jennifer K Harrison¹, Patricia Fearon², Anna H Noel-Storr³, Rupert McShane³, David J Stott⁴, Terry J Quinn⁴

¹Department of Cardiovascular Sciences, University of Leicester, Leicester, UK. ²Academic Section of Geriatric Medicine, University of Glasgow, Glasgow, UK. ³Radcliffe Department of Medicine, University of Oxford, Oxford, UK. ⁴Institute of Cardiovascular and Medical Sciences, University of Glasgow, Glasgow, UK

Contact address: Terry J Quinn, Institute of Cardiovascular and Medical Sciences, University of Glasgow, New Lister Building, Glasgow Royal Infirmary, Glasgow, Strathclyde, G4 0SFR, UK. Terry.Quinn@glasgow.ac.uk.

Editorial group: Cochrane Dementia and Cognitive Improvement Group.

Publication status and date: New, published in Issue 3, 2015.

Review content assessed as up-to-date: 16 January 2013.

Citation: Harrison JK, Fearon P, Noel-Storr AH, McShane R, Stott DJ, Quinn TJ. Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE) for the diagnosis of dementia within a secondary care setting. *Cochrane Database of Systematic Reviews* 2015, Issue 3. Art. No.: CD010772. DOI: 10.1002/14651858.CD010772.pub2.

Copyright © 2015 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

ABSTRACT

Background

The diagnosis of dementia relies on the presence of new-onset cognitive impairment affecting an individual's functioning and activities of daily living. The Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE) is a questionnaire instrument, completed by a suitable 'informant' who knows the patient well, designed to assess change in functional performance secondary to cognitive change; it is used as a tool to identifying those who may have dementia.

In secondary care there are two specific instances where patients may be assessed for the presence of dementia. These are in the general acute hospital setting, where opportunistic screening may be undertaken, or in specialist memory services where individuals have been referred due to perceived cognitive problems. To ensure an instrument is suitable for diagnostic use in these settings, its test accuracy must be established.

Objectives

To determine the diagnostic accuracy of the informant-based questionnaire IQCODE, for detection of all-cause (undifferentiated) dementia in adults presenting to secondary-care services.

Search methods

We searched the following sources on the 28th of January 2013: ALOIS (Cochrane Dementia and Cognitive Improvement Group), MEDLINE (Ovid SP), EMBASE (Ovid SP), PsycINFO (Ovid SP), BIOSIS Previews (Thomson Reuters Web of Science), Web of Science Core Collection (includes Conference Proceedings Citation Index) (Thomson Reuters Web of Science), CINAHL (EBSCOhost) and LILACS (BIREME). We also searched sources specific to diagnostic test accuracy: MEDION (Universities of Maastricht and Leuven); DARE (Database of Abstracts of Reviews of Effects - via the Cochrane Library); HTA Database (Health Technology Assessment Database via the Cochrane Library) and ARIF (Birmingham University). We also checked reference lists of relevant studies and reviews, used searches of known relevant studies in PubMed to track related articles, and contacted research groups conducting work on IQCODE for dementia diagnosis to try to find additional studies. We developed a sensitive search strategy; search terms were designed

Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE) for the diagnosis of dementia within a secondary care setting (Review)

Copyright © 2015 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

to cover key concepts using several different approaches run in parallel and included terms relating to cognitive tests, cognitive screening and dementia. We used standardised database subject headings such as MeSH terms (in MEDLINE) and other standardised headings (controlled vocabulary) in other databases, as appropriate.

Selection criteria

We selected those studies performed in secondary-care settings, which included (not necessarily exclusively) IQCODE to assess for the presence of dementia and where dementia diagnosis was confirmed with clinical assessment. For the 'secondary care' setting we included all studies which assessed patients in hospital (e.g. acute unscheduled admissions, referrals to specialist geriatric assessment services etc.) and those referred for specialist 'memory' assessment, typically in psychogeriatric services.

Data collection and analysis

We screened all titles generated by electronic database searches, and reviewed abstracts of all potentially relevant studies. Two independent assessors checked full papers for eligibility and extracted data. We determined quality assessment (risk of bias and applicability) using the QUADAS-2 tool, and reporting quality using the STARD tool.

Main results

From 72 papers describing IQCODE test accuracy, we included 13 papers, representing data from 2745 individuals ($n = 1413$ (51% with dementia)). Pooled analysis of all studies using data presented closest to a cut-off of 3.3 indicated that sensitivity was 0.91 (95% CI 0.86 to 0.94); specificity 0.66 (95% CI 0.56 to 0.75); the positive likelihood ratio was 2.7 (95% CI 2.0 to 3.6) and the negative likelihood ratio was 0.14 (95% CI 0.09 to 0.22).

There was a statistically significant difference in test accuracy between the general hospital setting and the specialist memory setting ($P = 0.019$), suggesting that IQCODE performs better in a 'general' setting.

We found no significant differences in the test accuracy of the short (16-item) versus the 26-item IQCODE, or in the language of administration.

There was significant heterogeneity in the included studies, including a highly varied prevalence of dementia (10.5% to 87.4%). Across the included papers there was substantial potential for bias, particularly around sampling of included participants and selection criteria, which may limit generalisability. There was also evidence of suboptimal reporting, particularly around disease severity and handling indeterminate results, which are important if considering use in clinical practice.

Authors' conclusions

The IQCODE can be used to identify older adults in the general hospital setting who are at risk of dementia and require specialist assessment; it is useful specifically for ruling out those without evidence of cognitive decline. The language of administration did not affect test accuracy, which supports the cross-cultural use of the tool. These findings are qualified by the significant heterogeneity, the potential for bias and suboptimal reporting found in the included studies.

PLAIN LANGUAGE SUMMARY

Assessment of changes in memory and everyday function in older people using a structured questionnaire, the IQCODE

Improving how we assess people who may have dementia is a health and social care priority, recent initiatives to increase dementia diagnosis rates have attracted considerable attention. At present we do not have an agreed approach to dementia testing. There are many tests which can help us identify people with the memory and thinking problems suggestive of dementia, but there is no agreement on which tests are best. It is possible that some tests may be better suited to certain healthcare settings than others.

Our review was interested in the accuracy of a questionnaire-based assessment for dementia, called the IQCODE (Informant Questionnaire for Cognitive Decline in the Elderly). We describe how useful the IQCODE is when used in a hospital setting. Under the umbrella term 'hospital' we include specialist memory clinics and old-age psychiatry units as well as general hospital clinics and wards and the older people's services within them.

We searched electronic databases of published research studies, looking for all studies of IQCODE in a hospital setting. We searched from the first available papers in scientific databases up to and including January 2013.

We found 13 relevant studies which had results suitable to be combined in a single analysis. Of these papers, six (1352 participants) described studies conducted in “specialist” services such as memory clinics or wards. Three papers (566 participants) described studies conducted in general older adult services and four studies (827 participants) included both specialist and general services.

Summarising the available papers, we found that IQCODE was useful for ‘ruling out’ possible dementia in the general hospital setting. This means if a person has a low score on IQCODE testing they probably do not have dementia. IQCODE was less useful in specialist memory clinics and psychiatry wards. We also found that a short version of the IQCODE gave similar results to the traditional longer version.

As part of our assessment we looked at whether the design of the available studies was suitable for the study question. We found several instances where the design of the study could be improved. For example, seven of the thirteen studies only included a selection of all the people attending the service who could have been assessed with IQCODE. We also looked at how well researchers reported the conduct and results of their studies. Again, there were many instances where the reporting could be improved. A common issue was not describing the severity of memory and thinking problems in those thought to have dementia, only reported in three of the included studies.

In summary, IQCODE may be a useful tool for assessing adults for possible dementia. There are still a number of unanswered questions around how useful IQCODE may be in hospital settings. For example, before we start using IQCODE routinely we need to describe if it is practical and acceptable to hospital staff, to patients and to their carers.

The review was performed by a team based in research centres in the UK (Glasgow, Leicester, Oxford). We had no external funding specific to this study and we have no conflicts of interest that may have influenced our assessment of the research data.

BACKGROUND

Dementia is a chronic, progressive, neurodegenerative syndrome that is a substantial and growing public health concern (Hebert 2003; Hebert 2013; Prince 2013). Depending on the case definition employed, contemporary estimates of dementia prevalence in the United States are in the range 2.5 to 4.5 million individuals. Dementia is predominantly a disease of older adults, with a 10% prevalence in adults aged over 65, increasing to around 30% in adults aged over 85 (Ferri 2005). Changes in population demographics will be accompanied by increases in dementia incidence and prevalence. Consensus opinion based on current epidemiological trends is of a doubling in dementia prevalence every 20 years, with a global prevalence of around 81 million cases by 2040. Dementia is not limited to ‘Western’ nations and an increasing prevalence is particularly marked in countries such as China and India (Ferri 2005). Recent population follow-up studies have cast doubt upon earlier estimates of increasing dementia incidence (Matthews 2013); however even with these lower predictions of incidence, the absolute number of individuals with dementia in society will be substantial, and accurate diagnosis remains a public health priority.

A key element of effective management in dementia is a firm diagnosis. Recent guidelines place emphasis on early diagnosis to facilitate improved management and to allow informed discussions and

planning with patients and carers. Given the projected global increase in dementia prevalence, there is a potential tension between the clinical requirements for robust diagnosis at the individual patient level and the need for equitable, easy access to diagnosis at a population level. The ideal would be expert, multidisciplinary assessment informed by various supplementary investigations. Such an approach may be possible for assessment of challenging cases in ‘specialist’ settings, but is not practical or feasible for all people with possible cognitive decline.

In practice a two-stage process is often employed, with initial screening or ‘triage’ assessments, suitable for use by non-specialists, used to select those people who require further detailed assessment (Boustani 2003). Various tools for initial cognitive screening have been described (Brodsky 2002; Folstein 1975; Galvin 2005). Regardless of the methods employed, there is scope for improvement, with observational work suggesting that many people with dementia are not diagnosed (Chodosh 2004; Valcour 2000). UK national dementia strategies have focused on secondary (hospital) care, particularly unscheduled admissions, as a setting where there may be scope for opportunistic dementia screening (Shenkin 2014).

Screening assessment often takes the form of brief, direct cognitive testing. Such an approach will only provide a ‘snapshot’ of cognitive function. However, a defining feature of dementia is cognitive

or neuropsychological change over time. Patients themselves may struggle to make an objective assessment of personal change, and so an attractive approach is to question collateral sources with sufficient knowledge of the patient. Informant-based interviews have been described that aim to retrospectively assess change in function. An instrument prevalent in research and clinical practice is the Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE) (Jorm 1988) and this is the focus of our review.

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

Target condition being diagnosed

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

Dementia is a syndrome characterised by cognitive or neuropsychological decline sufficient to interfere with usual functioning. The neurodegeneration and clinical manifestations of dementia are progressive and at present there is no 'cure', although numerous interventions to slow or arrest cognitive decline have been described, for example, pharmacotherapy such as acetylcholinesterase inhibitors; memantine; or cognitive rehabilitation therapies (Bahar-Fuchs 2013; Birks 2006; McShane 2006).

Dementia remains a clinical diagnosis, based on history from the patient and suitable collateral sources, and direct examination including cognitive assessment. We have chosen expert clinical diagnosis as our 'gold standard' (reference standard) for describing IQCODE properties, as we believe this is most in keeping with current diagnostic criteria and best practice. We recognise that there is no universally accepted, ante-mortem, gold standard diagnostic strategy. Although some would argue that the true gold standard would be neuropathological data, for the purposes of testing diagnostic accuracy in secondary care, limiting analysis to those studies with neuropathologically confirmed diagnosis is likely to yield limited and highly selected data. Furthermore, recent studies have suggested only a modest correlation between neuropathological changes and clinical cognitive phenotype in older age. There are several studies that have described cognitive impairment in 'normal brains' and multiple pathological changes with preserved cognition (Matthews 2009; Wharton 2015). We also recognise that clinical-neuropathological correlations are less apparent in mixed dementia and older people, who form the majority with dementia in the hospital setting (Savva 2009).

Criteria for diagnosis of dementia are evolving in line with improvements in our understanding of the underlying pathophysiological processes. Various biomarkers based on biological fluid assays or functional/quantitative neuroimaging have shown promise but to date are not accepted or validated as independent diagnostic tests (McKhann 2011). Here a distinction must be made between

dementia diagnosis in clinical practice and dementia diagnosis for clinical research. These novel biomarker and imaging techniques may be increasingly used in secondary-care settings and may be stipulated in research diagnostic criteria but are not absolutely required for clinical diagnosis.

The label of dementia encompasses varying pathologies, of which Alzheimer's disease is the most common (Savva 2009). For our reference standard of clinical diagnosis, we accepted a dementia diagnosis made according to any of the internationally accepted diagnostic criteria, with exemplars being the various iterations of the World Health Organization, International Classification of Diseases (ICD) and American Psychiatric Association Diagnostic and Statistical Manual of Mental Disorders (DSM) for all-cause dementia and subtypes (Appendix 1) and the various diagnostic criteria available for specific dementia subtypes, i.e. NINCDS-ADRDA (National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer's Disease and Related Disorders Association) criteria for Alzheimer's dementia (McKhann 1984); McKeith criteria for Lewy Body dementia (McKeith 2005); Lund criteria for frontotemporal dementias (McKhann 2001); and the NINDS-AIREN (National Institute of Neurological Disorders and Stroke and the Association Internationale pour la Recherche et l'Enseignement en Neurosciences) criteria for vascular dementia (Erkinjuntti 2000; Roman 1993). We considered all-cause dementia as the target condition for our primary analysis of diagnostic test accuracy, recognising that in a selected cohort referred to hospital services there may be a greater spectrum of differing dementia pathologies than is seen in unselected community cohorts. We have not defined preferred diagnostic criteria for rarer forms of dementia (e.g. alcohol-related; HIV-related; prion disease-related), which were considered under our rubric of 'all-cause dementia' and were not considered separately.

The label 'dementia' can also span a range of disease severities, from mild to end-stage disease. We recognise that the diagnostic properties of a tool such as IQCODE vary depending on disease stage; for example, a patient is more likely to screen positive when disease is advanced and diagnosis is clear. For our primary analysis we included any dementia diagnosis at any stage of disease. Definitions pertinent to various stages of the dementia 'journey' are also described: a preclinical stage occurring years before disease is manifest, which may be characterised by changes in one or more disease biomarkers (Sperling 2011); a stage of mild cognitive impairment (MCI) where problems with cognition are noticed by the patient or others but the disease is not sufficiently advanced to warrant a diagnostic label of dementia (Albert 2011); and finally established dementia as defined above (McKhann 2011). We have not included diagnoses of preclinical and MCI states in this review.

Index test(s)

Our index test was the Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE).

The IQCODE was originally described as a 26-item informant questionnaire that seeks to retrospectively ascertain change in cognitive and functional performance over a 10-year time period (Jorm 1988). IQCODE is designed as a brief screen for potential dementia, usually administered as a questionnaire given to the relevant proxy. For each item the chosen proxy scores change on a five-point ordinal hierarchical scale, with responses ranging from 1: 'has become much better' to 5: 'has become much worse'. This gives a sum-score of 26 to 130 that can be averaged by the total number of completed items to give a final score of 1.0 to 5.0, where higher scores indicate greater decline.

First described in 1989, use of IQCODE is prevalent in both clinical practice and research (Holsinger 2007). A literature describing the properties of IQCODE is available including studies of non-English IQCODE translations; studies in specific patient populations; and modifications to the original 26-item direct informant interview (Isella 2002; Jorm 1989; Jorm 2004). Versions of the IQCODE have been produced in other languages, including Chinese, Dutch, Finnish, French, Canadian French, German, Italian, Japanese, Korean, Norwegian, Polish, Spanish and Thai (www.anu.edu.au/iqcode/). A shortened 16-item version is also available (Jorm 1994); this modified IQCODE is common in clinical practice and has been recommended as the preferred IQCODE format (Jorm 2004).

For this review the term 'IQCODE' refers to the original 26-item questionnaire as described by Jorm 1988. Other versions of IQCODE were described according to the number of items and administration language (e.g. a 16-item IQCODE for Spanish speakers is described as 'IQCODE-16 Spanish'). Other authors have also shortened the timeframe for assessment with a two-year version of the IQCODE having been described (Ehrensperger 2010).

Although we describe the utility of IQCODE for dementia diagnosis, IQCODE used in isolation is not suitable for establishing a clinical diagnosis. The value of IQCODE is in selecting people who require more definitive assessment. Use of IQCODE in hospital settings is valid, as new diagnostic criteria for dementia make explicit reference to documenting decline and involving collateral informants, emphasising the potential utility of an informant interview tool such as IQCODE.

The full 26- and 16-item versions of IQCODE with scoring rules are available in [Appendix 2](#) and [Appendix 3](#).

The purpose of this review is to describe the diagnostic test accuracy of IQCODE. Other important psychometric properties for a tool that is to be used in clinical practice include reliability, responsiveness and acceptability. Contemporary reviews of the 26- and 16-item IQCODE suggest good inter-rater reliability with retest kappa 0.96 at three days and 0.75 at one year (Jorm 2004; Tang 2003). Internal consistency is uniformly high with Cronbach's alpha in the range 0.93 to 0.97 (Jorm 2004). Validation work has

included validation against measures of cognitive change, neuropathology, neuroimaging, and neuropsychological assessment (Cordoliani-Mackowiak 2003; Jorm 2000; Jorm 2004; Rockwood 1998). Factor analysis suggests that the scale measures a common factor of cognitive decline (Jorm 2004). There are fewer published data on the psychometric properties of other 'short' forms of IQCODE.

IQCODE cut-off scores suggestive of a potential dementia diagnosis will vary with the demographics of the population tested. In the original development and validation work, normative data were described, with a total score above 93 or an average score above 3.31 indicative of cognitive impairment (Jorm 2004). These data were based on community samples and the thresholds with greatest utility in a selected secondary-care cohort may differ. There is no consensus on the optimal threshold and various authors have described improved diagnostic accuracy with other cut-offs. In setting thresholds for any diagnostic test there is a trade-off between sensitivity and specificity, with the preferred values partly determined by the purpose of the test. For specialist memory services, a more sensitive test may be preferred as a degree of filtering of non-dementia diagnoses will have already occurred. For general hospital services, where there may be confounders of delirium or disability, a more specific test may be preferred.

IQCODE has a number of features that make it attractive for clinical and research use, particularly in a secondary-care setting. The questions have an immediacy and relevance that is likely to appeal to users. Assessment and (informant) scoring take around five to seven minutes and as the scale is not typically interviewer-administered it requires minimal training in application and scoring (Holsinger 2007). There are data to suggest that, compared to standard direct assessments, IQCODE may be less prone to bias from cultural norms and previous level of education (Jorm 2004).

Clinical pathway

Dementia develops over a trajectory of several years and screening tests may be performed at different stages in the dementia pathway. This review focuses on the secondary-care setting. This is effectively two related patient populations.

In 'general' secondary-care settings, people will have been referred for expert input but not exclusively due to memory complaints; there may have already been a degree of cognitive screening by the referrer. Opportunistic screening of adults presenting as unscheduled admissions to hospitals would be another secondary-care pathway.

The rubric of secondary care also includes those people referred to dementia/memory-specific services. This population will have a high prevalence of cognitive disorders and other physical and psychological health conditions; patients would be expected to have had a degree of cognitive assessment prior to referral. However, we recognise there will be no standardised approach to this pre-refer-

ral assessment and real-world studies have indicated a low level of pre-referral cognitive testing (Fisher 2007).

Alternative test(s)

Several other dementia screening and assessment tools have been described. Instruments commonly used in secondary-care settings include Folstein's mini-mental state examination (MMSE) (Folstein 1975); Montreal cognitive assessment (MoCA) (Nasreddine 2005); and the MiniCog (Borson 2000). These performance-based measures for cognitive screening all rely on comparing single or multi-domain cognitive testing against population-specific normative data. Copyright issues may preclude widespread use of certain tools.

Other informant interviews are also available. For example, the AD-8 is an eight-question tool, requiring dichotomous responses (yes or no) and testing for perceived change in memory, problem-solving, orientation and daily activities (Galvin 2005).

For this review we focused on papers that describe IQCODE diagnostic properties, and did not consider other cognitive screening/assessment tools. Where a paper describes IQCODE with an in-study comparison against another screening tool, we included the IQCODE data only. Where IQCODE is used in combination with another cognitive screening tool, we included the IQCODE data only.

Rationale

There is no consensus on the optimal screening test for dementia and the choice is currently dictated by experience with a particular instrument, time constraints and training. A better understanding of the diagnostic properties of various strategies would allow for an informed approach to testing. Critical evaluation of the evidence base for screening tests or other diagnostic markers is of major importance. Without a robust synthesis of the available information there is the risk that future research, clinical practice and policy will be built on erroneous assumptions about diagnostic validity. This is particularly pertinent to secondary care as healthcare systems increasingly see hospital admission as a window for opportunistic cognitive screening.

IQCODE is commonly used in practice and research; it is used internationally and is one of only a few validated informant-based screening/diagnostic tools. A literature describing the test accuracy of IQCODE in different settings is available, although some of these studies have been modest in size. Thus systematic review and, if possible, meta-analysis of the diagnostic properties of IQCODE is warranted.

Although we use the term 'diagnosis' in this review, we recognise that in practice IQCODE alone is not sufficient to make a diagnosis. Rather, IQCODE can be used to 'triage' people presenting with memory problems for further assessment or to inform a di-

agnosis in conjunction with direct patient assessment and investigations.

This review forms part of a body of work describing the diagnostic properties of commonly-used dementia tools. The Cochrane Dementia and Cognitive Improvement Group have reviews planned or underway for other commonly-employed dementia assessment scales (Appendix 4) and other IQCODE reviews are completed (Harrison 2014; Quinn 2014). At present we are conducting single-test review and meta-analysis. The intention, however, is then to collate these data, performing an overview allowing comparison of various test strategies.

OBJECTIVES

To determine the diagnostic accuracy of the informant-based questionnaire IQCODE, for detection of all-cause (undifferentiated) dementia in adults presenting to secondary-care services.

Secondary objectives

Where data were available we planned to describe the following:

1. The diagnostic accuracy of IQCODE at various prespecified thresholds. We recognise that various thresholds or cut-off scores have been used to define IQCODE screen-positive states. We described the properties of IQCODE for the following cut-off scores (rounded where necessary): 3.6; 3.5; 3.4; 3.3. These thresholds have been chosen to represent the range of cut-offs that are commonly used in practice and research; we have been inclusive in our choice of cut-off to maximise available data for review.
2. Accuracy of IQCODE for diagnosis of the commonest specific dementia subtype - Alzheimer's dementia.
3. Effects of heterogeneity on the reported diagnostic accuracy of IQCODE. Potential sources of heterogeneity that we aimed to explore included: age of cohort; case mix of cohort; reason for hospital consultation (dichotomised as 'memory' or 'non-memory' services); technical features of IQCODE; method of dementia diagnosis.

METHODS

Criteria for considering studies for this review

Types of studies

This review forms part of a suite of reviews describing IQCODE accuracy in various healthcare settings. We created a generic strategy for searching; selection; data extraction and analysis that would be applicable to all the proposed IQCODE reviews. For consistency with the other reviews we have used the same text descriptor in each, except where the methodology is specific to the setting of interest.

We included those studies concerned with secondary-care assessment that described the properties of IQCODE for diagnosis at a single time point in a population robustly and independently assessed for presence of dementia. This implies that the index and reference are performed contemporaneously.

An alternative approach is to perform the index test and then prospectively follow people for development of the condition of interest defined using a reference standard. This 'delayed verification' of dementia methodology is best suited to studies describing progression of mild cognitive impairment (MCI) to dementia and was not considered in this review.

Case-control studies are known to potentially overestimate properties of a test and we did not include such studies. Similarly we excluded case studies and samples with small numbers (for the purposes of this review, we defined 'small numbers' as fewer than 10 participants). Small samples were excluded due to the potential for bias in selection and lack of representativeness.

Where settings were mixed, for example, a population study 'enriched' with additional non-secondary-care cases, we did not consider such studies unless separate data were presented for participants from each setting. This design can suffer from similar biases to a case-control design.

Participants

All adults (aged over 18 years) presenting to secondary care were eligible.

Our definition of a secondary-care-based study setting was one where participants were referred to a hospital or outpatient specialist service, either due to perceived memory problems or due to another medical complaint; they may have had previous cognitive testing. There were no predefined exclusion criteria relating to the case mix of the population studied, but this aspect of the study was considered as part of our assessment of heterogeneity. Where there were concerns that the participants were not representative of a secondary-care sample we explored this at study level using our 'Risk of bias' assessment framework. Where studies focused on a specific population, for example, stroke survivors, we described these separately. Recognising that people referred to hospital for specific memory assessment may differ from those referred to hospital for other complaints, we presented these two settings separately.

Index tests

Studies had to include (not necessarily exclusively) IQCODE used as an informant questionnaire.

IQCODE has been translated into various languages to allow international administration (Isella 2002). The properties of a translated IQCODE in a cohort of non-English speakers may differ from properties of the original English-language questionnaire. We collected data on the principal language used for IQCODE assessment in studies to allow for assessment of heterogeneity in relation to language.

Since its original description, modifications to the administration of IQCODE have been described (Jorm 2004). Shorter forms of informant questionnaires that test fewer domains are available and properties may differ from the original 26-item IQCODE tool. We included all such versions of IQCODE, but present separate analysis limited to the commonest 26- and 16-item versions. A modified IQCODE for self assessment has been described (Cullen 2007). As our interest was informant interviews, we have not included self-assessment IQCODE in the review.

Target conditions

Papers reporting any clinical diagnosis of all-cause (unspecified) dementia were potentially eligible for inclusion. Defining a particular dementia subtype was not required, although where available these data were recorded.

Reference standards

Our reference standard was clinical diagnosis of dementia. We recognise that clinical diagnosis itself has a degree of variability but this is not unique to dementia studies and does not invalidate the basic diagnostic test accuracy approach. Clinical diagnosis included all-cause (unspecified) dementia, using any recognised diagnostic criteria (for example, International Classification of Diseases Edition 10 (ICD-10); Diagnostic and Statistical Manual of Mental Disorders Edition 4 (DSM-IV)). Dementia diagnosis may specify a pathological subtype and all dementia subtypes were included. Clinicians may use imaging, pathology, or other data to aid diagnosis; however, we did not include diagnosis based only on these data without corresponding clinical assessment. We recognise that different iterations of diagnostic criteria may not be directly comparable and that diagnosis may vary with the degree or manner in which the criteria have been operationalised (e.g. individual clinician versus algorithm versus consensus determination). We set no criteria relating to severity or stage of dementia diagnosis; instead we classified any clinical diagnosis of dementia (not mild cognitive impairment or its equivalents). We planned to explore stage/severity of dementia as a potential source of heterogeneity.

Search methods for identification of studies

We used a variety of information sources to ensure that we included all relevant studies. We devised terms for electronic database

searching in conjunction with the Trials Search Co-ordinator at the Cochrane Dementia and Cognitive Improvement Group. As part of a body of work looking at cognitive assessment tools, we created a sensitive search strategy designed to capture dementia test accuracy papers. We then assessed the output of the searches to select those papers that could be pertinent to IQCODE, with further selection for directly relevant papers and those papers with a secondary-care focus.

Electronic searches

We searched ALOIS, the specialised register of the Cochrane Dementia and Cognitive Improvement Group (which includes both intervention and diagnostic accuracy studies), MEDLINE (Ovid SP), EMBASE (Ovid SP), PsycINFO (Ovid SP), BIOSIS Previews (Thomson Reuters Web of Science), Web of Science Core Collection (includes Conference Proceedings Citation Index) (Thomson Reuters Web of Science), CINAHL (EBSCOhost) and LILACS (Bireme). See [Appendix 5](#) and [Appendix 6](#) for the search strategies run. The final search date was 28 January 2013.

We also searched sources specific to diagnostic accuracy and health-care research assessment:

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

We applied no language or date restrictions to the electronic searches, and used translation services as necessary.

A single researcher (ANS), with extensive experience of systematic reviews from the Cochrane Dementia and Cognitive Impairment Group, performed the initial screening of the search results. All subsequent searches of titles/abstracts/papers were performed by independent paired assessors (TJQ, PF).

Searching other resources

Grey literature: We identified 'grey' literature through searching of conference proceedings, theses or PhD abstracts in EMBASE, the Web of Science Core Collection and other databases already specified.

Handsearching: We did not perform handsearching. The evidence for the benefits of handsearching are not well defined, and we note that a study specific to diagnostic accuracy studies suggested little additional benefit of handsearching above a robust initial search strategy ([Glanville 2010](#)).

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

Correspondence: We contacted research groups who have published or are conducting work on IQCODE for dementia diagnosis, informed by the results of the initial search.

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

Data collection and analysis

Selection of studies

One review author (ANS) screened all titles generated by initial electronic database searches for relevance. The initial search was a sensitive, generic search, designed to include all potential dementia screening tools. Two review authors (ANS, TJQ) selected titles potentially relevant to IQCODE. Two review authors (TJQ, PF) independently conducted all further review and selection. We reviewed potential IQCODE-related titles, assessing all eligible studies as abstracts, and potentially relevant studies as full manuscripts against the inclusion criteria. We resolved disagreement by discussion, with the potential to involve a third review author (DJS) as arbiter if necessary.

We adopted a hierarchical approach to exclusion, first excluding on the basis of index test and reference standard, and then on the basis of sample size and study data. Finally we assessed all IQCODE papers with regard to setting.

Where a study may have included useable data but these were not presented in the published manuscript, or the data presented could not be extracted to a standard two-by-two table, we contacted the authors directly to request further information or source data. If authors did not respond or if the data were not available we did not include the study (labelled as 'data not suitable for analysis' on the study flowchart). If the same dataset was presented in more than one paper we included the primary paper.

We detailed the study selection process in a PRISMA flow diagram.

Data extraction and management

We extracted data to a study-specific pro forma that included clinical/demographic details of the participants (including details of reason for hospital referral - 'memory' or 'non-memory'), details of IQCODE administration, and details of the dementia diagnosis process. We extracted data for all IQCODE studies, before dividing them by setting (community, primary or secondary). We piloted the pro forma against two of the included papers before use.

Where IQCODE data were given for a number of cut-off points, we extracted data for each IQCODE threshold. Where thresholds were described to two decimal places, we chose the cutpoint closest to the point of interest (i.e. all scores less than 3.35 would be scored as 3.3, all scores 3.35 or greater would be scored as 3.4). We extracted data to a standard two-by-two table.

Two review authors (TJQ, PF) extracted data independently. The review authors were based in different centres and were blinded to each other's data until extraction was complete. We then compared and discussed data pro forma with reference to the original papers, resolving disagreements in data extraction by discussion, with the potential to involve a third review author (DJS) as arbiter if necessary.

For each included paper, we detailed the flow of participants (numbers recruited, included, assessed) in a flow diagram.

Assessment of methodological quality

As well as describing test accuracy, an important goal of the DTA (diagnostic test accuracy) process is to improve study design and reporting in dementia diagnostic studies. For this reason, we assessed both methodological and reporting quality.

We assessed the quality of study reporting using the Standards for the Reporting of Diagnostic Accuracy studies (STARD) checklist (Bossuyt 2003) (Appendix 7). We followed the guidance and principles outlined in the dementia-specific STARDdem extension to STARD grading. We present our results under the descriptor STARD, as at time of writing STARDdem guidance is not yet published and in the public domain. We advocate use of STARDdem (Noel-Storr 2014) for the assessment of diagnostic accuracy studies in dementia henceforth.

We assessed the methodological quality of each study using the Quality Assessment tool for Diagnostic Accuracy Studies (QUADAS-2) tool (www.bris.ac.uk/quadas/quadas-2). This tool incorporates domains specific to patient selection; index test; reference standard; and participant flow. Each domain is assessed for risk of bias and the first three domains are also assessed for applicability. Operational definitions describing the use of QUADAS-2 are detailed in Appendix 8. To create QUADAS-2 anchoring statements specific to studies of dementia test accuracy, we convened a multidisciplinary review of various test accuracy studies with a dementia reference standard (Davis 2013) (Appendix 9). Paired, independent raters (TJQ and PF or TJQ and JKH), blinded to each other's scores, performed both assessments. We resolved disagreements by further review and discussion, with the potential to involve a third review author (DJS) as arbiter if necessary.

We did not use QUADAS-2 data to form a summary quality score, but rather we chose to present a narrative summary describing studies that found high/low/unclear risk of bias/concerns regarding applicability with corresponding graphical displays.

Statistical analysis and data synthesis

We were principally interested in the test accuracy of IQCODE for the dichotomous variable 'dementia/no dementia'. Thus, we applied the current DTA framework for analysis of a single test and fitted the data extracted to a standard two-by-two data table showing binary test results cross-classified with a binary reference standard. We repeated this process for each IQCODE threshold score described.

We used Review Manager 5 (RevMan 2014) to calculate sensitivity, specificity and their 95% confidence intervals (CIs) from the two-by-two tables abstracted from the included studies. We present these data graphically in forest plots to allow basic visual inspection of individual studies only. Standard forest plots with graphical representation of summary estimates are not suited to quantitative synthesis of DTA data. Using software additional to Review Manager 5 (SAS release 9.1) we used the bivariate method to calculate summary values within each prespecified cut-off. The bivariate methods (Reitsma 2005) enabled us to calculate summary estimates of sensitivity and specificity while correctly dealing with the different sources of variation: (1) imprecision, by which sensitivity and specificity have been measured within each study; (2) variation beyond chance in sensitivity and specificity between studies; (3) any correlation that might exist between sensitivity and specificity. We describe the results for each chosen threshold as sensitivity and specificity and we estimate all accuracy measures with their 95% CI. Where data allowed, we chose to present individual study results graphically by plotting estimates of sensitivities and specificities in the receiver operating characteristic (ROC) space. We present the summary sensitivity and specificity points with a 95% confidence region. We have not fitted a ROC curve as we chose the bivariate model for the analysis rather than the hierarchical summary receiver-operator curve (HSROC) method. We also describe metrics of pooled positive and negative likelihood ratios. To allow an overview of IQCODE test accuracy, we performed a further analysis: pooling data at a common threshold (3.3 or closest), chosen to maximise the data available for inclusion.

Investigations of heterogeneity

Heterogeneity is to be expected in diagnostic test accuracy reviews and we did not perform formal analysis to quantify it.

The properties of a tool describe behaviour of the instrument under particular circumstances. Thus, for our assessment of potential sources of heterogeneity (where data allowed) we collected data to inform our prespecified areas of interest:

- a) clinical criteria used to reach dementia diagnosis (for example, ICD-10; DSM-IV) and the methodology used to reach dementia diagnosis (for example, individual assessment; group (consensus) assessment);
- b) technical features of the testing strategy. (version of IQCODE (language); number of items, for example traditional IQCODE;

16-item 'short' form etc).

c) reason for secondary-care consultation. We dichotomised this as attending for 'memory problem' or attending for 'other medical problem'.

Where data allowed we performed pooled analysis with these factors as covariates, and compared results of subgroups. We prespecified that we would present data from the specialised memory setting (memory) and general secondary-care setting (non-memory) separately, that we would present data from the traditional (26 questions) and short-form (16 questions) IQCODE separately, and that we would present data from studies using English language IQCODE against those using non-English-language versions.

Sensitivity analyses

Where appropriate (i.e. if not already explored in our analyses of heterogeneity) and as data allowed, we planned to explore the sensitivity of any summary accuracy estimates to aspects of study quality, guided by the anchoring statements developed in our QUADAS-2 exercise. We prespecified sensitivity analysis planned

to exclude studies of low quality (high likelihood of bias) to determine if the results are influenced by inclusion of the lower-quality studies; and sensitivity analysis excluding studies that may have unrepresentative populations.

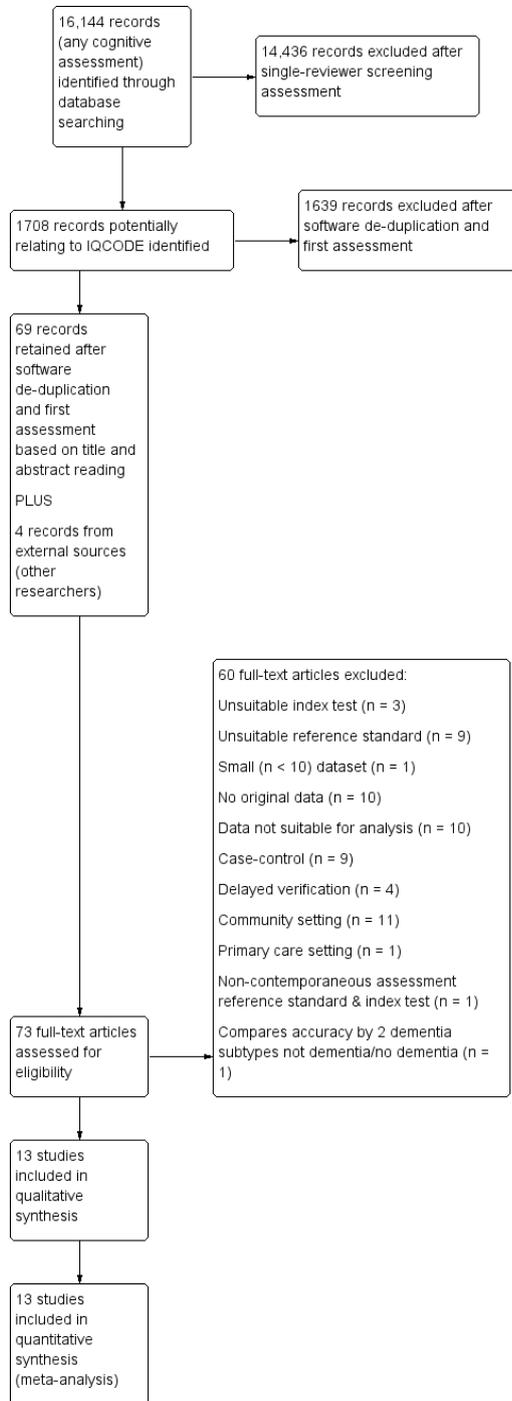
RESULTS

Results of the search

Our search resulted in 16,144 citations, from which we identified 73 full-text papers for eligibility.

We excluded 60 papers (Figure 1). Reasons for exclusion were: population not from a secondary-care setting; no IQCODE data or unsuitable IQCODE data; small numbers (< 10) of included participants; no clinical diagnosis of dementia; repeat datasets; data not suitable for analysis (described in more detail in [Selection of studies](#)) and case-control design (see [Characteristics of excluded studies](#)).

Figure 1. Study flow diagram.



Eight studies which we identified required translation. We contacted 16 authors to provide useable data, of whom 13 responded ([Acknowledgements](#)).

This review includes 13 studies, n = 2882 participants ([Summary of findings 1](#)).

Methodological quality of included studies

We described the risk of bias using the QUADAS-2 methodology ([Appendix 8](#)); our anchoring statements for the IQCODE are summarised in [Appendix 9](#). We did not rate any study at low risk of bias for all the categories of QUADAS-2 ([Figure 2](#)). Areas of particular concern for bias were around: participant sampling pro-

cedures (five papers graded low risk, with high rates of unclear or inappropriate sampling frames and inappropriate exclusions) and application of index test (two papers graded low risk of bias, with most papers failing to prespecify their cut-off for test positivity). There were also concerns around applicability, particularly concerning patient selection and index test. Only six papers recruited a sample of representative secondary-care attenders, either to a memory or a non-memory setting, and only five studies provided sufficient detail for their procedure for conducting the IQCODE for it to be considered consistent with the original methodology for use in clinical practice.

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

	<u>Risk of Bias</u>				<u>Applicability Concerns</u>		
	Patient Selection	Index Test	Reference Standard	Flow and Timing	Patient Selection	Index Test	Reference Standard
Flicker 1997	-	-	-	?	+	?	+
Garcia 2002	-	+	+	+	+	+	+
Goncalves 2011	+	+	-	+	+	+	+
Hancock 2009	+	-	+	+	+	?	+
Harwood 1997	+	-	?	+	+	-	+
Jorm 1991	-	-	+	-	-	+	+
Knaefelc 2003	+	-	+	-	+	+	?
Mackinnon 1998	-	-	+	?	?	?	+
Mulligan 1996	-	-	?	?	?	?	?
Narasimhalu 2008	-	-	-	-	-	?	?
Sikkes 2010	-	-	+	+	-	+	+
Siri 2006	-	-	+	-	?	-	?
Tang 2003	+	-	+	-	-	-	+

- High
 ? Unclear
 + Low

We described reporting quality using the STARD guidance (Appendix 7). There were limitations in reporting across all papers (Appendix 10). No paper included all the details recommended in the STARD statement; particular areas of study reporting that could be improved were: distribution of severity of disease (three papers reported on severity of dementia); handling of missing results (four papers explained, for example, how incomplete IQCODE questionnaires were scored) and estimates of variability of diagnostic accuracy (three papers considered variability between assessors or subgroups of participants).

Findings

We have described the individual included studies in [Characteristics of included studies](#) and Additional [Table 1](#). We have also presented tabulated data for test accuracy by IQCODE threshold ([Summary of findings 2](#)) and by covariate ([Summary of findings 3](#)).

The total number of participants across the studies was 2882 (range: 69 to 576), of whom 1413 (49%) had a clinical dementia diagnosis. We performed quantitative synthesis for 2745 participants, of whom 1413 (51%) had a clinical dementia diagnosis. This excludes 59 participants with mild cognitive impairment included by [Sikkes 2010](#), who was assessing the ability of IQCODE to diagnose mild cognitive impairment and presents data depending on diagnostic group. It also excludes 78 participants from the study by [Flicker 1997](#) who were not assessed in the specialist memory clinic setting, as this paper presented differing test accuracy data with respect to assessment location.

The included studies are international, including datasets from

eight countries (Australia, China, Singapore, Spain, Switzerland, Thailand, The Netherlands and the UK).

Nine different versions of IQCODE were used in the included studies and 10 different diagnostic thresholds (3.3, 3.4, 3.5, 3.6, 3.7, 3.8, 3.9, 4.0, 4.1, 4.2) were used to define a positive IQCODE. We limited our analysis to the validated forms of IQCODE that are in common clinical use, i.e. the 26- and 16-item questionnaires. Although [Siri 2006](#) only reported data relating to their 32-item modified IQCODE at an optimal cut-off of 3.4, the authors supplied data for use of the 26-item IQCODE to facilitate inclusion in the quantitative synthesis.

Within the prespecified thresholds chosen for analysis there was a spread of sensitivity and specificity (sensitivity range: 71% to 100%; specificity range: 39% to 88%). Additional [Table 1](#) provides a summary of test accuracy for each study, using the value closest to 3.3.

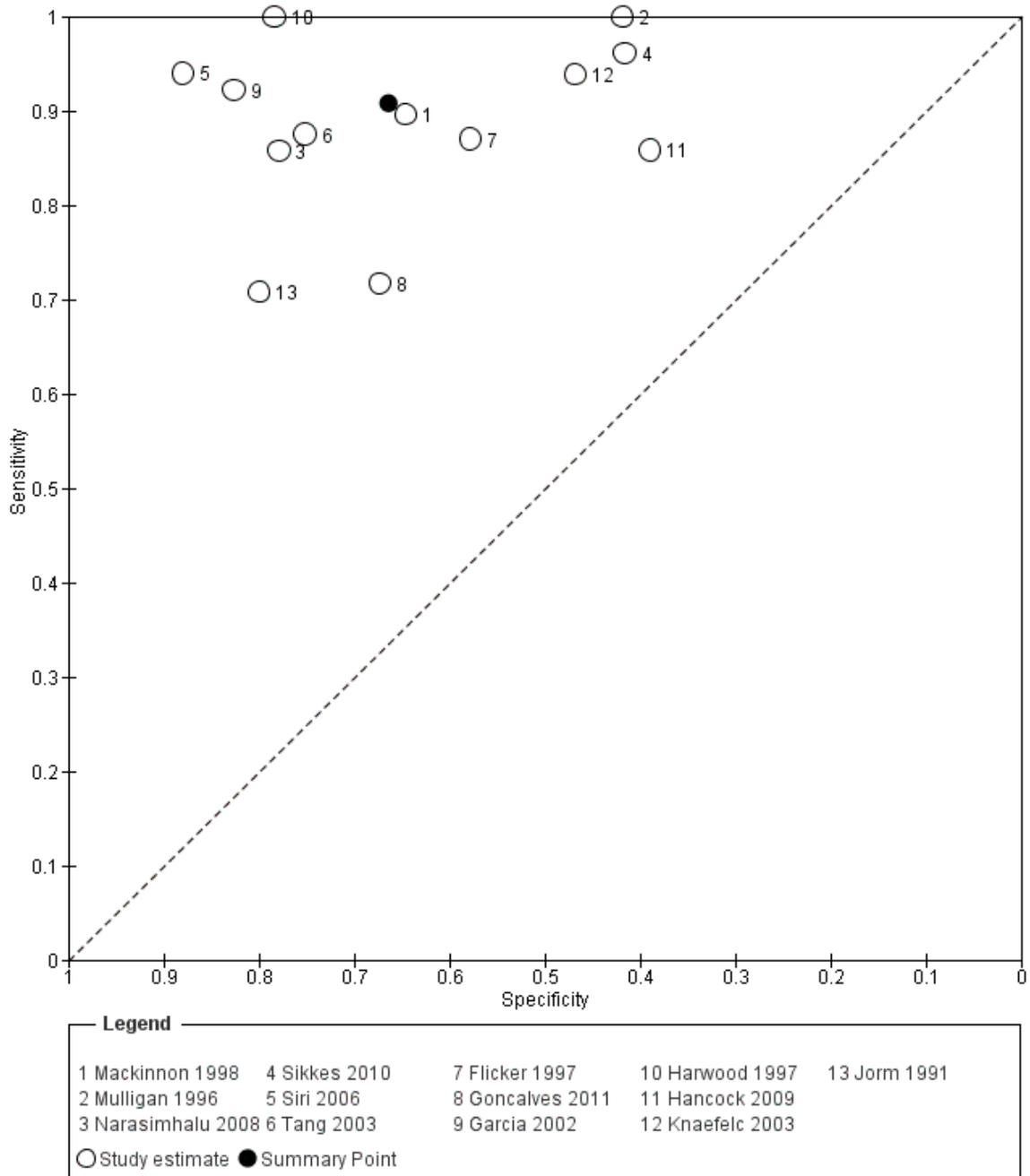
IQCODE (combined 16- and 26-item questionnaire)

Overview analysis - IQCODE using a 3.3 threshold or closest

From the 13 studies, 2745 participants are included in quantitative synthesis. Sensitivity was 0.91 (95% confidence interval (CI) 0.86 to 0.94); specificity 0.66 (95% CI 0.56 to 0.75). The overall positive likelihood ratio was 2.7 (95% CI 2.0 to 3.6) and the negative likelihood ratio was 0.14 (95% CI 0.09 to 0.22).

The summary receiver operating characteristic (ROC) curve describing test accuracy across the included studies is presented in [Figure 3](#).

Figure 3. Summary ROC Plot, IQCODE using a 3.3 threshold score or nearest. The dark point is a summary point, the broken line represents 95% confidence region



IQCODE 3.3 threshold or closest - comparing 26- and 16-item IQCODE

We used the overview dataset to examine the effect of heterogeneity relating to IQCODE format (traditional 26-item or short-form 16-item).

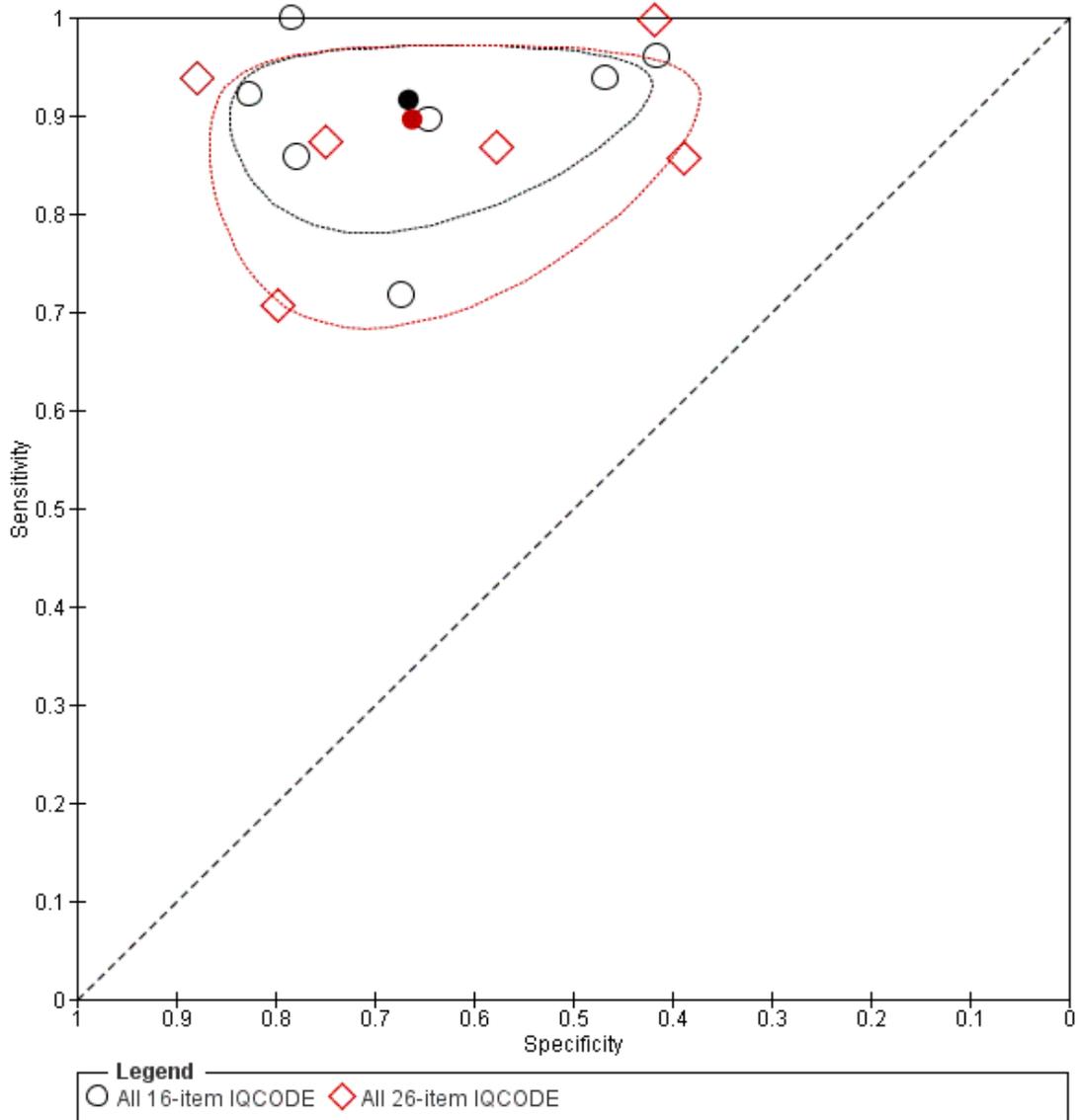
Analysis of the studies using the 26-item IQCODE (six datasets) gave sensitivity of 0.89 (95% CI 0.82 to 0.94); specificity 0.66 (95% CI 0.49 to 0.80). The overall positive likelihood ratio was 2.6 (95% CI 1.6 to 4.3) and the negative likelihood ratio was 0.2

(95% CI 0.1 to 0.3).

Analysis of the studies using the 16-item IQCODE (seven datasets) gave sensitivity of 0.92 (95% CI 0.85 to 0.96); specificity 0.66 (95% CI 0.54 to 0.77). The overall positive likelihood ratio was 2.7 (95% CI 1.9 to 3.8) and the negative likelihood ratio was 0.1. (95% CI 0.1 to 0.2).

Comparing the two, there were no differences in accuracy with a relative sensitivity of the 26-item versus 16-item of 0.98 (95% CI 0.89 to 1.07) and relative specificity of 0.99 (95% CI 0.75 to 1.33). (Figure 4)

Figure 4. Summary ROC Plot of IQCODE 3.3 threshold or nearest, comparing short form (16 item) and traditional IQCODE. The dark point is a summary point, the broken line represents 95% confidence region



As there was no difference we presented further data as the combined (26- and 16-item IQCODE together) test accuracy.

IQCODE 3.3 threshold or closest - comparing English and non-English language IQCODE

We coded the language of IQCODE administration as a covariate. Study numbers did not allow analysis by individual languages and so we compared the IQCODE in the original wording (English language) with all translated IQCODE forms (non-English language).

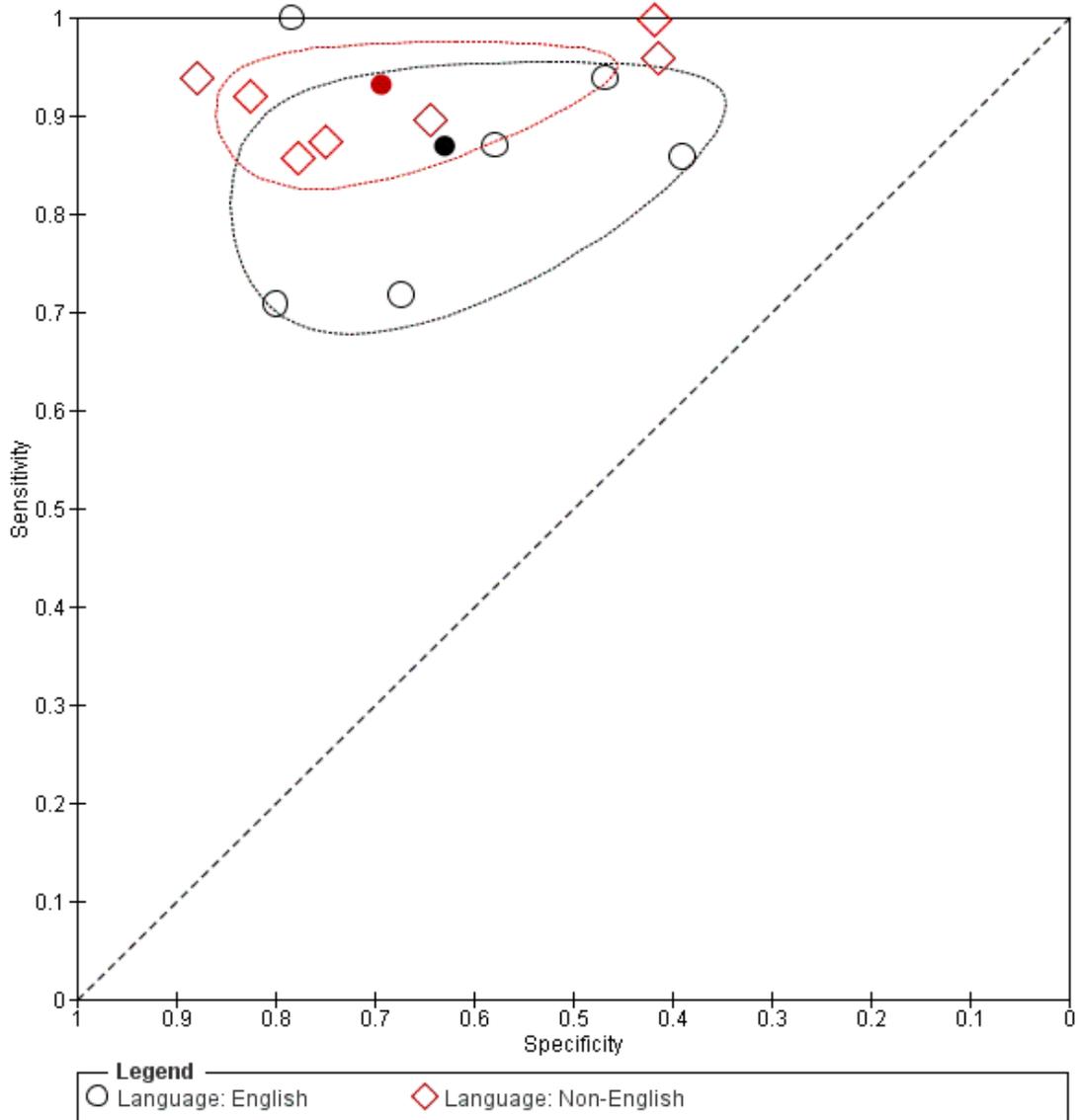
Analysis of studies using English language IQCODE (six datasets) gave sensitivity of 0.87 (95% CI 0.78 to 0.92); specificity 0.63

(95% CI 0.48 to 0.76). The overall positive likelihood ratio was 2.3 (95% CI 1.6 to 3.4) and the negative likelihood ratio was 0.2 (95% CI 0.1 to 0.3).

Analysis of studies using non-English language IQCODE (seven datasets) gave sensitivity of 0.93 (95% CI 0.88 to 0.96); specificity 0.69 (95% CI 0.56 to 0.80). The overall positive likelihood ratio was 3.0 (95% CI 2.1 to 4.5) and the negative likelihood ratio was 0.1 (95% CI 0.1 to 0.2).

Comparing the two, there were no differences in accuracy with a relative sensitivity of the non-English Language versus English language of 1.07 (95% CI 0.98 to 1.17) and relative specificity of 1.10 (95% CI 0.83 to 1.47). (Figure 5)

Figure 5. Summary ROC Plot of pooled IQCODE data at a 3.3 threshold (or nearest value), with language as covariate. The dark point is a summary point, the broken line represents 95% confidence region



As there was no difference we presented further data as the combined (English language and non-English language IQCODE together) test accuracy.

IQCODE test accuracy at differing diagnostic thresholds

We calculated test accuracy at our prespecified IQCODE thresholds. We chose to present a summary ROC curve for those analyses with more than three included studies:

IQCODE 3.3 threshold: there were four datasets* (n = 722) that contained relevant data. The sensitivity was 0.96 (95% CI 0.94 to 0.98); specificity 0.66 (95% CI 0.41 to 0.84). The overall positive likelihood ratio was 2.8 (95% CI 1.5 to 5.5) and the negative likelihood ratio was 0.1 (95% CI 0.03 to 0.1).

IQCODE 3.4 threshold: there were four datasets* (n = 1211) that contained relevant data. The sensitivity was 0.94 (95% CI 0.84 to 0.98); specificity 0.73 (95% CI 0.59 to 0.85). The overall positive likelihood ratio was 3.5 (95% CI 2.1 to 5.8) and the negative likelihood ratio was 0.1 (95% CI 0.03 to 0.2).

IQCODE 3.5 threshold: there was only one dataset (n = 269) that contained relevant data. The sensitivity was 0.92 and specificity was 0.63; we did not perform quantitative synthesis.

IQCODE 3.6 threshold: there were nine datasets* (n = 1576) that contained relevant data. The sensitivity was 0.89 (95% CI 0.85 to 0.92); specificity 0.68 (95% CI 0.56 to 0.79). The overall

positive likelihood ratio was 2.8 (95% CI 1.9 to 4.0) and the negative likelihood ratio was 0.2 (95% CI 0.1 to 0.2).

**Certain papers included more than one dataset*

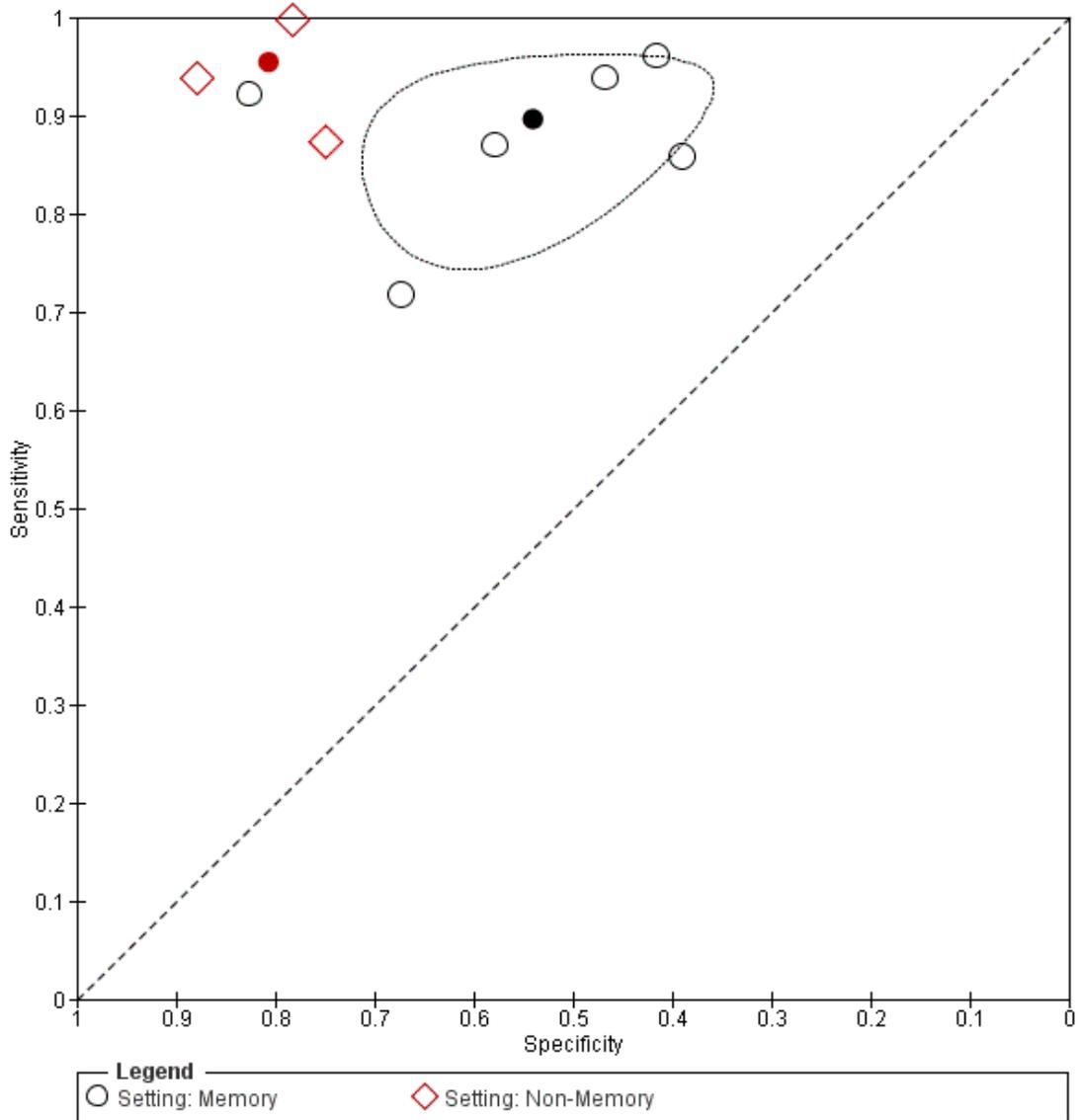
Heterogeneity relating to setting

Specialist memory setting: there were six datasets (n = 1352) that contained relevant data. The sensitivity was 0.90 (95% CI 0.83 to 0.94); specificity 0.54 (95% CI 0.44 to 0.64). The overall positive likelihood ratio was 1.9 (95% CI 1.6 to 2.4) and the negative likelihood ratio was 0.2 (95% CI 0.1 to 0.3). The dementia prevalence ranged from 55% to 87%.

Non-memory setting: there were three datasets (n = 566) that contained relevant data. The sensitivity was 0.95 (95% CI 0.88 to 0.98); specificity 0.81 (95% CI 0.71 to 0.88). The overall positive likelihood ratio was 4.9 (95% CI 3.3 to 7.4) and the negative likelihood ratio was 0.06 (95% CI 0.02 to 0.2). The dementia prevalence ranged from 11% to 50%.

Comparing the two, there is a significant difference in accuracy in the non-memory versus memory settings (P = 0.019) which is attributable to the higher specificity of the IQCODE in the non-memory setting. The relative sensitivity of non-memory versus memory setting is 1.06 (95% CI 0.99 to 1.15) and the relative specificity is 1.49 (95% CI 1.22 to 1.83). (Figure 6)

Figure 6. Summary ROC Plot of pooled IQCODE data at a 3.3 threshold (or nearest value), with setting as covariate. The dark point is a summary point, the broken line represents 95% confidence region



In four studies (Jorm 1991; Mackinnon 1998; Mulligan 1996; Narasimhalu 2008) participants were recruited both in specialist memory and in non-memory secondary-care settings and data were not available stratified by setting, so we could not include them in the quantitative synthesis.

Other sources of heterogeneity and sensitivity analyses

Our objective was to assess the diagnostic accuracy of IQCODE across the cut-off points commonly used in practice (3.3, 3.4, 3.5, 3.6). However, one study (Goncalves 2011) only reported data at an IQCODE cut-off of 4.1. We conducted a sensitivity analysis removing this study, which demonstrated a similar test accuracy (sensitivity was 0.92, 95% CI 0.88 to 0.94; specificity 0.66, 95% CI 0.55 to 0.76).

We performed a sensitivity analysis removing those studies which included participants with a low mean or median age (< 70 years) (Hancock 2009; Narasimhalu 2008; Sikkes 2010). Test accuracy was similar after exclusion of these studies, with an improvement in the specificity of IQCODE (sensitivity was 0.92, 95% CI 0.84 to 0.95; specificity 0.70, 95% CI 0.60 to 0.78 at a threshold of 3.3 or closest).

A quantitative analysis of the effect of dementia diagnosis criteria (reference standard) was not possible. Twelve studies used the American Psychiatric Association Diagnostic and Statistical Manual of Mental Disorders (DSM), one used the World Health Organization International Statistical Classification of Diseases and Related Health Problems (ICD) for diagnosis, and one used the National Institute of Neurological and Communicative Disorders and Stroke and Alzheimer's Disease and Related Disorders Associ-

ation (NINCDS-ADRDA) criteria. Only one study (Jorm 1991) used two diagnostic criteria. Jorm 1991 reported that using DSM-III-R at an IQCODE cut-off of 3.6, they obtained sensitivity of 69% and specificity of 80%, compared with using ICD-10 criteria which resulted in a sensitivity of 80% and specificity of 82%. As the DSM criteria were those most commonly used in other studies, we included these data for reporting of Jorm 1991.

A further original aim was to describe the accuracy of the IQCODE for diagnosis of Alzheimer's disease dementia. Although three studies reported assessing the IQCODE specifically in people with Alzheimer's disease (Goncalves 2011; Narasimhalu 2008; Sikkes 2010), suitable data were only available for two of the three studies (Narasimhalu 2008; Sikkes 2010) and thus we felt that quantitative synthesis would be inappropriate.

Two studies specifically assessed the IQCODE in a stroke population (Narasimhalu 2008; Tang 2003). Tang 2003 only recruited people who had experienced a stroke, while in Narasimhalu 2008 they were a subgroup of the total study population. However, data are not presented on IQCODE properties specific to the stroke population (Narasimhalu 2008) and the small number of studies would make quantitative synthesis inappropriate.

We considered an investigation of heterogeneity assessing the impact of a prespecified IQCODE threshold on test accuracy, from our QUADAS-2 assessment. However, only two studies (Garcia 2002; Goncalves 2011) were eligible for inclusion, and so quantitative synthesis was not appropriate.

Summary of findings

Study ID	Country	Subjects (n)	Mean Age (yrs)	IQCODE version	Language	Dementia diagnosis	Dementia prevalence N (%)	Other assessments
Flicker 1997	Australia	377 (299 from MC)	73.4 MC 79.7 ACAT	26 item	English	DSM-III-R	n = 248 (65.8)	AMT; MMSE
Garcia 2002	Spain	113	78	16 item	Spanish	DSM-III-R	n = 90 (87.4)	MMSE
Goncalves 2011	Australia	204	76.9	16 item	English	DSM-IV-TR	n = 152 (74.5)	RUDAS; SMMSE
Hancock 2009	UK	144	67 (median)	26 item	English	DSM-IV	n = 85 (59.0)	ACE-R; MMSE
Harwood 1997	UK	177	76	16 item	English	DSM-III-R	n = 21 (10.5)	AMT
Jorm 1991	Australia	69	80	26 item	English	DSM-III-R; ICD-10	n = 24 (34.8)	MMSE
Knaefelc 2003	Australia	323	74.7	16 item	English	DSM-IV	n = 229 (70.9)	CAMDEX
Mackinnon 1998	Switzerland	106	80.3	16 item	French	DSM-IV	n = 58 (54.7)	MMSE
Mulligan 1996	Switzerland	76	81.8	26 item	French	DSM-III-R	n = 33 (43.4)	AEMT; MMSE (French)
Narasimhalu 2008	Singapore	576	65 - 70 (mean by diagnosis)	16 item	Cantonese	DSM-IV	n = 169 (29.3)	MMSE (Singapore)
Sikkes 2010	The Netherlands	328 (59 known MCI)	68.4	16 item	Dutch	NINCDS-ADRDA	n = 180 (54.9)	MMSE
Siri 2006	Thailand	200	72.9	32 item & 26 item	Thai	DSM-IV	n = 100 (50.0)	BOMC; CDT; MIS; MMSE
Tang 2003	China	189	74.2	26 item	Chinese	DSM-IV	n = 24 (12.7)	CDR; MMSE

See [Characteristics of included studies](#) for more detailed study descriptors

Abbreviations: ACAT - Aged Care Assessment Team Group; ACE-R - Addenbrooke's Cognitive Examination-Revised; AEMT - Antisaccadic Eye Movement Test; AMT - Abbreviated Mental Test; BOMC - Blessed Orientation Memory; CAMDEX - Cambridge Mental Disorders of the Elderly Examination; CDR - Clinical Dementia Rating Scale; CDT - Clock Drawing Test; DSM - American Psychiatric Association Diagnostic and Statistical Manual of Mental Disorders; MC - Memory Clinic Group; MCI - Mild Cognitive Impairment; MIS - Memory Impairment Screen; MMSE - Mini Mental State Examination; NINCDS-ADRDA - National Institute of Neurological and Communicative Disorders and Stroke and Alzheimer's Disease and Related Disorders Association; NINDS-AIREN - National Institute of Neurological Disorders and Stroke and the Association Internationale pour la Recherche et l'Enseignement en Neurosciences; RUDAS - Rowland Universal Dementia Assessment Scale; SMMSE - Standardized Mini Mental State Examination

What is the accuracy of the Informant Questionnaire for Cognitive Decline in the Elderly (IQCODE) test for detection of dementia when differing thresholds are used to define IQCODE positive cases				
Population	Adults attending secondary-care services, with no restrictions on the case mix of recruited participants			
Setting	Our primary setting of interest was secondary care; within this rubric we included inpatient wards and hospital outpatient clinics			
Index test	Informant Questionnaire for Cognitive Decline in the Elderly (IQCODE) administered to a relevant informant. We restricted analyses to the traditional 26-item IQCODE and the commonly-used short form IQCODE with 16 items			
Reference Standard	Clinical diagnosis of dementia made using any recognised classification system			
Studies	We included cross-sectional studies but not case-control studies			
Test	Summary accuracy (95% CI)	No. of participants (studies)	Dementia prevalence	Implications, Quality and Comments
IQCODE cut-off 3.3 or nearest	sens: 0.91 (95% CI 0.86 to 0.94); spec: 0.66 (95% CI 0.56 to 0.75) +ve LR: 2.7 (95% CI 2.0 to 3.6) -ve LR: 0.14 (95% CI 0.09 to 0.22)	n = 2745 (13 studies)	n = 1413 (51%)	<p>Within the range of commonly used cut-offs for defining IQCODE positivity, there is no clearly optimal value for use in secondary care settings</p> <p>The sensitivity falls as the diagnostic threshold increases from 3.3-3.6, with a relative increase in the specificity</p> <p>The preferred balance between sensitivity and specificity is debatable</p> <p>Both false positive (person diagnosed with possible dementia and referred for further assessment) and false negative (person with dementia has diagnosis missed and is not referred to specialist services) are associated with potential harms</p> <p>The dementia prevalence was highly varied in included studies (10.5% to 87.4%) reflecting the heterogeneity of included participants within a ‘hospital’ setting</p> <p><i>This heterogeneity and associated “modelling” of real world implications of the test accuracy data presented are described in the next summary of findings table.</i></p>

IQCODE cut-off 3.3	sens: 0.96 (95% CI 0.94 to 0.98) spec: 0.66 (95% CI 0.41 to 0.84) +ve LR: 2.8 (95% CI 1.5 to 5.5) -ve LR: 0.1 (95% CI 0.03 to 0.1)	n = 722 (4 studies)	n = 334 (46%)
IQCODE cut-off 3.4	sens: 0.94 (95% CI 0.84 to 0.98) spec: 0.73 (95% CI 0.59 to 0.85) +ve LR: 3.5 (95% CI 2.1 to 5.8) -ve LR 0.1 (95% CI 0.03 to 0.2)	n = 1211 (4 studies)	n = 394 (33%)
IQCODE cut-off 3.5	sens: 0.92** spec: 0.63**	n = 269 (1 study)	n = 152 (57%)
IQCODE cut-off 3.6	sens: 0.89 (95% CI 0.85 to 0.92) spec: 0.68 (95% CI 0.56 to 0.79) +ve LR: 2.8 (95% CI 1.9 to 4.0) -ve LR: 0.2 (95% CI 0.1 to 0.2)	n = 1576 (9 studies)	n = 968 (61%)

CAUTION: The results in this table should not be interpreted in isolation from the results of the individual included studies contributing to each summary test accuracy measure. These are reported in the main body of the text of the review

** : quantitative synthesis not performed as only one study reported data at cut-off of 3.5

Abbreviations: sens - sensitivity; spec - specificity; +ve LR - positive likelihood ratio; -ve LR - negative likelihood ratio

What is the accuracy of the Informant Questionnaire for Cognitive Decline in the Elderly (IQCODE) test for detection of dementia using different versions of IQCODE and using different languages of administration				
Population	Adults attending secondary care services, with no restrictions on the case-mix of recruited participants			
Setting	<p>Our primary setting of interest was secondary care, within this rubric we included inpatient wards and hospital outpatient clinics</p> <p>Secondary care settings can be considered as two groups:</p> <p>(1) Studies conducted in a specialist memory/psychogeriatrics setting where participants were referred due to cognitive symptoms</p> <p>(2) Non-memory focused hospital services. These included unselected admissions of older adults, those referred to specialist older people's assessment teams, outpatient attenders and inpatients under the care of geriatricians</p>			
Index test	Informant Questionnaire for Cognitive Decline in the Elderly (IQCODE) administered to a relevant informant. We restricted analyses to the traditional 26-item IQCODE and the commonly used short form IQCODE with 16 items			
Reference Standard	Clinical diagnosis of dementia made using any recognised classification system			
Studies	Cross-sectional studies were included, we did not include case-control studies			
Comparative analyses				
Test	No. of participants (studies)	Dementia prevalence across studies	total Findings	Implications
26-item versus 16-item IQCODE	Total: n = 2745 (13) 26 item n = 977 (6)	Total n = 1413 (51%) 26 item n = 514 (53%) 16 item n = 899 (51%)	No significant difference in test accuracy Relative sensitivity of 26-item versus 16-item IQCODE: 0.98 (95% CI 0.89 to 1.07) Relative specificity of 26-item versus 16-item IQCODE: 0.99 (95% CI 0.75 to 1.33)	There was no difference in accuracy between IQCODE versions so it may be justifiable to advocate use of the short form to minimise responses required

English language versus Non-English	Total: n = 2745 (13) English language n = 1216 (6)	Total n = 1413 (51%) English language n = 759 (62%) Non-English language n = 654 (43%)	No significant difference in test accuracy Relative sensitivity of English language versus non-English language IQCODE: 1.07 (95% CI 0.98 to 1.17) Relative specificity of English language versus non-English language IQCODE: 1.10 (95% CI 0.83 to 1.47)	The language of administration does not significantly influence the diagnostic accuracy of IQCODE
Non-memory setting versus memory	Total: n = 1918 (9)* memory setting n = 1352 (6)	Total n = 1129 (59%) memory setting n = 984 (73%) non-memory setting n = 145 (26%)	Significant difference in test accuracy between settings (P = 0.019), due to higher specificity in non-memory settings Relative sensitivity of non-memory versus memory IQCODE: 1.06 (95% CI 0.99 to 1.15) Relative specificity of non-memory versus memory IQCODE: 1.49 (1.22 to 1.83)	The lower level of specificity in the specialist memory services is of limited clinical concern as other tests will be used in this setting and incorrectly diagnosing someone with dementia based on IQCODE alone would be unlikely In the non-memory setting it is likely a positive IQCODE would prompt referral to specialist services, and this may be associated with psychological harm and unnecessary expense Applying our non-memory findings to the UK; there are around 2 million unscheduled admissions annually in over-65s (Imison 2012) and a dementia prevalence of 42.4% in this group (Sampson 2009). Using the IQCODE alone to screen for dementia would result in 42,400 people with dementia not being identified and 218,880 dementia-free people being referred inappropriately for specialist assessment

CAUTION: The results in this table should not be interpreted in isolation from the results of the individual included studies contributing to each summary test accuracy measure. These are reported in the main body of the text of the review

*: Four studies included participants recruited in both specialist memory and non-memory settings, without reporting outcome data stratified by recruitment setting and are thus not included in the quantitative synthesis

DISCUSSION

Summary of main results

We present a review of the available evidence around test accuracy of the Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE) for dementia diagnosis in hospital/secondary-care settings. Our quantitative synthesis demonstrates summary sensitivity of 0.91 and specificity of 0.66 when IQCODE is used across all (undifferentiated) secondary-care settings for the diagnosis of dementia. The positive likelihood ratio was 2.7 and the negative likelihood ratio was 0.14; indicating that the IQCODE can be used as a 'rule-out' test of dementia in a secondary-care setting. These results represent a large dataset, comprising data from 13 international studies with over 2745 participants. We limited our review to studies concerning hospital-based healthcare systems; however even within this focused setting there was substantial heterogeneity and we must be cautious in our interpretation of the pooled data. Across the included papers there was substantial potential for bias and issues with limited generalisability and suboptimal reporting.

The prevalence of dementia in the included settings was highly varied, ranging from 10.5% to 87.4%. This marked difference in patient populations reflects in part the differing case mix that potentially can be included under the 'hospital' setting label. We explored this aspect of heterogeneity with prespecified subgroup and sensitivity analyses and found a significant difference in test accuracy depending on whether the hospital setting described was a specialist memory service (for example, old-age psychiatry ward; memory clinic) or a non-memory-specific hospital setting (for example, an acute admissions ward or general outpatient clinic). The clinical interpretation of such comparisons is challenging. One interpretation of these data is that IQCODE as a diagnostic tool may be more suited to general hospital settings rather than services with a cognitive focus. The pictorial summary analysis (Figure 6) illustrates that memory and non-memory groups seem to behave differently and perhaps should be treated as such in future analyses of cognitive test accuracy. These data come with several caveats (a modest number of included studies; heterogeneity within the memory/non-memory groups; issues with potential for bias) but our interpretation has a clinical validity as the case mix in a specialist service designed for those with suspected dementias is likely to be very different to the population presenting for assessment in an unscheduled acute admissions or medicine for the older adult ward. The difference between the groups was most apparent in specificity, with the data suggesting better specificity when IQCODE is used in non-memory settings. We can speculate on potential reasons for this difference: in the specialist memory service setting, the high prevalence of depression, either solely or co-existing with dementia (Knapskog 2014) may be an important consideration, as many of the IQCODE parameters are task-oriented and thus may be impacted by depressive symptoms giving

false positive IQCODE results. In separating the study settings of memory and non-memory we recognise that differences between these populations operate at many levels including potential availability of an informant to complete IQCODE scoring.

There is no universal value of sensitivity and specificity that is considered 'good' or 'poor'; the values that clinicians will accept as suitable for clinical use will vary with the implications of a false positive or false negative result. In the non-memory (often acute hospital) setting, delirium is prevalent (Ryan 2013), either alone or in association with cognitive impairment or dementia, but opportunities for in-depth patient-dependent cognitive testing may be limited and the more favourable test accuracy metrics of IQCODE in this setting are reassuring. It could be argued that the lower specificity for the instrument in a specialist memory service is less problematic than in other healthcare settings, as patients will receive additional assessments as determined by the specialist clinician and are unlikely to be misdiagnosed on the basis of an IQCODE result alone.

Applying summary test accuracy data to real-world settings can be illustrative of the potential strengths and limitations of a test in practice. Applying our non-memory summary data to the acute hospital admission setting, current UK data estimate around two million unscheduled admissions annually in the over-65s (Imison 2012) and a dementia prevalence of 42.4% in this group (Sampson 2009). Using the IQCODE alone to screen for dementia would result in 42,400 people with dementia not being identified and 218,880 people without dementia being referred inappropriately for specialist assessment. Both false positive and false negative results have potential for harm. It is not certain that those whose dementia is missed with IQCODE will eventually receive a diagnosis and opportunities for early intervention may be lost, while inappropriately labelling a person as having cognitive decline based on IQCODE will also be associated with potential psychological harm and economic implications of need for further investigation. We acknowledge that the UK NHS-based figures we quote may not be applicable to other countries or healthcare systems. We have presented UK data in this review as we have access to reasonably robust input data and their inclusion illustrates important points about the real-world implications of our summary test accuracy data

Our results do not indicate an optimal cut-off for the IQCODE in a hospital setting. A range of diagnostic thresholds were reported with significant overlap between the included studies, with the commonest diagnostic threshold being 3.6; this is higher than the IQCODE cut-offs employed when the tool is used in community settings. Only one study reported data using a cut-off outside our prespecified range of 3.3 to 3.6 (Goncalves 2011; cut-off 4.1) To allow us to use the maximum available data we included these data in our summary analysis, with a sensitivity analysis demonstrating no significant effect of excluding them.

We recognise that IQCODE can be applied using various methods and we prespecified analyses to try and describe the effects on

test accuracy. Our finding of no difference in the diagnostic accuracy between assessments conducted in the English language and those conducted in six other languages (grouped together as 'non-English language' to allow analysis) is reassuring, and supports the cross-cultural use of IQCODE. Similarly, the length of instrument (26-item versus 16-item) had no significant effect on test accuracy, a result in keeping with previous narrative review findings in [Jorm 2004](#) and with previous review of IQCODE properties when used in a community setting ([Quinn 2014](#)).

Since IQCODE was originally designed for use in the older adult population, we felt it was important to ensure that a tool used to aid the diagnosis of dementia would be robust to the difficulties of assessment in younger age groups with potential early-onset dementia ([Vieira 2013](#)). To explore age effects, we performed a sensitivity analysis removing studies with a low average age of included participants, and found test accuracy to be broadly similar.

We prespecified two other analyses based on diagnostic features. We accepted any validated clinical assessment system for our reference standard of dementia diagnosis but recognised that differing classifications operationalise dementia in slightly different ways. One included study used two diagnostic criteria in direct comparison ([Jorm 1991](#)). Using the same cut-off of 3.6, the DSM III-R resulted in a sensitivity of 0.69 and a specificity of 0.80 compared with ICD-10 which produced a sensitivity of 0.80 and specificity of 0.82. As the majority of the other included studies used DSM criteria only, it was not possible to further describe any potential effect of diagnostic criteria on IQCODE accuracy. As a recognition of the different effects that subtypes of dementia have on the individual ([Gure 2010](#)), we felt it was reasonable to analyse the diagnostic properties of IQCODE with respect to specific subtypes of dementia. It had previously been demonstrated that the IQCODE performs differently in people with Alzheimer's disease dementia and those with frontotemporal dementia ([Larner 2010](#)). However, there was a lack of data available on dementia pathology in the included studies in our review and we were unable to offer subgroup analysis by dementia subtype.

Our 'Risk of bias' assessment using the QUADAS-2 tool identified significant potential for bias in the included studies as described below. Given the modest number of included studies, we did not perform subgroup or sensitivity analyses to quantitatively explore these effects for each QUADAS-2 domain.

Strengths and weaknesses of the review

Strengths and weaknesses of included studies

Our QUADAS-2 and STARD assessments suggested potential problems of bias, poor generalisability and suboptimal reporting across the included studies. Areas of particular concern are highlighted in the text of the [Characteristics of included studies](#) table and summarised in Figure 2.

A key aspect of our QUADAS-2 assessments was establishing

whether authors prespecified the cut-off used to define IQCODE positivity. Where authors calculate test accuracy across the range of potential IQCODE thresholds, they are not reflecting clinical practice, and test accuracy may be inflated if only the best-performing cutpoints are reported. Thus, where cutpoints were not prespecified, we classified the paper as being at high risk of bias for the conduct of the index test. Only two of our included studies ([Garcia 2002](#); [Goncalves 2011](#)) were deemed to be at low risk of bias for this domain.

In order that the findings of our analysis are applicable in practice it is essential that the recruited participants to the included studies are representative. Details about sampling procedures, particularly non-consecutive or non-random samples being recruited and studies inappropriately excluding those with relevant co-morbidities, were an area of concern in our 'Risk of bias' assessment. A further concern about using the IQCODE is that it relies on the assessment of an informant, and not all patients have someone who can fulfil this role. Four of the included studies only recruited participants where an informant was present at the consultation ([Garcia 2002](#); [Goncalves 2011](#); [Hancock 2009](#); [Sikkes 2010](#)). The significance of attending the memory clinic unaccompanied has previously been demonstrated to be a specific predictor of the individual not having a clinical diagnosis of dementia ([Larner 2009](#)). Nonetheless, reliance on an informant is a relevant factor when considering test accuracy as a screening tool, as those studies did not recruit participants who attended unaccompanied. One author adopted a broader approach and permitted the completion of the IQCODE by post or telephone ([Harwood 1997](#)).

We wanted to ensure that case-control methodologies were not included in this review, given the propensity to falsely estimate test accuracy, as the prevalence has been artificially fixed. One of the included studies ([Siri 2006](#)) reported exactly equivalent numbers of those with dementia and those without ($n = 100$). The methodology described does not suggest that a case-control design was used, but there is a lack of detail as to how the final sample was obtained.

Reporting quality impacts on the 'Risk of bias' assessment as, where procedures are not fully described, this limits the potential for judging the rigor of the methodology. The STARD assessments revealed a lack of reporting around disease severity and the handling of indeterminate results. Both of these have implications for the use of the tool in clinical practice. None of the included studies reported properties of IQCODE in relation to disease severity or stage. Intuitively test properties will differ comparing subtle, early dementia with later stage advanced disease; the optimal cut-off may also change as the disease progresses.

Strengths and weaknesses of review process

In common with the other reviews in the suite ([Harrison 2014](#); [Quinn 2014](#)), the review benefits from a structured and thorough search strategy created and conducted by an experienced Tri-

als Search Co-ordinator. We adopted an inclusive approach and identified relevant studies in a formal and standardised manner. We recognise that our search was performed in January 2013 and this may have led to the potential exclusion of relevant studies published more recently. Quality assessment was guided by our dementia-specific QUADAS-2 anchoring statements which were devised for use in diagnostic test accuracy studies that compare a cognitive index test and clinical reference standard (Davis 2013). In addition, our quality assessment was complemented by formal assessment of reporting quality using the STARD methodology (Appendix 7), an approach which has been shown to add rigor in test accuracy evaluation (Oliveira 2011). Had it been available at the time of analysis, the dementia-specific STARDdem guidance on reporting may have better described the challenges inherent in reporting research around dementia tests (Noel-Storr 2014). We were inclusive in our initial search of the literature and assessed study reports which were not available in English, making use of translation services to facilitate study selection and data extraction. Although only one paper written in Spanish met the final inclusion criteria, this approach meant studies were not inappropriately excluded due to their language of presentation.

Contacting study authors was highly productive, allowing for clarification of methodology, for example, to ensure case-control designs were not included; updating citations identified as abstracts to allow for the subsequent full-text publications to be cited; and provision of data in a format suitable for inclusion in the quantitative synthesis.

Our review question was focused to facilitate the assessment of the test properties of IQCODE in a secondary-care setting. Where a study included a non-secondary-care setting, we excluded the data from this review but considered them for reviews of IQCODE in other healthcare settings. We excluded studies concerned with the diagnosis of mild cognitive impairment from our quantitative synthesis, as our objective was to assess the diagnostic accuracy of IQCODE for the diagnosis of dementia. We prespecified a series of subgroup and sensitivity analyses to look at hospital settings, IQCODE application and dementia diagnosis. Not all of these analyses were possible due to limited data; we were mindful of not over-analysing what was a modest dataset and did not perform post hoc analyses or analyses relating to QUADAS-2 domains.

Comparisons with previous research

Our findings are in keeping with reviews assessing the test accuracy of the IQCODE in other healthcare settings. In the review describing IQCODE as used in the community setting, summary sensitivity was 0.80 (95% confidence interval (CI) 0.75 to 0.85) and specificity was 0.84 (95% CI 0.78 to 0.90). In the community review, the form of IQCODE (26 versus 16 items) similarly had no effect on accuracy and there was no obvious optimal cut-off for IQCODE across the range 3.3 to 3.6 (Quinn 2014). The third in the suite of IQCODE reviews, assessing accuracy in primary care,

had no quantitative synthesis as we found only one relevant study (Harrison 2014). The IQCODE has been assessed in comparison to other informant or self-completed instruments, although without presenting quantitative synthesis (Cherbuin 2008). Other authors have concluded that a combined approach of tools, often with direct patient assessment and informant assessment, is required in view of the complexity of diagnosis and disease subtypes (Stephan 2010; Cullen 2007).

Applicability of findings to the review question

Our focused review question concerned the accuracy of IQCODE for dementia diagnosis in a secondary-care/hospital setting. We believe our robust search and clear operationalisation of the hospital setting have allowed us to comprehensively collate all available evidence on this question. Although the number of included studies was modest with substantial heterogeneity, we were still able to offer quantitative summary analyses of IQCODE test accuracy.

AUTHORS' CONCLUSIONS

Implications for practice

From its inception, the IQCODE was intended as a screening test designed to detect deterioration in cognition with less potential bias from educational attainment (Jorm 1988). As such, it has never purported to be able to diagnose dementia per se; instead its value was in indicating those who have features suggestive of new-onset cognitive decline who merit further specialist assessment. With this in mind and based on our summary data, the use of IQCODE as an initial assessment of older adults in general hospital settings seems reasonable. Our data offer less support for the utility of IQCODE as a diagnostic tool in hospital settings that have a cognition/dementia focus, albeit it seems unlikely that IQCODE would be used as the sole diagnostic instrument in such settings.

Although the focus of our review was specific to test accuracy, it is worth emphasising that IQCODE has other features that make it attractive for clinical use. IQCODE offers a practical insight into the impacts of cognitive impairment on the individual from the perspective of an informant. Although people may complain of subjective memory loss, many do not notice their cognitive deficits and, crucially, the impacts on their daily life (Derouesné 1999). IQCODE may allow a greater degree of perspective on the functional impairments arising from cognitive decline, and this in turn may help guide therapy and intervention. These advantages could equally apply to other multi-item informant assessment tools, and it is important to recognise that the IQCODE is not the only informant tool for dementia screening. Other such tools are increasingly used, particularly in North America, for example the AD-8

(Galvin 2005), with reviews of these tools planned or underway (Hendry 2014).

The simplest approach to cognitive screening is the use of single-question tools which ask the patient or informant if they have noticed any memory-related problems. In the UK NHS a single-question screening assessment is recommended for all older adults admitted to hospital (Hendry 2015). Informant reports have been described as superior (Carr 2000). This has led to some recommending the use of a single-informant response ahead of structured tools such as the IQCODE as a means of screening, due to the lesser burden of a single question (Ayalon 2011). These approaches, however, have largely been tested for the detection of mild dementia or cognitive impairment, whereas the secondary-care population includes the spectrum of disease severity.

Although our review considers the test accuracy of the IQCODE, we are not able to present data on the acceptability or feasibility of using the tool in this setting. Of particular relevance to secondary care is the means of IQCODE completion and how obtaining an appropriate informant account can be operationalised in busy acute hospital environments. It has been previously noted that the relationship of the informant to the patient is important in ensuring responses are reliable over the 10-year time course (Jorm 2004). Other means of completion include by telephone or by post. Postal completion methods are associated with missing data (Smeeth 2001) and may lead to delays in obtaining data. In small-scale studies with dedicated research assistants IQCODE completion rates have been around 80% (Lees 2013), but numbers may be considerably less where the IQCODE assessment is part of routine care.

Implications for research

In view of the significant difference in diagnostic accuracy using the IQCODE in a specialist memory setting compared to a non-memory setting, it is essential that future diagnostic test accuracy studies present results stratified by the recruitment setting of included participants, and consider these two populations as discrete entities.

Research into the feasibility and acceptability of the instrument is also needed to allow us to define the potential role of IQCODE in the clinical pathway for assessing individuals for dementia. Patients' and carers' experiences of cognitive screening and the impact of test results have not been well described and more research in this field would help inform test strategy.

ACKNOWLEDGEMENTS

We thank the following researchers who assisted with translation: Salvador Fudio, EMM van de Kamp-van de Glind, Anja Hayden.

We thank the following researchers who responded to requests for original data: Dr D Salmon, Dr S Sikkens, Dr G Potter, Dr M Razavi, Prof H Henon, Dr JFM de Jonghe, Dr V Isella, Dr AJ Larner, Dr B Rovner, Dr M Krogseth, Dr V Valcour, Dr K Okanurak, Dr D Goncalves.

We would like to thank Dr Yemisi Takwoingi for providing one-to-one training with two of the review authors (JKH and TJQ) to facilitate data analysis.

REFERENCES

References to studies included in this review

Flicker 1997 *{published data only}*

Flicker L, Logiudice D, Carlin JB, Ames D. The predictive value of dementia screening instruments in clinical populations. *International Journal of Geriatric Psychiatry* 1997;**12**(2):203–9.

Garcia 2002 *{published data only}*

Forcano Garcia M, Perlado Ortiz de Pinedo F. Cognitive deterioration: use of the short version of the Informant Test (IQCODE) in the geriatrics consultations. *Revista Española de Geriatria y Gerontología* 2002;**37**:81–5.

Goncalves 2011 *{published data only}*

Goncalves DC, Arnold E, Appadurai K, Byrne GJ. Case finding in dementia: comparative utility of three brief instruments in the memory clinic setting. *International Psychogeriatrics* 2011;**23**(5):788–96.

Hancock 2009 *{published data only}*

Hancock P, Larner AJ. Diagnostic utility of the Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE) and its combination with the Addenbrooke's Cognitive Examination-Revised (ACE-R) in a memory clinic-based population. *International Psychogeriatrics* 2009;**21**(3):526–30.

Harwood 1997 *{published data only}*

Harwood DMJ, Hope T, Jacoby R. Cognitive impairment in medical inpatients: I screening for dementia - is history better than mental state?. *Age and Ageing* 1997;**26**(1):31–5.

Jorm 1991 *{published data only}*

Jorm AF, Scott R, Cullen JS, MacKinnon AJ. Performance of the Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE) as a screening test for dementia. *Psychological Medicine* 1991;**21**(3):785–90.

Knaefelc 2003 *{published data only}*

Knaefelc R, Lo Giudice D, Harrigan S, Cook R, Flicker L,

- Mackinnon A, et al. The combination of cognitive testing and an informant questionnaire in screening for dementia. *Age and Ageing* 2003;**32**(5):541–7.
- Mackinnon 1998** *{published data only}*
Mackinnon A, Mulligan R. Combining cognitive testing and informant report to increase accuracy in screening for dementia. *American Journal of Psychiatry* 1998;**155**(11):1529–35.
- Mulligan 1996** *{published data only}*
Mulligan R, Mackinnon A, Jorm AF, Giannakopoulos P, Michel JP. A comparison of alternative methods of screening for dementia in clinical settings. *Archives of Neurology* 1996;**53**(6):532–6.
- Narasimhalu 2008** *{published data only}*
Narasimhalu K, Lee J, Auchus AP, Chen CP. Improving detection of dementia in Asian patients with low education combining the Mini-Mental State Examination and the Informant Questionnaire on Cognitive Decline in the Elderly. *Dementia and Geriatric Cognitive Disorders* 2008;**25**(1):17–22.
- Sikkes 2010** *{published data only}*
Sikkes SA, Van den Berg MT, Knol DL, De-Lange-de Klerk ES, Scheltens P, Uitdehaag BM, et al. How useful is IQCODE for discriminating between Alzheimer's disease, mild cognitive impairment and subjective memory complaints?. *Dementia and Geriatric Cognitive Disorders*. 2010;**30**(5):411–6.
- Siri 2006** *{published data only}*
* Siri S, Okanurak K, Chansirikanjana S, Kitiyaporn D, Jorm AF. Modified Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE) as a screening test for dementia for Thai elderly. *Southeast Asian Journal of Tropical Medicine and Public Health* 2006;**37**(3):587–94.
Siri, S. Dementia Screening Test for Thai Elderly. Faculty of Graduate Studies, Mahidol University, Thailand 2007.
- Tang 2003** *{published data only}*
Tang WK, Chan SS, Chiu HF, Wong KS, Kwok TC, Mok V, et al. Can IQCODE detect post stroke dementia? . *International Journal of Geriatric Psychiatry* 2003;**18**(8):706–10.
- References to studies excluded from this review**
- Abreu 2008** *{published data only}*
Abreu ID, Nunes PV, Diniz BS, Forlenza OV. Combining functional scales and cognitive tests in screening for mild cognitive impairment at a university-based memory clinic in Brazil. *Revista Brasileira de Psiquiatria* 2008;**30**(4):346–9.
- Bustamante 2003** *{published data only}*
Bustamante SE, Bottino CM, Lopes MA, Azevedo D, Hototatian SR, Litvoc J, et al. Combined instruments on the evaluation of dementia in the elderly preliminary results. *Arquivos de Neuro-Psiquiatria* 2003;**61**(3A):601–6.
- Butt 2008** *{published data only}*
Butt Z. Sensitivity of the informant questionnaire on cognitive decline: an application of item response theory. *Neuropsychology, Development and Cognition: Section B: Aging, Neuropsychology and Cognition* 2008;**15**(5):642–55.
- Cherbuin 2008** *{published data only}*
Cherbuin N, Anstey KJ, Lipnicki DM. Screening for dementia: a review of self- and informant-assessment instruments. *International Psychogeriatrics* 2008;**20**(3):431–58.
- Cherbuin 2013** *{published data only}*
Cherbuin N, Jorm AF. Chapter 8. The Informant Questionnaire for Cognitive Decline in the Elderly. In: Larner AJ editor(s). *Cognitive Screening Instruments*. London: Springer-Verlag, 2013:166–79.
- De Jonge 1997** *{published data only}*
De Jonghe JF. Differentiating between demented and psychiatric patients with the Dutch version of IQCODE. *International Journal of Geriatric Psychiatry* 1997;**12**(4):462–5.
- Dekkers 2009** *{published data only}*
Dekkers M, Joosten-Weyn Banningh EW, Eling PA. Awareness in patients with mild cognitive impairment (MCI). *Tijdschrift voor Gerontologie en Geriatrie* 2009;**40**(1):17–23.
- Diesfeldt 2007a** *{published data only}*
Diesfeldt HF. Discrepancies between the IQCODE (Informant Questionnaire on Cognitive Decline in the Elderly) and cognitive test performance. *Tijdschrift voor Gerontologie en Geriatrie* 2007;**38**(5):199–209.
- Diesfeldt 2007b** *{published data only}*
Diesfeldt HF. Informant based measures may over estimate cognitive impairment in elderly patients. *International Journal of Geriatric Psychiatry* 2007;**22**(11):1166–70.
- Ehrensperger 2010** *{published data only}*
Enrensperger MM, Berres M, Taylor KI, Monsch AU. Screening properties of the German IQCODE with a two-year time frame in MCI and early Alzheimer's disease. *International Psychogeriatrics* 2010;**22**(1):91–100.
- Farias 2002** *{published data only}*
Farias ST, Mungas D, Reed B, Haan MN, Jagust WJ. Everyday imaging in relation to cognitive functioning and neuroimaging in community-dwelling Hispanic and non-Hispanic older adults. *Journal of the International Neuropsychological Society* 2004;**10**(3):342–54.
- Finneli 2009** *{published data only}*
Finelli L, Kunze U, Gautier A, Gomez-Mancilla B, Monsch A. Algorithms to retrospectively diagnose mild cognitive impairment and dementia in a longitudinal study of ageing and dementia (Abstract). *Alzheimer's and Dementia*. 2009; Vol. 5:455.
- Fuh 1995** *{published data only}*
Fuh JL, Teng EL, Lin KN, Larson EB, Wang SJ, Liu CY, et al. The Informant Questionnaire on Cognitive Decline in the Elderly as a screening tool for dementia for a predominantly illiterate Chinese population. *Neurology* 1995;**45**(1):92–6.

- Hayden 2003** *{published data only}*
Hayden KM, Khachaturian AS, Tschanz JT, Corcoran C, Nortond M, Breitner JC, Cache Country Study Group. Characteristics of a two-stage screen for incident dementia. *Journal of Clinical Epidemiology* 2003;**56**(11):1038–45.
- Hénon 2001** *{published data only}*
Hénon H, Durieu I, Guerouaou D, Lebert F, Pasquier F, Leys D. Poststroke dementia: incidence and relationship to prestroke cognitive decline. *Neurology* 2001;**57**(7):1216–22.
- Isella 2002** *{published data only}*
Isella V, Villa ML, Frattola L, Appollonio I. Screening cognitive decline in dementia preliminary data on the Italian version of the IQCODE. *Neurological Sciences* 2002;**23 Suppl 2**:s79–80.
- Isella 2006** *{published data only}*
Isella V, Villa L, Russo A, Regazzoni R, Ferrarese C, Appollonio IM. Discriminative and predictive power of an informant report in mild cognitive impairment. *Journal of Neurology, Neurosurgery, and Psychiatry* 2006;**77**(2):166–71.
- Jorm 1989** *{published data only}*
Jorm AF, Jacomb PA. The Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE) socio-demographic correlates, reliability, validity and some norms. *Psychological Medicine* 1989;**19**(4):1015–22.
- Jorm 1994** *{published data only}*
Jorm AF. A short form of the Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE): development and cross-validation. *Psychological Medicine* 1994;**24**(1):145–53.
- Jorm 1996** *{published data only}*
Jorm AF, Christensen H, Henderson AS, Jacomb PA, Korten AE, Mackinnon A. Informant ratings of cognitive decline of elderly people: relationships to longitudinal change on cognitive tests. *Age and Ageing* 1996;**25**(2):125–9.
- Jorm 1997** *{published data only}*
Jorm AF. Methods of screening for dementia: a meta-analysis of studies comparing an informant interview with a brief cognitive test. *Alzheimer Disease and Associated Disorders* 1997;**11**(3):158–62.
- Jorm 2000** *{published data only}*
Jorm AF, Christensen H, Korten AE, Jacomb PA, Henderson AS. Informant ratings of cognitive decline in old age: validation against change on cognitive tests over 7 to 8 years. *Psychological Medicine* 2000;**30**(4):981–5.
- Jorm 2003** *{published data only}*
Jorm AF. The value of informant reports for assessment and prediction of dementia. *Journal of the American Geriatrics Society* 2003;**51**(6):881–2.
- Jorm 2004** *{published data only}*
Jorm AF. The Informant Questionnaire on cognitive decline in the elderly (IQCODE): a review. *International Psychogeriatrics* 2004;**16**(3):275–93.
- Kathriarachchi 2001** *{published data only}*
Kathriarachchi ST, Sivayogan S, Jayaratna SD, Dharmasena SR. Comparison of three instruments used in the assessment of dementia in Sri Lanka. *Indian Journal of Psychiatry* 2005;**47**(2):109–12.
- Khachaturian 2000** *{published data only}*
Khachaturian AS, Gallo JJ, Breitner JC. Performance characteristics of a two-stage dementia screen in a population sample. *Journal of Clinical Epidemiology* 2000;**53**(5):531–40.
- Krogseth 2011** *{published data only}*
Krogseth M, Wyller TB, Engedal K, Juliebø V. Delirium is an important predictor of incident dementia among elderly hip fracture patients. *Dementia and Geriatric Cognitive Disorders* 2011;**31**(1):63–70.
- Larner 2010** *{published data only}*
Larner AJ. Can IQCODE differentiate Alzheimer's disease from frontotemporal dementia?. *Age and Ageing* 2010;**39**(3):392–4.
- Law 1995** *{published data only}*
Law S, Wolfson C. Validation of a French version of an informant-based questionnaire as a screening test for Alzheimer's disease. *British Journal of Psychiatry* 1995;**167**(4):541–4.
- Li 2012** *{published data only}*
Li F, Jia XF, Jia J. The Informant Questionnaire on Cognitive Decline in the Elderly individuals in screening mild cognitive impairment with or without functional impairment. *Journal Geriatric Psychiatry and Neurology* 2012;**25**(4):227–32.
- Louis 1999** *{published data only}*
Louis B, Harwood D, Hope T, Jacoby R. Can an informant questionnaire be used to predict the development of dementia in medical inpatients?. *International Journal of Geriatric Psychiatry* 1999;**14**(11):941–5.
- Mackinnon 2003** *{published data only}*
Mackinnon A, Khalilian A, Jorm AF, Korten AE, Christensen H, Mulligan R. Improving screening accuracy for dementia in a community sample by augmenting cognitive testing with informant report. *Journal of Clinical Epidemiology* 2003;**56**(4):358–66.
- Mimori 2010** *{published data only}*
Mimori Y. Cognitive decline and detection of dementia among the Japanese population: analysis with CASI and IQCODE (Abstract). Conference proceedings. 2010:s451.
- Morales 1995** *{published data only}*
Morales JM, Gonzalez-Montalvo JJ, Bermejo F, Del-Ser T. The screening of mild dementia with a shortened Spanish version of the "Informant Questionnaire on Cognitive Decline in the Elderly". *Alzheimer Disease and Associated Disorders* 1995;**9**(2):105–1.
- Morales 1997** *{published data only}*
Morales JM, Bermejo F, Romero M, Del-Ser T. Screening of dementia in community dwelling elderly through informant

- report. *International Journal of Geriatric Psychiatry* 1997;**12**(8):808–16.
- Morales-González 1992** *{published data only}*
Morales-González JM, González-Montalvo JI, Del Ser Quijano T, Bermejo Pareja F. Validation of the S-IQCODE: the Spanish version of the informant questionnaire on cognitive decline in the elderly. *Archivos de neurobiología* 1992;**55**(6):262–6.
- Ozel-Kizel 2010** *{published data only}*
Ozel-Kizel ET, Turan ED, Yilmaz E, Cangoz B, Uluc S. Discriminant validity and reliability of the Turkish version of the Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE-T). *Archives of Clinical Neuropsychology* 2010;**25**(2):139–45.
- Perroco 2009** *{published data only}*
Perroco TR, Bustamente SE, Moreno M del P, Hototian SR, Lopes MA, Azevedo D, et al. Performance of Brazilian long and short IQCODE on the screening of dementia in elderly people with low education. *International Psychogeriatrics* 2009;**21**(3):531–8.
- Potter 2009** *{published data only}*
Potter GG, Plassman BL, Burke JR, Kabeto MU, Langa KM, Llewellyn DJ, et al. Cognitive performance and informant reports in the diagnosis of cognitive impairment and dementia in African Americans and whites. *Alzheimer's & Dementia* 2009;**5**(6):445–53.
- Razavi 2014** *{published data only}*
Razavi M, Tolea MI, Margrett J, Martin P, Oakland A, Tscholl DW, et al. Comparison of 2 informant questionnaire screening tools for dementia and mild cognitive impairment: AD8 and IQCODE. *Alzheimer Disease and Associated Disorders* 2014;**28**(2):156–61.
- Ritchie 1992** *{published data only}*
Ritchie K, Fuhrer R. A comparative study of the performance of screening tests for senile dementia using receiver operating characteristics analysis. *Journal of Clinical Epidemiology* 1992;**45**(6):627–37.
- Rodriguez-Molinero 2010** *{published data only}*
Rodriguez-Molinero A, Lopez-Dieguez M, Medina IP, Tabuenca AI, De la Cruz JJ, Banegas JR. Cognitive assessment of elderly patients in the emergency department. *Revista Española de Geriatria y Gerontología* 2010;**45**:183–8.
- Rovner 2012** *{published data only}*
Rovner BW, Casten RJ, Arenson C, Salzman B, Kornsey EB. Racial differences in the recognition of cognitive dysfunction in older persons. *Alzheimer Disease and Associated Disorders* 2012;**26**(1):44–9.
- Sanchez 2009** *{published data only}*
Sanchez MA, Lourenço RA. Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE): cross-cultural adaptation for use in Brazil. *Cadernos de Saúde Pública Rio de Janeiro* 2009;**25**(7):1455–65.
- Schofield 2006** *{published data only}*
Schofield PW. Discrepancies in cognitive history from patient and informant in relation to cognitive function. *Research and Practice in Alzheimer's Disease* 2006;**11**:328–31.
- Senanorong 2001** *{published data only}*
Senanorong V, Assavisaraporn S, Sivasiriyanonds N, Printarakul T, Jamjumrus S, Udompunthurunk S, et al. The IQCODE: an alternative screening test for dementia for low educated Thai elderly. *Journal of the Medical Association of Thailand* 2001;**84**(5):648–55.
- Silpakit 2007** *{published data only}*
Silpakit O, Silpakit C, Pukdeenaul P. A comparison study of cognitive impairment screening tools: CDT, IQCODE vs MMSE. *Siriraj Medical Journal* 2007;**59**:361–3.
- Srikanth 2006** *{published data only}*
Srikanth V, Thrift AG, Fryer JL, Saling MM, Dewey HM, Sturm JW, et al. The validity of brief screening cognitive assessments in the diagnosis of cognitive impairments and dementia after first-ever stroke. *International Psychogeriatrics* 2006;**18**(2):295–305.
- Starr 2000** *{published data only}*
Starr JM, Nicolson C, Anderson K, Dennis MS, Deary IJ. Correlates of informant-rated cognitive decline after stroke. *Cerebrovascular Diseases* 2000;**10**(3):214–20.
- Thomas 1994** *{published data only}*
Thomas LD, Gonzales MF, Chamberlain A, Beyreuther K, Masters CL, Flicker L. Comparison of clinical state retrospective informant interview and the neuropathological diagnosis of Alzheimer's disease. *International Journal of Geriatric Psychiatry* 1994;**9**:233–6.
- Tokuhara 2006** *{published data only}*
Tokuhara KG, Valcour VG, Masaki KH, Blanchette PL. Utility of the Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE) for dementia in a Japanese American population. *Hawaii Medical Journal* 2006;**65**(3):72–5.
- Wiederholt 1999** *{published data only}*
Wiederholt WC, Galasko D, Salmon DP. Utility of CASI and IQCODE as screening instruments for dementia in natives of Guam (Abstract). *Journal of the Neurological Sciences*. 1997; Vol. 150 (supplement):s89.
- Wolf 2009** *{published data only}*
Wolf SA, Kubatschek K, Henry M, Harth S, Edbert AD, Wallesch CW. Informant report of cognitive changes in the elderly. A first evaluation of the German version of the IQCODE. *Der Nervenarzt* 2009;**80**(10):1178–80.
- Yamada 2000** *{published data only}*
Yamada M, Mimori Y, Sasaki H, Ikeda J, Nakamura S, Kodama K. Cognitive dysfunction among the elderly evaluated by the cognitive abilities screening instrument. *Nihon Ronen Iqakkai Zasshi* 2000;**37**:56–62.
- Zhang 2003** *{published data only}*
Zhang XQ, Zhou JS, Wang LD, Meng C, Chen B. Memory complaints in the clinical diagnosis of dementia. *Chinese Journal of Clinical Rehabilitation* 2003;**7**:4254–5.
- Zhou 2002** *{published data only}*
Zhou JS, Zhang XQ, Wang L. Telephone questionnaire: a new method for screening dementia. *Chinese Journal of Clinical Rehabilitation* 2002;**6**:3166–7.

Zhou 2003 *{published data only}*

Zhou J, Xinqing Z, Wang L, Meng C, Chu C, Chen B. Orientation memory concentration test and short IQCODE in the elderly screen dementia by telephone. *Chinese Journal of Clinical Rehabilitation* 2003;7:1529–31.

Zhou 2004 *{published data only}*

Zhou JS, Zhang XQ, Mundt JC, Wang L, Meng C, Chu C, Yang J, Chan P. Comparison of three dementia screening instruments administered by telephone in China. *Dementia (The International Journal of Social Research and Practice)* 2004;3:69–81.

Additional references**Albert 2011**

Albert MS, DeKosky ST, Dickson D, Dubois B, Feldman HH, Fox NC, et al. The diagnosis of mild cognitive impairment due to Alzheimer's disease: recommendations from the National Institute on Ageing and Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's & Dementia* 2011 May;7(3):270–9.

Ayalon 2011

Ayalon L. The IQCODE versus a single-item informant measure to discriminate between cognitively intact individuals and individuals with dementia or cognitive impairment. *Journal of Geriatric Psychiatry and Neurology* 2011;24(3):168–73.

Bahar-Fuchs 2013

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

Birks 2006

Birks J. Cholinesterase inhibitors for Alzheimer's disease. *Cochrane Database of Systematic Reviews* 2006, Issue 1. [DOI: 10.1002/14651858.CD005593]

Borson 2000

Borson S, Scanlan J, Brush M, Vitaliano P, Dokmak A. The mini-cog: a cognitive 'vital signs' measure for dementia screening in multi-lingual elderly. *International Journal of Geriatric Psychiatry* 2000;15(11):1021–7.

Bossuyt 2003

Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, et al. Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. *BMJ* 2003;326(7379):41–4.

Boustani 2003

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

Brody 2002

Brody H, Pond D, Kemp NM, Luscombe G, Harding L, Berman K, et al. The GPCOG: a new screening test

for dementia designed for general practice. *Journal of the American Geriatrics Society* 2002;50(3):530–4.

Carr 2000

Carr DB, Gray S, Baty J, Morris JC. The value of informant versus individual's complaints of memory impairment in early dementia. *Neurology* 2000;55(11):1724–6.

Chodosh 2004

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

Cordell 2013

Cordell CB, Borson S, Boustani M, Chodosh J, Reuben D, Verghese JB, et al: Medicare Detection of Cognitive Impairment Workgroup. Alzheimer's Association recommendations for operationalizing the detection of cognitive impairment during the Medicare Annual Wellness Visit in a primary care setting. *Alzheimer's & Dementia* 2013;9(2):141–50.

Cordoliani-Mackowiak 2003

Cordoliani-Mackowiak MA, Hénon H, Pruvo JP, Pasquier F, Leys D. Poststroke dementia influence of hippocampal atrophy. *Archives of Neurology* 2003;60(4):585–90.

Cullen 2007

Cullen B, O'Neill B, Evans JJ, Coen RF, Lawlor BA. A review of screening tests for cognitive impairment. *Journal of Neurology, Neurosurgery, and Psychiatry* 2007;78(8):790–9.

Davis 2013

Davis DHJ, Creavin ST, Noel-Storr A, Quinn TJ, Smailagic N, Hyde C, et al. Neuropsychological tests for the diagnosis of Alzheimer's disease dementia and other dementias: a generic protocol for cross-sectional and delayed-verification studies. *Cochrane Database of Systematic Reviews* 2013, Issue 3. [DOI: 10.1002/14651858.CD010460]

Derouesné 1999

Derouesné C, Thibault S, Lagha-Pierucci S, Baudouin-Madec V, Ancri D, Lacomblez L. Decreased awareness of cognitive deficits in patients with mild dementia of the Alzheimer type. *International Journal of Geriatric Psychiatry* 1999;14(12):1019–30.

Erkinjuntti 2000

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

Ferri 2005

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

Fisher 2007

Fisher CA, Larner AJ. Frequency and diagnostic utility of cognitive test instrument use by GPs prior to memory clinic referral. *Family Practice* 2007;24(5):495–7.

Folstein 1975

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

Galvin 2005

Galvin JE, Roe CM, Powlisha KK, Coats MA, Muich SJ, Grant E, et al. A brief informant interview to detect dementia. *Neurology* 2005;**65**(4):559-64.

Glanville 2010

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

Greenhalgh 2005

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

Gure 2010

Gure TR, Kabeto MU, Plassman BL, Piette JD, Langa KM. Differences in Functional Impairment Across Subtypes of Dementia. *Journals of Gerontology Series A: Biological Sciences and Medical Sciences* 2010;**65A**(4):434-41.

Harrison 2014

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

Hebert 2003

Hebert LE, Scherr PA, Bienas JL, Bennett DA, Evans DA. Alzheimers disease in the US population: prevalence estimates using the 2000 census. *Archives of Neurology* 2003;**60**(8):1119-22.

Hebert 2013

Hebert LE, Weuve J, Scherr PA, Evans DA. Alzheimer's disease in the United States (2010-2050) estimated using the 2010 census. *Neurology* 2013;**80**:1778-83. [DOI: 10.1212/WNL.0b013e31828726f5]

Hendry 2014

Hendry K, Lees RA, McShane R, Noel-Storr AH, Stott DJ, Quinn TJ. AD-8 for diagnosis of dementia across a variety of healthcare settings. *Cochrane Database of Systematic Reviews* 2014, Issue 5. [DOI: 10.1002/14651858.CD011121]

Hendry 2015

Hendry K, Hill E, Quinn TJ, Evans J, Stott DJ. Single screening questions for cognitive impairment in older people: a systematic review. *Age and Ageing* 2015;**44**(2): 322-6.

Holsinger 2007

Holsinger T, Deveau J, Boustani M, Williams JW Jr. Does this patient have dementia?. *JAMA* 2007;**297**(21): 2391-404.

Imison 2012

Imison C, Poteliakhoff E, Thompson J. Older people and emergency bed use. Exploring variation. www.kingsfund.org.uk/sites/files/kf/field/field_publication_file/older-people-and-emergency-bed-use-aug-2012.pdf 2012 (accessed 1st March 2015).

Jorm 1988

Jorm AF, Korten AE. Assessment of cognitive decline in the elderly by informant interview. *British Journal of Psychiatry* 1988;**152**:209-13.

Knapskog 2014

Knapskog A-B, Barca ML, Engedal K. Prevalence of depression among memory clinic patients as measured by the Cornell Scale of Depression in Dementia. *Ageing & Mental Health* 2014;**18**(5):579-587.

Larner 2009

Larner AJ. 'Attended alone' sign: validity and reliability for the exclusion of dementia. *Age and Ageing* 2009;**38**(4): 476-8.

Lees 2013

Lees R, Corbet S, Johnston C, Moffitt E, Shaw G, Quinn TJ. Test accuracy of short screening tests for diagnosis of delirium or cognitive impairment in an acute stroke unit setting. *Stroke* 2013;**44**(11):3078-83.

Matthews 2009

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

Matthews 2013

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

McKeith 2005

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

McKhann 1984

McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM. Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology* 1984;**34**(7): 939-44.

McKhann 2001

McKhann GM, Albert MS, Grossman M, Miller B, Dickson D, Trojanowski JQ. Work Group on Frontotemporal Dementia and Pick's Disease. Clinical and pathological diagnosis of frontotemporal dementia: report of the Work

- Group on Frontotemporal Dementia and Pick's Disease. *Archives of Neurology* 2001;**58**(11):1803–9.
- McKhann 2011**
McKhann GM, Knopman DS, Chertkow H, Hyman BT, Jack CR Jr, Kawas CH, et al. The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging and the Alzheimer's Association Workgroup. *Alzheimer's & Dementia* 2011;**7**(3):263–9.
- McShane 2006**
McShane R, Areosa Sastre A, Minakaran N. Memantine for dementia. *Cochrane Database of Systematic Reviews* 2006, Issue 2. [DOI: 10.1002/14651858.CD003154.pub5]
- Nasreddine 2005**
Nasreddine ZS, Phillips NA, Bédirian V, Charbonneau S, Whitehead V, Collin I, et al. The Montreal Cognitive Assessment, MoCA: A Brief Screening Tool for Mild Cognitive Impairment. *Journal of the American Geriatrics Society* 2005;**53**(4):695–9.
- Noel-Storr 2014**
Noel-Storr AH, McCleery JM, Richard E, Ritchie CW, Flicker L, Cullum SJ, et al. Reporting standards for studies of diagnostic test accuracy in dementia: The STARDdem Initiative. *Neurology* 2014;**83**(4):364–73.
- Oliveira 2011**
Oliveira MR, Gomes AC, Toscano CM. QUADAS and STARD: evaluating the quality of diagnostic accuracy studies. *Revista de Saúde Pública* 2011;**45**(2):416–22.
- Prince 2013**
Prince M, Bryce R, Albanese E, Wimo A, Ribeiro W, Ferri CP. The global prevalence of dementia: A systematic review and metaanalysis. *Alzheimer's & Dementia* 2013;**9**(1):63–75.
- Quinn 2014**
Quinn TJ, Fearon P, Noel-Storr AH, Young C, McShane R, Stott DJ. Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE) for the diagnosis of dementia within community dwelling populations. *Cochrane Database of Systematic Reviews* 2014, Issue 4. [DOI: 10.1002/14651858.CD010079.pub2]
- Reitsma 2005**
Reitsma JB, Glas AS, Rutjes AW, Scholten RJ, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *Journal of Clinical Epidemiology* 2005;**58**(10):982–90.
- RevMan 2014**
The Nordic Cochrane Centre, The Cochrane Collaboration. Review Manager (RevMan). 5.3. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2014.
- Rockwood 1998**
Rockwood K, Howard K, Thomas VS, Mallery L, MacKnight C, Sangalang V, et al. Retrospective diagnosis of dementia using an informant interview based on the Brief Cognitive Rating Scale. *International Psychogeriatrics* 1998;**10**(1):53–60.
- Roman 1993**
Roman GC, Tatemichi TK, Erkinjuntti T, Cummings JL, Masdeu JC, Garcia JH, et al. Vascular dementia: diagnostic criteria for research studies. Report of the NINDS-AIREN International Workshop. *Neurology* 1993;**43**(2):250–60.
- Ryan 2013**
Ryan DJ, O'Regan NA, Caoimh RO, Clare J, O'Connor M, Leonard M, et al. Delirium in an adult acute hospital population: predictors, prevalence and detection. *BMJ Open* 2013;**3**(1):e001772.
- Sampson 2009**
Sampson EL, Blanchard MR, Jones L, Tookman A, King M. Dementia in the acute hospital: prospective cohort study of prevalence and mortality. *British Journal of Psychiatry* 2009;**195**(1):61–6.
- Savva 2009**
Savva GM, Wharton SB, Ince PG, Forster G, Matthews FE, Brayne C, et al. Age, neuropathology and dementia. *New England Journal of Medicine* 2009;**360**(22):2302–9.
- Shenkin 2014**
Shenkin SD, Russ TC, Ryan TM, MacLulich AMJ. Screening for dementia and other causes of cognitive impairment in general hospital in-patients. *Age and Ageing* 2014;**43**(2):166–168.
- Smeeth 2001**
Smeeth L, Fletcher AE, Stirling S, Nunes M, Breeze E, Ng E, et al. Randomised comparison of three methods of administering a screening questionnaire to elderly people: findings from the MRC trial of the assessment and management of older people in the community. *BMJ* 2001;**323**(7326):1403–7.
- Sperling 2011**
Sperling RA, Aisen PS, Beckett LA, Bennett DA, Craft S, Fagan AM, et al. Toward defining the pre-clinical stages of Alzheimer's disease: recommendations from the National Institute on Aging and the Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's & Dementia* 2011;**7**(3):280–92.
- Stephan 2010**
Stephan BC, Kurth T, Matthews FE, Brayne C, Dufouil C. Dementia risk prediction in the population: are screening models accurate?. *Nature Reviews. Neurology* 2010;**6**(6):318–26.
- Valcour 2000**
Valcour VG, Masaki KH, Curb JD, Blanchette PL. The detection of dementia in the primary care setting. *Archives of Internal Medicine* 2000;**160**(19):2964–8.
- Vieira 2013**
Vieira RT, Caixeta L, Machado S, Silva AC, Nardi AE, Arias-Carrión O, et al. Epidemiology of early-onset dementia: a review of the literature. *Clinical Practice and Epidemiology in Mental Health* 2013;**9**:88–95.
- Wharton 2015**
Wharton SB, Simpson JE, Brayne C, Ince PG. Age-associated white matter lesions: The MRC Cognitive

Function and Ageing Study. *Brain Pathology* 2015;25(1):
35–43.

References to other published versions of this review

Quinn 2013

Quinn TJ, Fearon P, Young C, Noel-Storr AH, McShane R, Stott DJ. IQCODE for the diagnosis of Alzheimer's disease dementia and other dementias within a secondary care setting. *Cochrane Database of Systematic Reviews* 2013, Issue 10.

* *Indicates the major publication for the study*

CHARACTERISTICS OF STUDIES

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

Flicker 1997

Study characteristics			
Patient sampling	Participants were consecutive patients seen at the memory clinic who had complete data available and 100 patients randomly selected who had been reviewed by the aged-care assessment team (ACAT) who consented to participate		
Patient characteristics and setting	Included participants came from 2 sources: those referred to the memory clinic in a hospital in Melbourne, Australia and those referred to the ACAT over a 6-month period		
Index tests	IQCODE 26 item, English language		
Target condition and reference standard(s)	Clinical dementia diagnosis using DSM-III-R		
Flow and timing	Of 437 possible consecutive memory-clinic participants, 299 were included on grounds of complete available data and where the assessments could be made in English, without involving translation services. From 100 possible ACAT patients, 78 were included based on willingness to consent in assessments and informant interviews It is not clear if the results of the index test informed the conclusions of the assessments to form a clinical diagnosis (reference standard)		
Comparative			
Notes			
Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	No		
			Low
DOMAIN 2: Index Test All tests			

Flicker 1997 (Continued)

If a threshold was used, was it pre-specified?	No			
				Unclear
DOMAIN 3: Reference Standard				
Is the reference standards likely to correctly classify the target condition?	Yes			
				Low
DOMAIN 4: Flow and Timing				
Was there an appropriate interval between index test and reference standard?	Yes			
Were all patients included in the analysis?	Unclear			

Garcia 2002

Study characteristics	
Patient sampling	Consecutive sample of patients attending a Geriatric External Facility accompanied by a family member
Patient characteristics and setting	113 participants not previously diagnosed with dementia referred due to memory loss, behavioural disorder and/or cognitive deterioration. Geriatric external facility in Spain
Index tests	IQCODE 16 item, Spanish
Target condition and reference standard(s)	Clinical dementia diagnosis using DSM-III-R
Flow and timing	All selected patients underwent IQCODE assessment. This was administered by a different physician the same day, blinded to clinical diagnosis
Comparative	
Notes	
Methodological quality	

Garcia 2002 (Continued)

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	No		
			Low
DOMAIN 2: Index Test All tests			
If a threshold was used, was it pre-specified?	Yes		
			Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
			Low
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Were all patients included in the analysis?	Yes		

Goncalves 2011

Study characteristics	
Patient sampling	Cross-sectional study of consecutive memory clinic attenders attending with an informant

Goncalves 2011 (Continued)

Patient characteristics and setting	Participants were referred by their primary care physicians Memory clinic, city hospital in Brisbane Australia, n = 204		
Index tests	IQCODE 16 item, English language		
Target condition and reference standard(s)	Clinical diagnosis of dementia using DSM-IV-TR criteria including IQCODE result		
Flow and timing	Of 243 potential subjects, 208 attended with an informant. A further 4 were excluded due to missing cognitive test data, final sample n = 204 Index test and initial assessment performed together; clinical assessment by psychiatrist performed 2 weeks later with knowledge of all results and reference standard determined at that point		
Comparative			
Notes			
Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
			Low
DOMAIN 2: Index Test All tests			
If a threshold was used, was it pre-specified?	Yes		
			Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
			Low

Goncalves 2011 (Continued)

DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Were all patients included in the analysis?	No		

Hancock 2009

Study characteristics			
Patient sampling	Consecutive new patient referrals to memory clinic attending with an informant		
Patient characteristics and setting	Memory clinics in a psychiatric hospital and cognitive function clinic based in a regional neuroscience centre in the UK, n = 144		
Index tests	IQCODE 16 item, English language		
Target condition and reference standard(s)	Clinical dementia diagnosis using DSM-IV		
Flow and timing	144 included, no figures to quote how many ineligible over study period Index test performed independently from clinical assessment and not used to assess reference standard. Both tests performed on same day		
Comparative			
Notes			
Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		

Hancock 2009 (Continued)

					Low
DOMAIN 2: Index Test All tests					
If a threshold was used, was it pre-specified?	No				
					Unclear
DOMAIN 3: Reference Standard					
Is the reference standards likely to correctly classify the target condition?	Yes				
					Low
DOMAIN 4: Flow and Timing					
Was there an appropriate interval between index test and reference standard?	Yes				
Were all patients included in the analysis?	Yes				

Harwood 1997

Study characteristics	
Patient sampling	Random sample of urgent admissions in over-65s. Sample was determined by assigning a number on a card to each admission and selecting half of the cards each day at random
Patient characteristics and setting	Unscheduled admissions (aged > 65 years) admitted to an acute medical unit of a UK teaching hospital in Oxford (n = 201)
Index tests	IQCODE 16 item, English language
Target condition and reference standard(s)	Clinical dementia diagnosis using DSM-III-R
Flow and timing	Of 223 potential participants, 13 died prior to assessment or were excluded due to being 'too ill' to participate, 7 declined consent, 2 were excluded due to absence at time of assessment - result in n = 201 participants Index test and reference standard conducted contemporaneously, but not blinded

Harwood 1997 (Continued)

Comparative			
Notes			
Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
			Low
DOMAIN 2: Index Test All tests			
If a threshold was used, was it pre-specified?	No		
			High
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
			Low
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Were all patients included in the analysis?	Yes		

Jorm 1991

Study characteristics			
Patient sampling	Participants were recruited as part of a study investigating the reliability and validity of the Canberra Interview for the Elderly (CIE); recruited as a consecutive series. An additional non-consecutive group were recruited elsewhere (principally psychogeriatricians)		
Patient characteristics and setting	Patients were a series of those assessed by geriatricians as inpatients, day-hospital attenders or outpatient clinic attenders (n = 64). As the authors were interested in assessing the score in depression, they also recruited a subgroup of patients, principally via psychogeriatricians (n = 12). Study conducted in Canberra Australia		
Index tests	IQCODE 26 item, English language		
Target condition and reference standard(s)	Clinical dementia diagnosis using DSM-III-R and ICD-10		
Flow and timing	Of 72 potential participants from the geriatric settings 64 agreed to participate. 12 patients were recruited from other sources. Of the 76 potential participants, 7 were excluded due to missing data (index test or MMSE) so total sample was 69 Reference standard determined by the treating clinician. Index test administered by lay interviewers at an interval afterwards. Unclear if this was blinded to clinical diagnosis. Further index test assessment performed to assess test-retest reliability		
Comparative			
Notes			
Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
			High
DOMAIN 2: Index Test All tests			
If a threshold was used, was it pre-specified?	No		

Jorm 1991 (Continued)

					Low
DOMAIN 3: Reference Standard					
Is the reference standards likely to correctly classify the target condition?	Yes				
					Low
DOMAIN 4: Flow and Timing					
Was there an appropriate interval between index test and reference standard?	Unclear				
Were all patients included in the analysis?	Unclear				

Knaefelc 2003

Study characteristics					
Patient sampling	Memory clinic attenders over a 10-year period				
Patient characteristics and setting	Memory clinic based in a geriatric hospital in Melbourne Australia, n = 323				
Index tests	IQCODE 16 item, English language				
Target condition and reference standard(s)	Clinical dementia diagnosis using DSM-IV				
Flow and timing	Of 519 potential participants; 426 had both assessments; 103 were excluded as they were non-English speakers Index test was administered independent from reference standard. Clinical diagnosis (reference standard) was made by a geriatrician and a psychiatrist, without knowledge of index test result				
Comparative					
Notes					
Methodological quality					
Item	Authors' judgement	Risk of bias	Applicability concerns		

Knaefelc 2003 (Continued)

DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
			Low
DOMAIN 2: Index Test All tests			
If a threshold was used, was it pre-specified?	No		
			Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
			Unclear
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Were all patients included in the analysis?	Unclear		

Mackinnon 1998

Study characteristics	
Patient sampling	Participants from the geriatric hospital or memory clinic; sampling frame unclear
Patient characteristics and setting	Participants came from a university hospital in Switzerland. The sample included geriatric patients and those referred to the memory clinic; n = 106, no breakdown by setting

Mackinnon 1998 (Continued)

Index tests	IQCODE 16 item, French language		
Target condition and reference standard(s)	Clinical dementia diagnosis using DSM-IV		
Flow and timing	Participants underwent cognitive testing and IQCODE assessment prior to assessment using the reference standard. However, the index test was conducted contemporaneously with reference standard. Clinical diagnosis (reference standard) conducted blinded to results of index test		
Comparative			
Notes			
Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
			Unclear
DOMAIN 2: Index Test All tests			
If a threshold was used, was it pre-specified?	No		
			Unclear
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
			Low
DOMAIN 4: Flow and Timing			

Mackinnon 1998 (Continued)

Was there an appropriate interval between index test and reference standard?	Yes		
Were all patients included in the analysis?	Yes		

Mulligan 1996

Study characteristics			
Patient sampling	Mixed group of those admitted to geriatric hospital or outpatients at the memory clinic; sampling frame unclear		
Patient characteristics and setting	University hospital in Switzerland - inpatient geriatric admissions and outpatients referred to the memory clinic; total sample n = 76 (no breakdown available by recruitment setting)		
Index tests	IQCODE 26 item, French language		
Target condition and reference standard(s)	Clinical dementia diagnosis using DSM-III-R		
Flow and timing	Index test and reference standard conducted contemporaneously. Index test administrators blinded to results of diagnosis. Diagnosis made by senior psychiatrists		
Comparative			
Notes			
Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
			Unclear

Mulligan 1996 (Continued)

DOMAIN 2: Index Test All tests	
If a threshold was used, was it pre-specified?	No
	Unclear
DOMAIN 3: Reference Standard	
Is the reference standards likely to correctly classify the target condition?	Yes
	Unclear
DOMAIN 4: Flow and Timing	
Was there an appropriate interval between index test and reference standard?	Yes
Were all patients included in the analysis?	Yes

Narasimhalu 2008

Study characteristics	
Patient sampling	Patients were obtained from 2 sources: (1) Consecutive referrals to hospital dementia clinic (2) Participants in another study investigating cognition following stroke The study only included those who had completed an MMSE & IQCODE assessment
Patient characteristics and setting	General hospital setting in Singapore; group (1) were referred to the dementia clinic; included 237 out of 695 evaluated. For group (2) 355 included out of 398 who received both tests from 843 total enrolled participants
Index tests	IQCODE 16 item, Cantonese language
Target condition and reference standard(s)	Clinical dementia diagnosis using DSM-IV
Flow and timing	Unclear order of conduct of the index test and reference standard
Comparative	

Notes			
Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	No		
			High
DOMAIN 2: Index Test All tests			
If a threshold was used, was it pre-specified?	No		
			Unclear
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
			Unclear
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Were all patients included in the analysis?	Unclear		

Sikkes 2010

Study characteristics			
Patient sampling	All consecutive patients diagnosed with probable AD, MCI (mild cognitive impairment) or SMC (subjective memory complaint) whose informant completed an IQCODE were included		
Patient characteristics and setting	Alzheimer Centre at a University Hospital in the Netherlands; n = 328		
Index tests	IQCODE 16 item, Dutch language		
Target condition and reference standard(s)	Clinical dementia diagnosis using NINCDS-ADRDA		
Flow and timing	328 participants included - no record given of numbers screened but not included Informants completed IQCODE while participants were assessed for clinical diagnosis and this was done independently of IQCODE result		
Comparative			
Notes			
Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	No		
			High
DOMAIN 2: Index Test All tests			
If a threshold was used, was it pre-specified?	No		
			Low
DOMAIN 3: Reference Standard			

Sikkes 2010 (Continued)

Is the reference standards likely to correctly classify the target condition?	Yes		
Low			
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Were all patients included in the analysis?	Yes		

Siri 2006

Study characteristics			
Patient sampling	Participants were those referred to a geriatric clinic aged > 60 years, sampling frame unclear		
Patient characteristics and setting	Geriatric clinic attenders at a university hospital in Bangkok Thailand. 200 elderly people, divided into 100 'normal' and 100 'demential'		
Index tests	IQCODE 32 item, Thai language		
Target condition and reference standard(s)	Clinical dementia diagnosis using DSM-IV		
Flow and timing	Participants had to have an informant in order to be eligible. No report of how many potential participants compared with n = 200 (100 'normal', 100 'demential') Assessment of IQCODE conducted blinded to clinical diagnosis		
Comparative			
Notes			
Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		

Siri 2006 (Continued)

Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	No		
Unclear			
DOMAIN 2: Index Test All tests			
If a threshold was used, was it pre-specified?	No		
High			
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Unclear			
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Were all patients included in the analysis?	Unclear		

Tang 2003

Study characteristics	
Patient sampling	Consecutive admissions of first-ever or recurrent stroke
Patient characteristics and setting	484 were admitted to the Acute Stroke Unit of a general teaching hospital in Hong Kong in the study period; n = 189 included
Index tests	IQCODE 26 item, Chinese language
Target condition and reference standard(s)	Clinical dementia diagnosis using DSM-IV

Tang 2003 (Continued)

Flow and timing	Of 484 potential participants, 471 had an informant available. 95 were excluded due to their response to IQCODE items (> 20% scored 'I don't know'); 17 were excluded due to 'physical frailty' and 18 due to 'prolonged hospitalisation' IQCODE assessment was performed independently from clinical assessment by a psychiatrist, both were conducted 3 months after index stroke. The clinical assessment was performed in outpatient clinic setting		
Comparative			
Notes			
Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
			High
DOMAIN 2: Index Test All tests			
If a threshold was used, was it pre-specified?	No		
			High
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
			Low
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		

Tang 2003 (Continued)

Were all patients included in the analysis?	No		

Characteristics of excluded studies [ordered by study ID]

Study	Reason for exclusion
Abreu 2008	Case-control design
Bustamente 2003	Case-control methodology
Butt 2008	Community population
Cherbuin 2008	No original data
Cherbuin 2013	No original data - review article
De Jonge 1997	Data not suitable for analysis
Dekkers 2009	Data not suitable for analysis
Diesfeldt 2007a	No dementia diagnosis reference standard
Diesfeldt 2007b	No original data - repeat dataset
Ehrensperger 2010	Uses unvalidated (2-year) IQCODE
Farias 2002	No dementia diagnosis reference standard
Finneli 2009	Data not suitable for analysis
Fuh 1995	Case-control methodology
Hayden 2003	Small dataset; IQCODE administered to < 10 participants
Hénon 2001	Uses a delayed verification analysis
Isella 2002	Uses a delayed verification analysis
Isella 2006	Data not suitable for analysis
Jorm 1989	Data not suitable for analysis, and no dementia diagnosis reference standard

(Continued)

Jorm 1994	Community setting
Jorm 1996	Community setting
Jorm 1997	No original data
Jorm 2000	No dementia diagnosis reference standard
Jorm 2003	No original data
Jorm 2004	No original data - review article
Kathirarachi 2001	Community setting
Khachaturian 2000	No IQCODE index test data
Krogseth 2011	Uses a delayed verification analysis
Larner 2010	Looks at diagnostic accuracy comparing 2 dementia types rather than dementia/no dementia dichotomy
Law 1995	Community setting
Li 2012	No dementia diagnosis reference standard
Louis 1999	Uses a delayed verification analysis
Mackinnon 2003	Community setting
Mimori 2010	No original data
Morales 1995	Community setting
Morales 1997	Community setting
Morales-González 1992	Case-control methodology
Ozel-Kizel 2010	Case-control methodology
Perroco 2009	Case-control methodology
Potter 2009	Data not suitable for analysis
Razavi 2014	Case-control methodology; data not available without added controls
Ritchie 1992	No IQCODE index test data

(Continued)

Rodriguez-Molinero 2010	No dementia diagnosis reference standard
Rovner 2012	Data not suitable for analysis
Sanchez 2009	No dementia diagnosis reference standard
Schofield 2006	Data not suitable for analysis
Senanorong 2001	Community setting
Silpakit 2007	Case-control methodology
Srikanth 2006	Community setting
Starr 2000	No dementia diagnosis reference standard
Thomas 1994	Non-contemporaneous assessment of reference standard and index test
Tokuhara 2006	Primary-care setting
Wierderholt 1999	Data not suitable for analysis
Wolf 2009	No dementia diagnosis reference standard
Yamada 2000	Community setting
Zhang 2003	Data not suitable for analysis
Zhou 2002	Case-control methodology
Zhou 2003	No original data - repeat dataset
Zhou 2004	No original data - repeat dataset

DATA

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

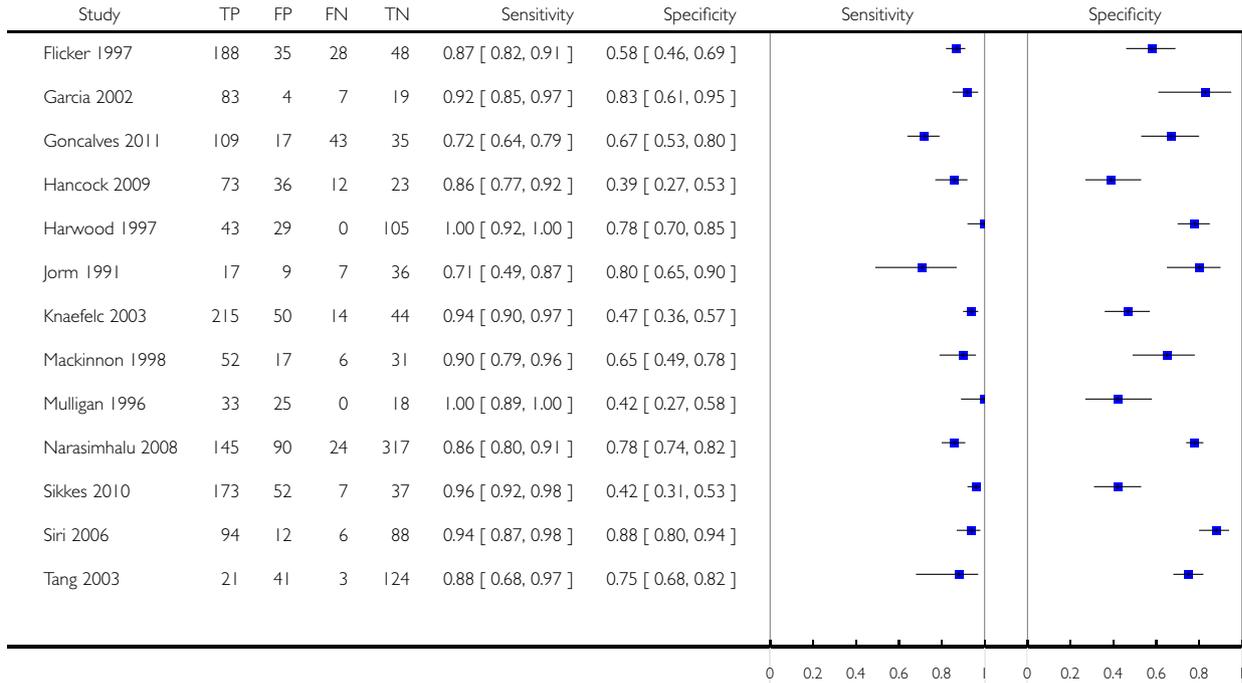
Tests. Data tables by test

Test	No. of studies	No. of participants
1 All studies IQCODE 3.3 or closest	13	2745
2 All 16-item IQCODE	7	1768
3 All 26-item IQCODE	6	977
4 IQCODE 3.3 Threshold	4	722
5 IQCODE 3.4 Threshold	4	1211
6 IQCODE 3.5 Threshold	1	269
7 IQCODE 3.6 Threshold	9	1576
8 IQCODE >3.6 Threshold	3	772
9 16-item IQCODE 3.3 Threshold	2	446
10 16-item IQCODE 3.4 Threshold	3	1022
11 16-item IQCODE 3.5 Threshold	1	269
12 16-item IQCODE 3.6 Threshold	5	988
13 26-item IQCODE 3.3 Threshold	2	276
14 26-item IQCODE 3.4 Threshold	1	189
15 26-item IQCODE 3.6 Threshold	4	588
16 Sensitivity analysis removing Goncalves	12	2541
17 Sensitivity analysis removing low average age	10	1756

Test 1. All studies IQCODE 3.3 or closest.

Review: Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE) for the diagnosis of dementia within a secondary care setting

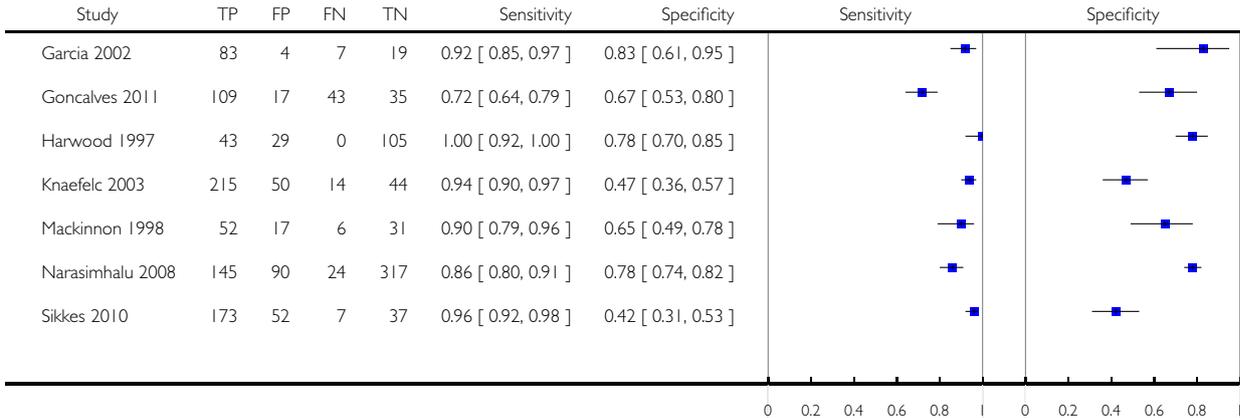
Test: I All studies IQCODE 3.3 or closest



Test 2. All 16-item IQCODE.

Review: Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE) for the diagnosis of dementia within a secondary care setting

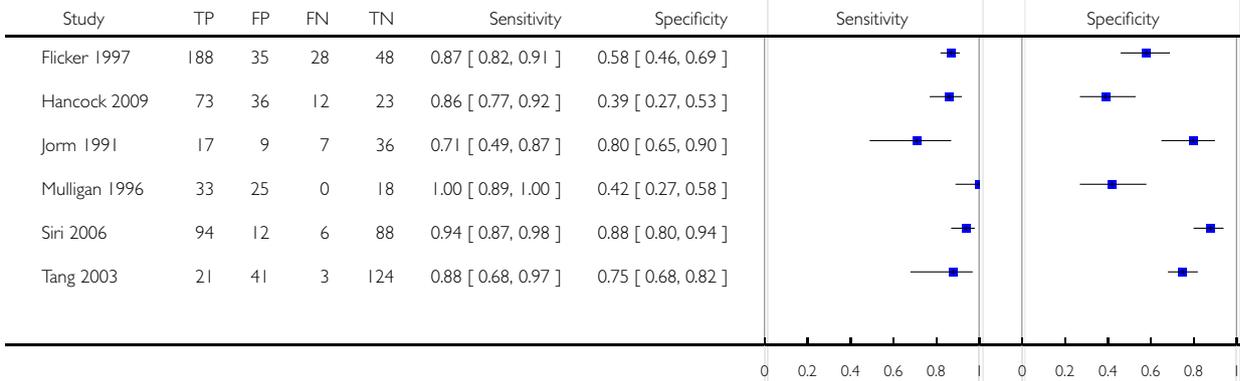
Test: 2 All 16-item IQCODE



Test 3. All 26-item IQCODE.

Review: Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE) for the diagnosis of dementia within a secondary care setting

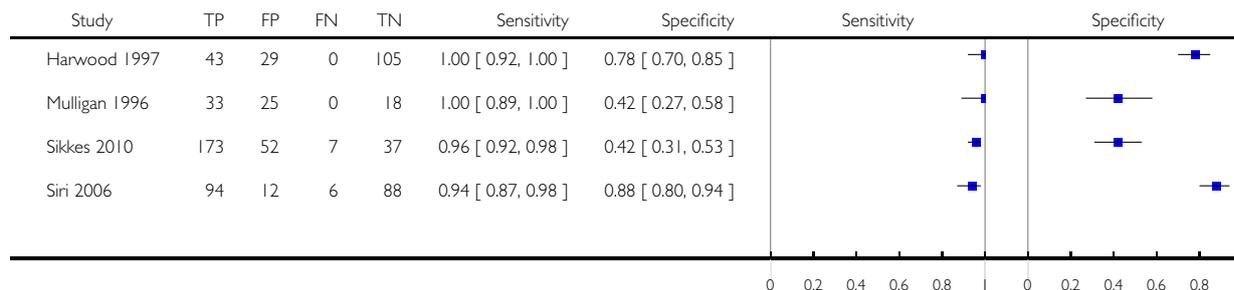
Test: 3 All 26-item IQCODE



Test 4. IQCODE 3.3 Threshold.

Review: Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE) for the diagnosis of dementia within a secondary care setting

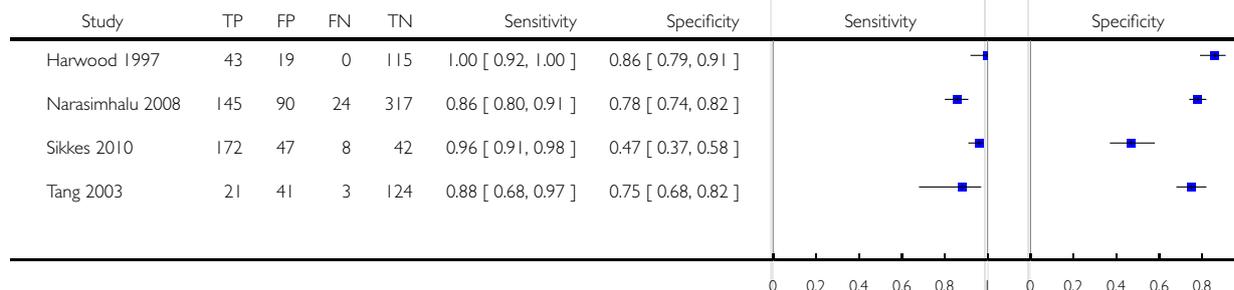
Test: 4 IQCODE 3.3 Threshold



Test 5. IQCODE 3.4 Threshold.

Review: Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE) for the diagnosis of dementia within a secondary care setting

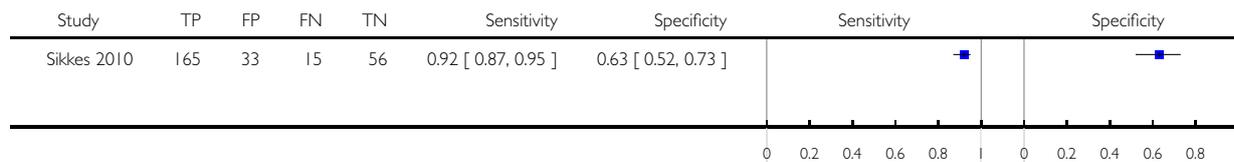
Test: 5 IQCODE 3.4 Threshold



Test 6. IQCODE 3.5 Threshold.

Review: Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE) for the diagnosis of dementia within a secondary care setting

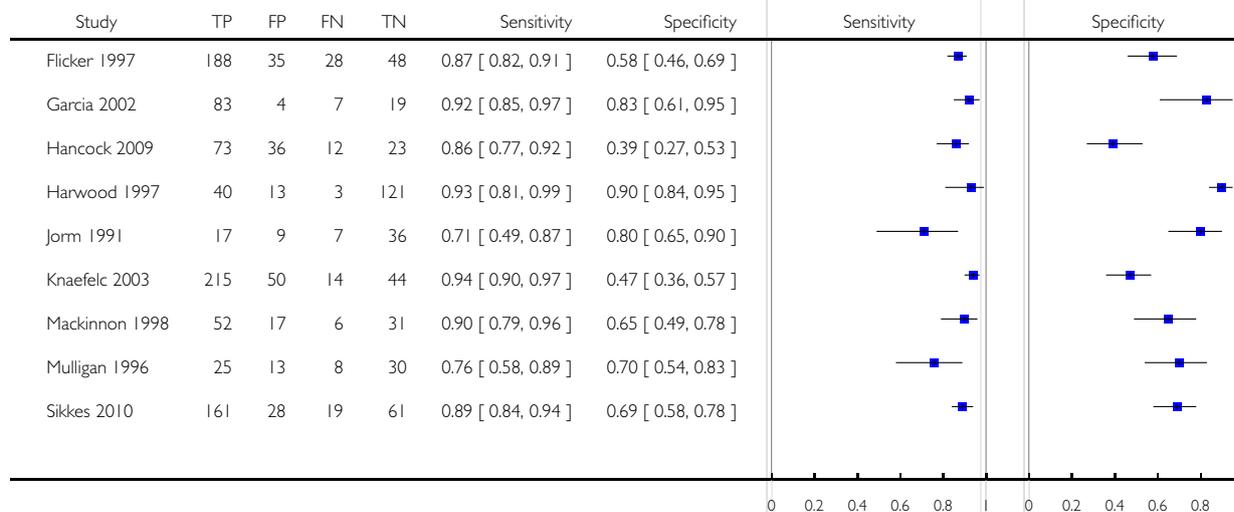
Test: 6 IQCODE 3.5 Threshold



Test 7. IQCODE 3.6 Threshold.

Review: Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE) for the diagnosis of dementia within a secondary care setting

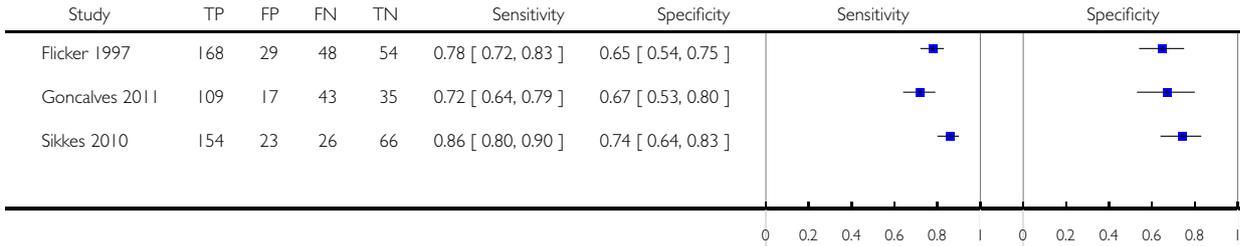
Test: 7 IQCODE 3.6 Threshold



Test 8. IQCODE >3.6 Threshold.

Review: Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE) for the diagnosis of dementia within a secondary care setting

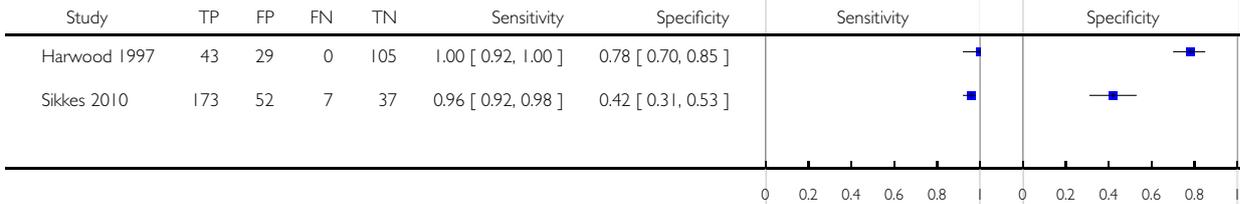
Test: 8 IQCODE >3.6 Threshold



Test 9. 16-item IQCODE 3.3 Threshold.

Review: Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE) for the diagnosis of dementia within a secondary care setting

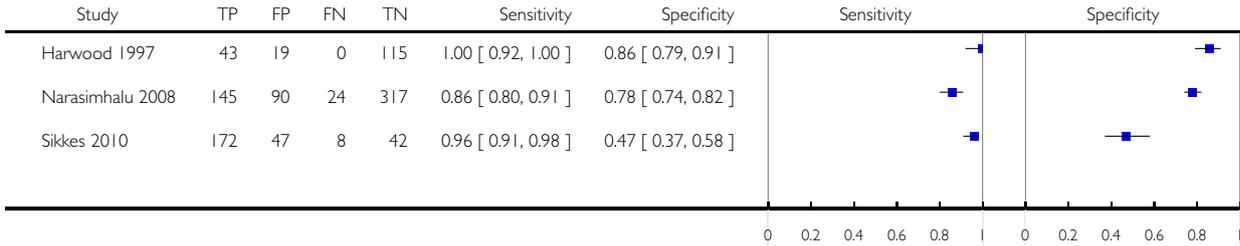
Test: 9 16-item IQCODE 3.3 Threshold



Test 10. 16-item IQCODE 3.4 Threshold.

Review: Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE) for the diagnosis of dementia within a secondary care setting

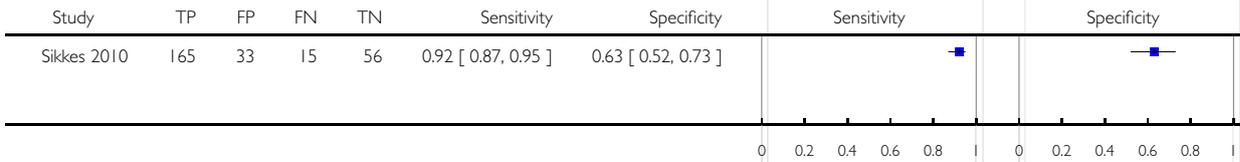
Test: 10 16-item IQCODE 3.4 Threshold



Test 11. 16-item IQCODE 3.5 Threshold.

Review: Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE) for the diagnosis of dementia within a secondary care setting

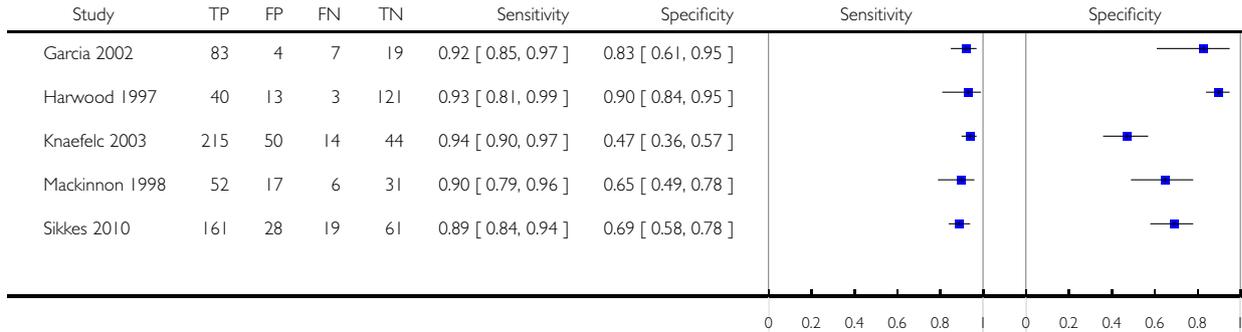
Test: 11 16-item IQCODE 3.5 Threshold



Test 12. 16-item IQCODE 3.6 Threshold.

Review: Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE) for the diagnosis of dementia within a secondary care setting

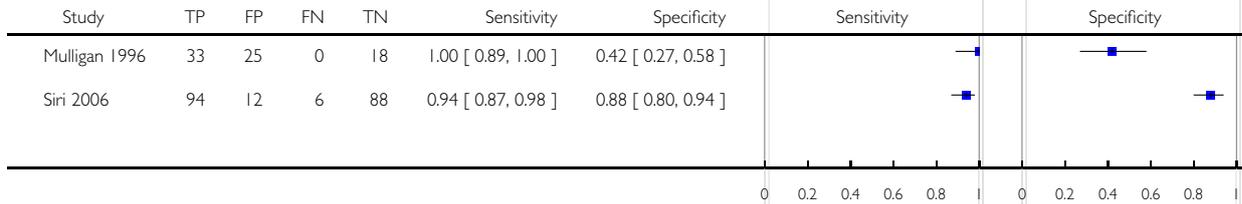
Test: 12 16-item IQCODE 3.6 Threshold



Test 13. 26-item IQCODE 3.3 Threshold.

Review: Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE) for the diagnosis of dementia within a secondary care setting

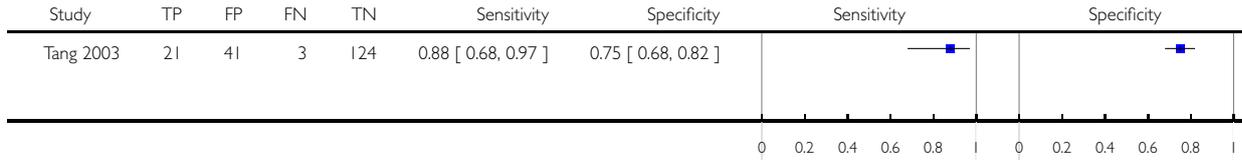
Test: 13 26-item IQCODE 3.3 Threshold



Test 14. 26-item IQCODE 3.4 Threshold.

Review: Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE) for the diagnosis of dementia within a secondary care setting

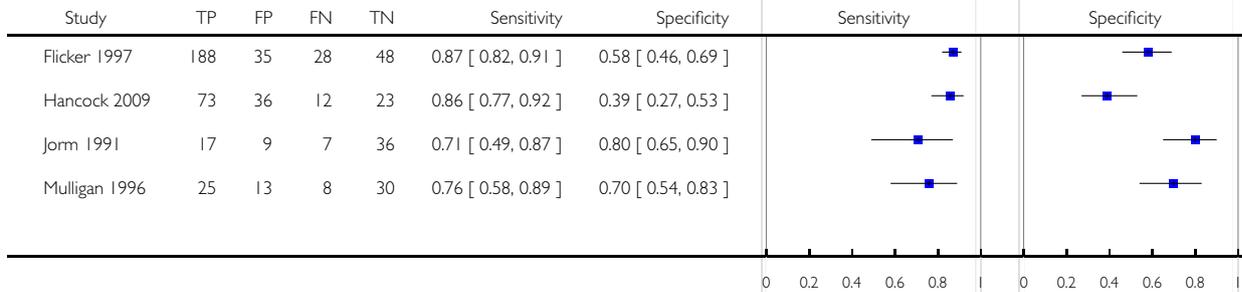
Test: 14 26-item IQCODE 3.4 Threshold



Test 15. 26-item IQCODE 3.6 Threshold.

Review: Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE) for the diagnosis of dementia within a secondary care setting

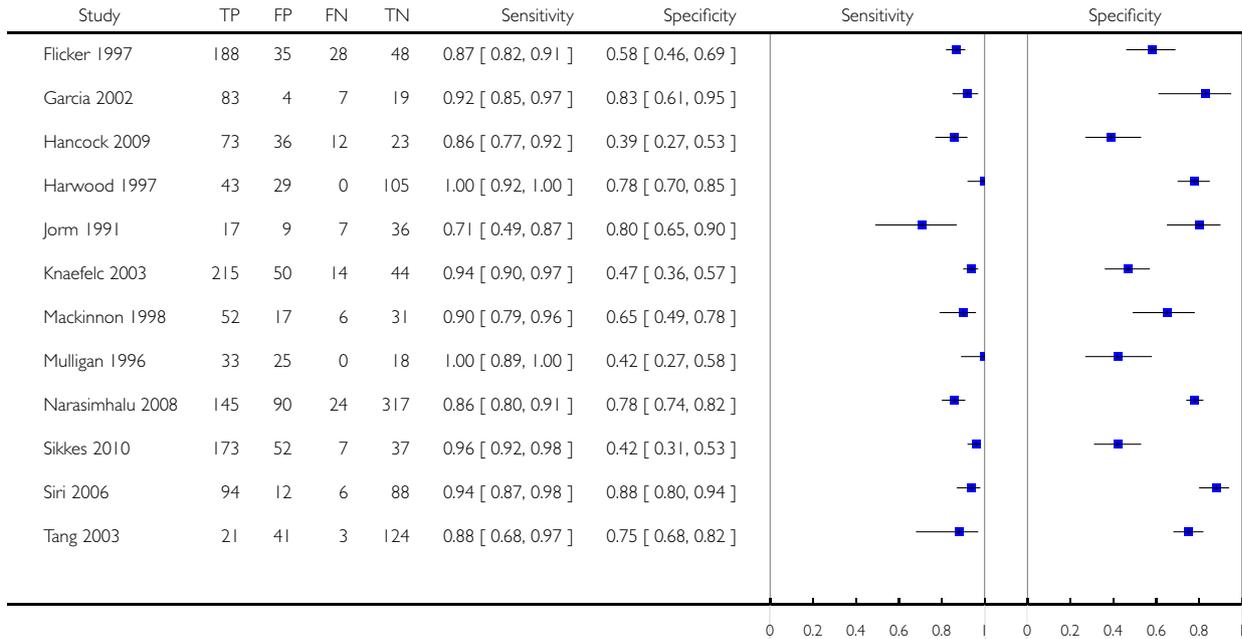
Test: 15 26-item IQCODE 3.6 Threshold



Test 16. Sensitivity analysis removing Goncalves.

Review: Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE) for the diagnosis of dementia within a secondary care setting

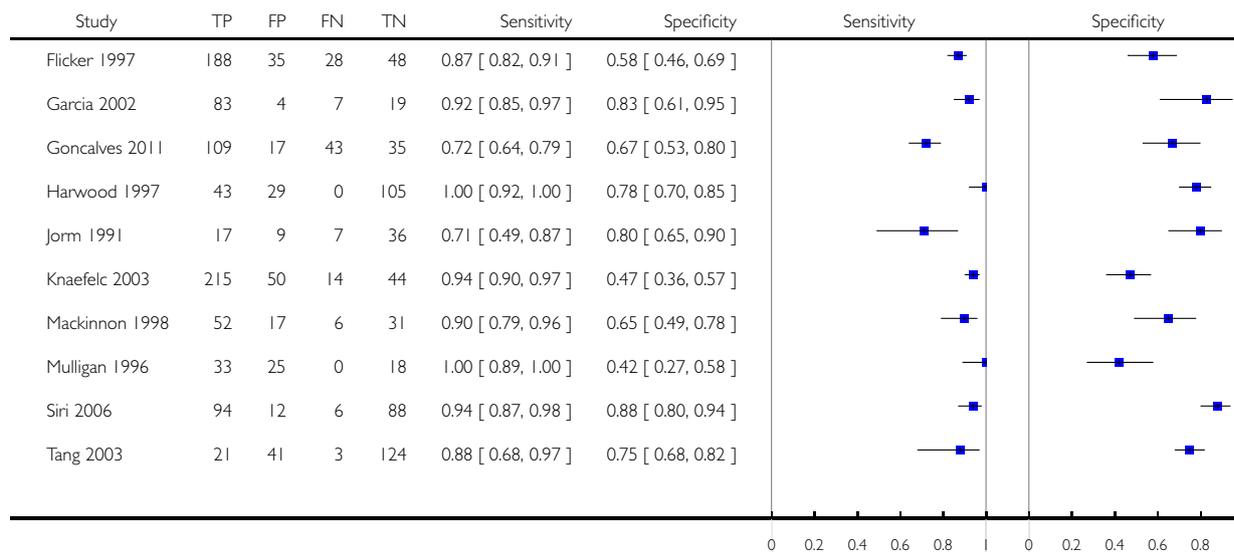
Test: 16 Sensitivity analysis removing Goncalves



Test 17. Sensitivity analysis removing low average age.

Review: Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE) for the diagnosis of dementia within a secondary care setting

Test: 17 Sensitivity analysis removing low average age



ADDITIONAL TABLES

Table 1. Summary of test accuracy at study level

Study ID	Participants (n)	Primary threshold	Sensitivity (%)	Specificity (%)
Flicker 1997	299*	3.6	87	58
Garcia 2002	103	3.6	92	81
Goncalves 2011	204	4.1	72	67
Hancock 2009	144	3.6	86	39
Harwood 1997	177	3.3	100	78
Jorm 1991	69	3.6	71	80
Knaefelc 2003	323	3.6	94	47
Mackinnon 1998	106	3.6	90	65
Mulligan 1996	76	3.3	100	42

Table 1. Summary of test accuracy at study level (Continued)

Narasimhalu 2008	576	3.4	86	78
Sikkas 2010	269*	3.3	96	42
Siri 2006	200	3.3	94	88
Tang 2003	189	3.4	88	75

Where multiple thresholds were reported, we used the value closest to 3.3 to populate this table

*Total number of participants adjusted to reflect numbers included in quantitative synthesis

APPENDICES

Appendix I. WHO International Classification of Disease - Dementia

World Health Organization International Classification of Diseases 10

F00 - F09 ORGANIC, INCLUDING SYMPTOMATIC, MENTAL DISORDERS

DEMENTIA

G1. Evidence of each of the following:

(1) A decline in memory, which is most evident in the learning of new information, although in more severe cases, the recall of previously learned information may be also affected. The impairment applies to both verbal and non-verbal material. The decline should be objectively verified by obtaining a reliable history from an informant, supplemented, if possible, by neuropsychological tests or quantified cognitive assessments. The severity of the decline, with mild impairment as the threshold for diagnosis, should be assessed as follows:

Mild: a degree of memory loss sufficient to interfere with everyday activities, though not so severe as to be incompatible with independent living. The main function affected is the learning of new material. For example, the individual has difficulty in registering, storing and recalling elements in daily living, such as where belongings have been put, social arrangements, or information recently imparted by family members.

Moderate: A degree of memory loss which represents a serious handicap to independent living. Only highly learned or very familiar material is retained. New information is retained only occasionally and very briefly. The individual is unable to recall basic information about where he lives, what he has recently been doing, or the names of familiar persons.

Severe: a degree of memory loss characterized by the complete inability to retain new information. Only fragments of previously learned information remain. The subject fails to recognise even close relatives.

(2) A decline in other cognitive abilities characterized by deterioration in judgement and thinking, such as planning and organizing, and in the general processing of information. Evidence for this should be obtained when possible from interviewing an informant, supplemented, if possible, by neuropsychological tests or quantified objective assessments. Deterioration from a previously higher level of performance should be established. The severity of the decline, with mild impairment as the threshold for diagnosis, should be assessed as follows:

Mild. The decline in cognitive abilities causes impaired performance in daily living, but not to a degree making the individual dependent on others. More complicated daily tasks or recreational activities cannot be undertaken.

Moderate. The decline in cognitive abilities makes the individual unable to function without the assistance of another in daily living, including shopping and handling money. Within the home, only simple chores are preserved. Activities are increasingly restricted and poorly sustained.

Severe. The decline is characterized by an absence, or virtual absence, of intelligible ideation. The overall severity of the dementia is best expressed as the level of decline in memory or other cognitive abilities, whichever is the more severe (e.g. mild decline in memory and moderate decline in cognitive abilities indicate a dementia of moderate severity).

G2. Preserved awareness of the environment during a period of time long enough to enable the unequivocal demonstration of G1. When there are superimposed episodes of delirium the diagnosis of dementia should be deferred.

G3. A decline in emotional control or motivation, or a change in social behaviour, manifest as at least one of the following:

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

G4. For a confident clinical diagnosis, G1 should have been present for at least six months; if the period since the manifest onset is shorter, the diagnosis can only be tentative.

Comments: The diagnosis is further supported by evidence of damage to other higher cortical functions, such as aphasia, agnosia, apraxia.

Judgment about independent living or the development of dependence (upon others) need to take account of the cultural expectation and context.

Dementia is specified here as having a minimum duration of six months to avoid confusion with reversible states with identical behavioural syndromes, such as traumatic subdural haemorrhage (S06.5), normal pressure hydrocephalus (G91.2) and diffuse or focal brain injury (S06.2 and S06.3).

A fifth character may be used to indicate the presence of additional symptoms, in the categories F00-F03 (F00 Dementia in Alzheimer's disease; F01 Vascular dementia; F02 Dementia in diseases classified elsewhere; and F03 Unspecified dementia), as follows:

- .x0 without additional symptoms
- .x1 with other symptoms, predominantly delusional
- .x2 with other symptoms, predominantly hallucinatory
- .x3 with other symptoms, predominantly depressive
- .x4 with other mixed symptoms

A sixth character may be used to indicate the severity of the dementia:

- .xx0 mild
- .xx1 moderate
- .xx2 severe

As mentioned above the overall severity of the dementia depends on the level of memory or intellectual impairment, whichever is the more severe.

F00 DEMENTIA IN ALZHEIMER'S DISEASE

A. The general criteria for dementia (G1 to G4) must be met.

B. There is no evidence from the history, physical examination or special investigations for any other possible cause of dementia (e.g. cerebrovascular disease, Parkinson's disease, Huntington's disease, normal pressure hydrocephalus), a systemic disorder (e.g. hypothyroidism, vit. B12 or folic acid deficiency, hypercalcaemia), or alcohol- or drug-abuse.

Comments: The diagnosis is confirmed by post mortem evidence of neurofibrillary tangles and neuritic plaques in excess of those found in normal ageing of the brain.

The following features support the diagnosis, but are not necessary elements: Involvement of cortical functions as evidenced by aphasia, agnosia or apraxia; decrease of motivation and drive, leading to apathy and lack of spontaneity; irritability and disinhibition of social behaviour; evidence from special investigations that there is cerebral atrophy, particularly if this can be shown to be increasing over time. In severe cases there may be Parkinson-like extrapyramidal changes, logoclonia, and epileptic fits.

Specification of features for possible subtypes. Because of the possibility that subtypes exist, it is recommended that the following characteristics be ascertained as a basis for a further classification: age at onset; rate of progression; the configuration of the clinical features, particularly the relative prominence (or lack) of temporal, parietal or frontal lobe signs; any neuropathological or neurochemical abnormalities, and their pattern.

The division of AD into subtypes can at present be accomplished in two ways: first by taking only the age of onset and labelling AD as either early or late, with an approximate cut-off point at 65 years; or secondly, by assessing how well the individual conforms to one of the two putative syndromes, early or late onset type. It should be noted that it is unlikely that a sharp distinction exists between early and late onset type. Early onset type may occur in late life, just as late onset type may occasionally have an onset under the age of 65.

The following criteria may be used to differentiate F00.0 from F00.1, but it should be remembered that the status of this subdivision is still controversial.

F00.0 Dementia in Alzheimer's disease with early onset

1. The criteria for dementia in Alzheimer's disease (F00) must be met, and the age at onset being under 65 years.

2. In addition, at least one of the following requirements must be met:

(a) evidence of a relatively rapid onset and progression;

(b) in addition to memory impairment, there is aphasia (amnesic or sensory), agraphia, alexia, acalculia, or apraxia (indicating the presence of temporal, parietal and/or frontal lobe involvement).

F00.1 Dementia in Alzheimer's disease with late onset

1. The criteria for dementia in Alzheimer's disease (F00) must be met and the age at onset must be 65 or more.

2. In addition, at least one of the following requirements must be met:

(a) evidence of a very slow, gradual onset and progression (the rate of the latter may be known only retrospectively after a course of 3 years or more);

(b) predominance of memory impairment G1.1, over intellectual impairment G1.2 (see general criteria for dementia).

F00.2 Dementia in Alzheimer's disease, atypical or mixed type

Use this term and code for dementias that have important atypical features or that fulfil criteria for both early and late onset type of Alzheimer's disease. Mixed Alzheimer's and vascular dementia is also included here.

F00.9 Dementia in Alzheimer's disease, unspecified

F01 VASCULAR DEMENTIA

G1. The general criteria for dementia (G1 to G4) must be met.

G2. Unequal distribution of deficits in higher cognitive functions, with some affected and others relatively spared. Thus memory may be quite markedly affected while thinking, reasoning and information processing may show only mild decline.

G3. There is clinical evidence of focal brain damage, manifest as at least one of the following:

(1) unilateral spastic weakness of the limbs;

(2) unilaterally increased tendon reflexes;

(3) an extensor plantar response;

(4) pseudobulbar palsy.

G4. There is evidence from the history, examination, or tests, of a significant cerebrovascular disease, which may reasonably be judged to be etiologically related to the dementia (e.g. a history of stroke; evidence of cerebral infarction).

The following criteria may be used to differentiate subtypes of vascular dementia, but it should be remembered that the usefulness of this subdivision may not be generally accepted.

F01.0 Vascular dementia of acute onset

A. The general criteria for vascular dementia (F01) must be met.

B. The dementia develops rapidly (i.e. usually within one month, but within no longer than three months) after a succession of strokes, or (rarely) after a single large infarction.

F01.1 Multi-infarct dementia

A. The general criteria for vascular dementia (F01) must be met.

B. The onset of the dementia is gradual (i.e. within three to six months), following a number of minor ischaemic episodes.

Comments: It is presumed that there is an accumulation of infarcts in the cerebral parenchyma. Between the ischaemic episodes there may be periods of actual clinical improvement.

F01.2 Subcortical vascular dementia

A. The general criteria for vascular dementia (F01) must be met.

B. A history of hypertension.

C. Evidence from clinical examination and special investigations of vascular disease located in the deep white matter of the cerebral hemispheres, with preservation of the cerebral cortex.

F01.3 Mixed cortical and subcortical vascular dementia

Mixed cortical and subcortical components of the vascular dementia may be suspected from the clinical features, the results of investigations (including autopsy), or both.

F01.8 Other vascular dementia

F01.9 Vascular dementia, unspecified

F02 DEMENTIA IN OTHER DISEASES CLASSIFIED ELSEWHERE

F02.0 Dementia in Pick's disease

A. The general criteria for dementia (G1 to G4) must be met.

B. Slow onset with steady deterioration.

C. Predominance of frontal lobe involvement evidenced by two or more of the following:

- (1) emotional blunting;
- (2) coarsening of social behaviour;
- (3) disinhibition;
- (4) apathy or restlessness;
- (5) aphasia.

D. Relative preservation, in the early stages, of memory and parietal lobe functions.

F02.1 Dementia in Creutzfeldt-Jakob disease

A. The general criteria for dementia (G1 to G4) must be met.

B. Very rapid progression of the dementia, with disintegration of virtually all higher cerebral functions.

C. The emergence, usually after or simultaneously with the dementia, of one or more of the following types of neurological symptoms and signs:

- (1) pyramidal symptoms;
- (2) extrapyramidal symptoms;
- (3) cerebellar symptoms;
- (4) aphasia;
- (5) visual impairment.

Comments: An akinetic and mute state is the typical terminal stage. An amyotrophic variant may be seen, where the neurological signs precede the onset of the dementia. A characteristic electroencephalogram (periodic spikes against a slow and low voltage background), if present in association with the above clinical signs, will increase the probability of the diagnosis. However, the diagnosis can be confirmed only by neuropathological examination (neuronal loss, astrocytosis, and spongiform changes). Because of the risk of infection, this should be carried out only under special protective conditions.

F02.2 Dementia in Huntington's disease

A. The general criteria for dementia (G1 to G4) must be met.

B. Subcortical functions are affected first and dominate the picture of dementia throughout; manifest as slowness of thinking or movement and personality alteration with apathy or depression.

C. Presence of involuntary choreiform movements, typically of the face, hands or shoulders, or in the gait. The patient may attempt to conceal them by converting them into a voluntary action.

D. A history of Huntington's disease in one parent or a sibling; or a family history which suggests the disorder.

E. The absence of clinical features otherwise accounting for the abnormal movements.

Comments: In addition to involuntary choreiform movements there may be development of extrapyramidal rigidity or spasticity with pyramidal signs.

F02.3 Dementia in Parkinson's disease

A. The general criteria for dementia (G1 to G4) must be met.

B. Diagnosis of Parkinson's disease.

C. Absence of cognitive impairment attributable to anti-parkinsonian medication.

D. There is no evidence from the history, physical examination or special investigations for any other possible cause of dementia, including other forms of brain disease, damage or dysfunction (e.g. cerebrovascular disease, HIV disease, Huntington's disease, normal pressure hydrocephalus), a systemic disorder (e.g. hypothyroidism, vit. B12 or folic acid deficiency, hypercalcaemia), or alcohol or drug abuse.

If criteria are also fulfilled for dementia in Alzheimer's disease with late onset (F00.1), this category F00.1 should be used in combination with Parkinson's disease G20.

F02.4 Dementia in human immunodeficiency (HIV) disease

A. The general criteria for dementia (G1 to G4) must be met.

B. Diagnosis of HIV infection.

C. There is no evidence from the history, physical examination or special investigations for any other possible cause of dementia, including other forms of brain disease, damage or dysfunction (e.g. Alzheimer's disease, cerebrovascular disease, Parkinson's disease, Huntington's disease, normal pressure hydrocephalus), a systemic disorder (e.g. hypothyroidism, vit. B12 or folic acid deficiency, hypercalcaemia), or alcohol or drug abuse.

F02.8 Dementia in other specified diseases classified elsewhere

Dementia can occur as a manifestation or consequence of a variety of cerebral and somatic conditions. To specify the etiology, the ICD-10 code for the underlying condition should be added.

F03 UNSPECIFIED DEMENTIA

This category should be used when the general criteria for dementia are met, but when it is not possible to identify one of the specific types (F00.0-F02.9).

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

Dementia Codes

Dementia of the Alzheimer's Type, with early onset

294.10 Without behavioural disturbance

294.11 With behavioural disturbance

Dementia of the Alzheimer's Type, with late onset

294.10 Without behavioural disturbance

294.11 With behavioural disturbance

Vascular dementia

290.40 Uncomplicated

290.41 With delirium

290.42 With delusions

290.43 With depressed mood

Dementia due to HIV disease

294.10 Without behavioural disturbance

294.11 With behavioural disturbance

Dementia due to head trauma

294.10 Without behavioural disturbance

294.11 With behavioural disturbance

Dementia due to Parkinson's disease

294.10 Without behavioural disturbance

294.11 With behavioural disturbance

Dementia due to Huntington's disease

294.10 Without behavioural disturbance

294.11 With behavioural disturbance

Dementia due to Pick's disease

294.10 Without behavioural disturbance

294.11 With behavioural disturbance

Dementia due to Creutzfeldt-Jakob Disease

294.10 Without behavioural disturbance

294.11 With behavioural disturbance

Dementia due to... [indicate other general medical condition]

294.10 Without behavioural disturbance

294.11 With behavioural disturbance

294.8 Dementia NOS

Appendix 2. Twenty-six item IQCODE

Instructions: Now we want you to remember what your friend or relative was like 10 years ago and to compare it with what he/she is like now. 10 years ago was 19^{xx}. On the next page are situations where this person has to use his/her memory or intelligence and we want you to indicate whether this has improved, stayed the same or got worse than in that situation over the past 10 years. Note the importance of comparing his/her present performance with 10 years ago. So if 10 years ago this person always forgot where he/she had left things and he/she still does this, then this would be considered 'Not much change'. Please indicate the changes you have observed by circling the appropriate answer.

		1	2	3	4	5
1	Remembering the names of family and friends	Much improved	A bit improved	Not much change	A bit worse	Much worse
2	Remembering the faces of family and friends	Much improved	A bit improved	Not much change	A bit worse	Much worse
3	Remembering things about family and friends, e.g. occupations, birthdays, addresses	Much improved	A bit improved	Not much change	A bit worse	Much worse
4	Remembering things that have happened recently	Much improved	A bit improved	Not much change	A bit worse	Much worse
5	Recalling conversations a few days later	Much improved	A bit improved	Not much change	A bit worse	Much worse
6	Forgetting what he / she wanted to say in the middle of a conversation	Much improved	A bit improved	Not much change	A bit worse	Much worse
7	Remembering her/his address and telephone number	Much improved	A bit improved	Not much change	A bit worse	Much worse
8	Remembering what day and month it is	Much improved	A bit improved	Not much change	A bit worse	Much worse
9	Remembering where things are usually kept	Much improved	A bit improved	Not much change	A bit worse	Much worse
10	Remembering where to find things which have been put in a different place from usual	Much improved	A bit improved	Not much change	A bit worse	Much worse

(Continued)

11	Adjusting to any change in his / her day to day routine	Much improved	A bit improved	Not much change	A bit worse	Much worse
12	Knowing how to work familiar machines around the house	Much improved	A bit improved	Not much change	A bit worse	Much worse
13	Learning to use a new gadget or machine around the house	Much improved	A bit improved	Not much change	A bit worse	Much worse
14	Learning new things in general	Much improved	A bit improved	Not much change	A bit worse	Much worse
15	Remembering things that happened to him/her when he/she was young	Much improved	A bit improved	Not much change	A bit worse	Much worse
16	Remembering things that he/she learned when he/she was young	Much improved	A bit improved	Not much change	A bit worse	Much worse
17	Understanding the meaning of unusual words	Much improved	A bit improved	Not much change	A bit worse	Much worse
18	Understanding magazine or newspaper articles	Much improved	A bit improved	Not much change	A bit worse	Much worse
19	Following a story in a book or on TV	Much improved	A bit improved	Not much change	A bit worse	Much worse
20	Composing a letter to friends or for business purposes	Much improved	A bit improved	Not much change	A bit worse	Much worse
21	knowing about important historical events if the past	Much improved	A bit improved	Not much change	A bit worse	Much worse

(Continued)

22	Making decisions on everyday matters	Much improved	A bit improved	Not much change	A bit worse	Much worse
23	Handling money for shopping	Much improved	A bit improved	Not much change	A bit worse	Much worse
24	Handling financial matters, e.g. the pension, dealing with the bank	Much improved	A bit improved	Not much change	A bit worse	Much worse
25	Handling other everyday arithmetic problems, e.g. knowing how much food to buy, knowing how long between visits from family or friends	Much improved	A bit improved	Not much change	A bit worse	Much worse
26	Using his/her intelligence to understand what's going on and to reason things through	Much improved	A bit improved	Not much change	A bit worse	Much worse

Appendix 3. Sixteen-item IQCODE

Instructions: Now we want you to remember what your friend or relative was like 10 years ago and to compare it with what he/she is like now. 10 years ago was 1977. On the next page are situations where this person has to use his/her memory or intelligence and we want you to indicate whether this has improved, stayed the same or got worse than in that situation over the past 10 years. Note the importance of comparing his/her present performance with 10 years ago. So if 10 years ago this person always forgot where he/she had left things and he/she still does this, then this would be considered 'Not much change'. Please indicate the changes you have observed by circling the appropriate answer.

		1	2	3	4	5
1	Remembering things about family and friends, e.g. occupations, birthdays, addresses	Much improved	A bit improved	Not much change	A bit worse	Much worse

(Continued)

2	Remembering things that have happened recently	Much improved	A bit improved	Not much change	A bit worse	Much worse
3	Recalling conversations a few days later	Much improved	A bit improved	Not much change	A bit worse	Much worse
4	Remembering her/his address and telephone number	Much improved	A bit improved	Not much change	A bit worse	Much worse
5	Remembering what day and month it is	Much improved	A bit improved	Not much change	A bit worse	Much worse
6	Remembering where things are usually kept	Much improved	A bit improved	Not much change	A bit worse	Much worse
7	Remembering where to find things which have been put in a different place from usual	Much improved	A bit improved	Not much change	A bit worse	Much worse
8	Knowing how to work familiar machines around the house	Much improved	A bit improved	Not much change	A bit worse	Much worse
9	Learning to use a new gadget or machine around the house	Much improved	A bit improved	Not much change	A bit worse	Much worse
10	Learning new things in general	Much improved	A bit improved	Not much change	A bit worse	Much worse
11	Following a story in a book or on TV	Much improved	A bit improved	Not much change	A bit worse	Much worse
12	Making decisions on everyday matters	Much improved	A bit improved	Not much change	A bit worse	Much worse

(Continued)

13	Handling money for shopping	Much improved	A bit improved	Not much change	A bit worse	Much worse
14	Handling financial matters, e.g. the pension, dealing with the bank	Much improved	A bit improved	Not much change	A bit worse	Much worse
15	Handling other everyday arithmetic problems, e.g. knowing how much food to buy, knowing how long between visits from family or friends	Much improved	A bit improved	Not much change	A bit worse	Much worse
16	Using his/her intelligence to understand what's going on and to reason things through	Much improved	A bit improved	Not much change	A bit worse	Much worse

Appendix 4. Commonly used cognitive assessments / screening tools

TEST	Cochrane DTA review in process
Mini-mental state examination (MMSE)	YES
GPcog	Still available
Minicog	Yes
Memory Impairment Screen (MIS)	Still available
Abbreviated mental testing	Still available
Clock drawing tests (CDT)	Still available
Montreal Cognitive Assessment (MoCA)	YES
AD-8 (informant interview)	YES

Appendix 5. Search strategies

Source	Search strategy	Hits retrieved
1. Medline In-process and other non-indexed citations and MEDLINE 1950-present (Ovid SP)	<ol style="list-style-type: none"> 1. IQCODE.ti,ab. 2. “informant questionnaire on cognitive decline in the elderly”.ti,ab 3. “IQ code”.ti,ab. 4. (“informant* questionnair*” adj3 (dement* or screening)).ti,ab 5. (“screening test*” adj2 (dement* or alzheimer*)).ti,ab. 6. or/1-5 	<p>Apr 2011: 291 Jul 2012: 39 Jan 2013: 19</p>
2. Embase 1980-2013 January 10 (Ovid SP)	<ol style="list-style-type: none"> 1. IQCODE.ti,ab. 2. “informant questionnaire on cognitive decline in the elderly”.ti,ab 3. “IQ code”.ti,ab. 4. (“informant* questionnair*” adj3 (dement* or screening)).ti,ab 5. (“screening test*” adj2 (dement* or alzheimer*)).ti,ab. 6. or/1-5 	<p>Apr 2011: 356 Jul 2012: 49 Jan 2013: 44</p>
3. PsycINFO 1806-January week 2 2013 (Ovid SP)	<ol style="list-style-type: none"> 1. IQCODE.ti,ab. 2. “informant questionnaire on cognitive decline in the elderly”.ti,ab 3. “IQ code”.ti,ab. 4. (“informant* questionnair*” adj3 (dement* or screening)).ti,ab 5. (“screening test*” adj2 (dement* or alzheimer*)).ti,ab. 6. or/1-5 	<p>Apr 2011: 215 Jul 2012: 28 Jan 2013: 17</p>
4. BIOSIS Previews 1926 to present (Thomson Reuters Web of Science)	<p>Topic=(IQCODE OR “informant questionnaire on cognitive decline in the elderly” OR “IQ code”) AND Topic=(dement* OR alzheimer* OR FTLD OR FTD OR “primary progressive aphasia” OR “progressive non-fluent aphasia” OR “frontotemporal lobar degeneration” OR “frontolobar degeneration” OR “frontal lobar degeneration” OR “pick* disease” OR “lewy bod*”)</p> <p>Timespan=All Years. Databases=SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH, BKCI-S, BKCI-SSH Lemmatization=On</p>	<p>Apr 2011: 84 Jul 2012: 12 Jan 2013: 2</p>

(Continued)

<p>5. Web of Science Core Collection (includes Conference Proceedings Citation Index) 1945-present (Thomson Reuters Web of Science)</p>	<p>Topic=(IQCODE OR “informant questionnaire on cognitive decline in the elderly” OR “IQ code”) AND Topic=(dement* OR alzheimer* OR FTLD OR FTD OR “primary progressive aphasia” OR “progressive non-fluent aphasia” OR “frontotemporal lobar degeneration” OR “frontolobar degeneration” OR “frontal lobar degeneration” OR “pick* disease” OR “lewy bod*”) Timespan=All Years. Databases=SCI-EXPANDED, SSCI, A&HCI, CPCIS, CPCIS-SH, BKCI-S, BKCI-SSH Lemmatization=On</p>	<p>Apr 2011: 184 Jul 2012: 24 Jan 2013: 13</p>
<p>6. LILACS (BIREME)</p>	<p>“short-IQCODE” OR IQCODE OR “IQ code” OR “Informant Questionnaire” OR “Informant Questionnaires”</p>	<p>Apr 2011: 10 Jul 2012: 0 Jan 2013: 0</p>
<p>7. CINAHL (EBSCOhost)</p>	<p>S1 TX IQCODE S2 TX “informant questionnaire” S3 TX “IQ code” S4 TX screening instrument S5 S1 or S2 or S3 or S4 S6 (MM “Dementia+”) S7 TX dement* S8 TX alzheimer* S9 S6 or S7 or S8 S10 S5 and S9</p>	<p>Apr 2011: 231 Jul 2012: 53 Jan 2013: 12</p>
<p>8. Additional other review sources: MEDION database (searched 31 Jan for all dates); Database of Abstracts of Reviews of Effects (searched Issue 1 of the Cochrane Library 2013); Health Technology Assessment Database (searched Issue 1 of the Cochrane Library 2013); ARIF: Aggressive Research Intelligence Facility www.arif.bham.ac.uk (searched 31 Jan for all dates)</p>	<p>Jan 2013: 3</p>	
<p>9 ALOIS (see Appendix 6 for the Medline strategy used to populate ALOIS)</p>	<p>Jan 2013: 22</p>	
<p>TOTAL before de-duplication of search results</p>	<p>Apr 2011: 1361 Jul 2012: 215 Jan 2013: 107 (+3 from additional review sources) TOTAL: 1708</p>	

(Continued)

TOTAL after de-duplication and first-assess by the Trials Search Co-ordinator	71	
---	----	--

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

Search strategy (MEDLINE Ovid SP) run for specialised register (ALOIS)

Search narrative: The searches detailed above are very simple, single concept strategies based on the index test (IQCODE). This is a sensitive approach to take. More complex and developed searches are run each month for the dementia group.

Every month the following strategy is run in Medline (via Ovid SP), with similar strategies run in Embase (via Ovid SP) and PsycINFO (via Ovid SP). The results are screened based on a reading of title and abstract. The full texts (where there is one) are then obtained and a few key details about each study are extracted including Index test/s and details of population and setting. For this review it was expected that most studies would be identified through a search of multiple sources based on one concept (the index test in question). However, we felt it was worth also searching ALOIS for any studies which had evaluated the accuracy of IQCODE but had not referred to it in the title or abstract of the reference.

MEDLINE In-process and other non-indexed citations and
MEDLINE 1950-present (Ovid SP)

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

48. "syndrome kurztest".ti,ab.
49. "six item screener".ti,ab.
50. "short memory questionnaire".ti,ab.
51. ("short memory questionnaire" and "SMQ").ti,ab.
52. "short orientation memory concentration test".ti,ab.
53. "s-omc".ti,ab.
54. "short blessed test".ti,ab.
55. "short portable mental status questionnaire".ti,ab.
56. "spmsq".ti,ab.
57. "short test of mental status".ti,ab.
58. "telephone interview of cognitive status modified".ti,ab.
59. "tics-m".ti,ab.
60. "trail making test".ti,ab.
61. "verbal fluency categories".ti,ab.
62. "WORLD test".ti,ab.
63. "general practitioner assessment of cognition".ti,ab.
64. "GPCOG".ti,ab.
65. "Hopkins verbal learning test".ti,ab.
66. "HVL".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 bod*" or DLB or LBD).ti,ab.
81. "cognit* impair*".ti,ab.
82. (cognit* adj4 (disorder* or declin* or fail* or function*)).ti,ab.
83. (memory adj3 (complain* or declin* or function*)).ti,ab.
84. or/75-83
85. exp "sensitivity and specificity"/
86. "reproducibility of results"/
87. (predict* adj3 (dement* or AD or alzheimer*)).ti,ab.
88. (identif* adj3 (dement* or AD or alzheimer*)).ti,ab.
89. (discriminat* adj3 (dement* or AD or alzheimer*)).ti,ab.
90. (distinguish* adj3 (dement* or AD or alzheimer*)).ti,ab.
91. (differenti* adj3 (dement* or AD or alzheimer*)).ti,ab.
92. diagnos*.ti.
93. di.fs.
94. sensitivit*.ab.
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
 124. 74 or 123
 125. 110 and 124
 126. 74 or 123
 127. 84 and 106 and 126
 128. 74 and 106
 129. 125 or 127 or 128
 130. (animals not (humans and animals)).sh.
 131. 129 not 130
 The concepts for this are:
A Specific neuropsychological tests
B General terms (both free text and MeSH) for tests/testing/
 screening
C Outcome: dementia diagnosis (unfocused MeSH with diagnos-
 tic sub-headings)
D Condition of interest: Dementia (general dementia terms both
 free text and MeSH - exploded and unfocused)
E Methodological filter: not used to limit all search
 The concept combinations are:
 1. (A OR B) AND C

- 2. (A OR B) AND D AND E
- 3. A AND E

Appendix 7. Assessment of reporting quality - STARD checklist

Section and Topic		
TITLE/ABSTRACT KEYWORDS	1	Identify the article as a study of diagnostic accuracy (recommend MeSH heading 'sensitivity and specificity')
INTRODUCTION	2	State the research questions or study aims, such as estimating diagnostic accuracy or comparing accuracy between tests or across participant groups
METHODS		
<i>Participants</i>	3	The study population: The inclusion and exclusion criteria, setting and locations where data were collected
	4	Participant recruitment: Was recruitment based on presenting symptoms, results from previous tests, or the fact that the participants had received the index tests or the reference standard?
	5	Participant sampling: Was the study population a consecutive series of participants defined by the selection criteria in item 3 and 4? If not, specify how participants were further selected
	6	Data collection: Was data collection planned before the index test and reference standard were performed (prospective study) or after (retrospective study)?
<i>Test methods</i>	7	The reference standard and its rationale.
	8	Technical specifications of material and methods involved including how and when measurements were taken, and/or cite references for index tests and reference standard
	9	Definition of and rationale for the units, cut-offs and/or categories of the results of the index tests and the reference standard
	10	The number, training and expertise of the persons executing and reading the index tests and the reference standard
	11	Whether or not the readers of the index tests and reference standard were blind (masked) to the results of the other test and describe any other clinical information available to the readers
<i>Statistical methods</i>	12	Methods for calculating or comparing measures of diagnostic accuracy, and the statistical methods used to quantify uncertainty (e.g. 95% confidence intervals)

(Continued)

	13	Methods for calculating test reproducibility, if done.
RESULTS		
<i>Participants</i>	14	When study was performed, including beginning and end dates of recruitment
	15	Clinical and demographic characteristics of the study population (at least information on age, gender, spectrum of presenting symptoms)
	16	The number of participants satisfying the criteria for inclusion who did or did not undergo the index tests and/or the reference standard; describe why participants failed to undergo either test (a flow diagram is strongly recommended)
<i>Test results</i>	17	Time-interval between the index tests and the reference standard, and any treatment administered in between
	18	Distribution of severity of disease (define criteria) in those with the target condition; other diagnoses in participants without the target condition
	19	A cross tabulation of the results of the index tests (including indeterminate and missing results) by the results of the reference standard; for continuous results, the distribution of the test results by the results of the reference standard
	20	Any adverse events from performing the index tests or the reference standard
<i>Estimates</i>	21	Estimates of diagnostic accuracy and measures of statistical uncertainty (e.g. 95% confidence intervals)
	22	How indeterminate results, missing data and outliers of the index tests were handled
	23	Estimates of variability of diagnostic accuracy between subgroups of participants, readers or centres, if done
	24	Estimates of test reproducibility, if done.
DISCUSSION	25	Discuss the clinical applicability of the study findings.

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

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

Appendix 9. Anchoring statements for quality assessment of IQCODE diagnostic studies

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

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

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

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

Anchoring statements to assist with assessment for risk of bias

Selection

Was a case-control or similar design avoided?

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

Was the sampling method appropriate?

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

Are exclusion criteria described and appropriate?

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

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

Index Test

Was IQCODE assessment performed without knowledge of clinical dementia diagnosis?

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

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

Were IQCODE thresholds prespecified?

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

Were sufficient data on IQCODE application given for the test to be repeated in an independent study?

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

Reference Standard

Is the assessment used for clinical diagnosis of dementia acceptable?

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

Was clinical assessment for dementia performed without knowledge of IQCODE?

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

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

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

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

Flow

Was there an appropriate interval between IQCODE and clinical dementia assessment?

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

Did all get the same assessment for dementia regardless of IQCODE result?

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

Were all who received IQCODE assessment included in the final analysis?

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

Were missing IQCODE results or un-interpretable IQCODE results reported?

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

Applicability

Were those included representative of the general population of interest?

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

Was IQCODE performed consistently and in a manner similar to its use in clinical practice?

IQCODE studies will be judged against the original description of its use.

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

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

Appendix 10. STARD (reporting quality) results

Study ID	STARD Item Assessment			
	Yes	No	Partially	Unclear
Flicker 1997	1, 3-5, 7, 9, 12, 15, 25	2, 11, 13, 14, 18, 20, 22, 24	8, 10, 16, 17, 19, 21, 23	6
Garcia 2002	4-7, 11, 12, 14, 15, 17, 19, 25	1, 9, 10, 13, 18, 20, 23, 24	2, 3, 8, 21	16, 22
Goncalves 2011	2-9, 12, 14, 17, 18, 21, 25	1, 11, 13, 16, 20, 22-24	10, 15, 19	-
Hancock 2009	1, 3-9, 11, 12, 14, 17, 19, 21, 25	10, 13, 16, 18, 20, 22-24	2, 15	-
Harwood 1997	1-5, 7, 9, 14, 16-18, 25	11-13, 20, 23, 24	8, 10, 15, 19, 21	6, 22
Jorm 1991	2-5, 7-10, 13, 17, 23-25	1, 11, 12, 14, 18-20, 22	15, 16, 21	6
Knaefelc 2003	2-4, 7, 10-12, 14, 15, 21, 22, 25	1, 5, 13, 16, 17, 20, 23, 24	8, 18, 19	6, 9
Mackinnon 1998	2, 4, 7-9, 11, 12, 15, 21, 25	1, 10, 13, 14, 16-18, 20, 22-24	3, 19	5, 6
Mulligan 1996	2, 7, 9-12, 15, 19, 22, 25	1, 4, 5, 13, 14, 17, 18, 20, 23, 24	3, 8, 21	6, 16
Narasimhalu 2008	1-4, 7, 9, 12, 16, 19, 23, 25	11, 13, 17, 20, 22, 24	5, 8, 10, 14, 15, 18, 21	6
Sikkes 2010	2-5, 7-10, 12, 14, 17, 19, 21-23, 25	1, 11, 13, 16, 18, 20, 24	15	6
Siri 2006	1, 2, 7, 9, 11, 12, 14, 15, 19, 25	10, 13, 16-18, 20, 22-24	3, 4, 8, 21	5, 6
Tang 2003	1-5, 7-12, 15-18, 22, 25	13, 14, 19, 20, 23, 24	21	6

CONTRIBUTIONS OF AUTHORS

JKH drafted initial manuscript, extracted data and performed analyses.

PF assisted with data extraction, quality assessment and analysis.

ANS assisted with search strategy, searching, and provided input to protocol and review.

DJS and RMcS provided supervision and input to protocol and review.

TJQ drafted protocol, assisted with searching, data extraction, quality assessment and analysis.

DECLARATIONS OF INTEREST

JKH - None known.

PF - None known.

ANS - None known.

DJS - None known.

RMcS - None known.

TJQ - None known.

SOURCES OF SUPPORT

Internal sources

- No sources of support supplied

External sources

- RCPSG Travelling Scholarship, UK.

RCPSG Travelling Scholarship supported TJQ in a period working with the Cochrane Dementia and Cognitive Improvement Group

- Graham Wilson Travel Fund, UK.

Graham Wilson Travel Fund supported TJQ to work with Test Accuracy group in University of Birmingham, UK

- The Orr Bequest, UK.

The Orr Bequest (University of Glasgow) has supported DJS and TJQ to pursue a body of work around cognitive assessment.

• This review was supported by the National Institute for Health Research, via a Cochrane Programme Grant to the Cochrane Dementia and Cognitive Improvement group. The views and opinions expressed therein are those of the authors and do not necessarily reflect those of the Systematic Reviews Programme, NIHR, NHS or the Department of Health, UK.

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

The review title was changed from *IQCODE for the diagnosis of Alzheimer's disease dementia and other dementias within a secondary care setting* to *Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE) for the diagnosis of dementia within a secondary care setting* for consistency with the other reviews in the suite.

Following discussions with DTA statisticians and in common with the analysis performed and reported in the IQCODE Community Review, we used the bivariate method for meta-analysis, rather than the HSROC, as stated in the original protocol.

Analysis looking at diagnostic accuracy stratified by dementia subtype and the effect of diagnostic criteria were not possible due to the lack of data in included studies.

We amended the methodology adopted for conducting sensitivity analysis by age due to the availability of data. The protocol had stated that where studies had more than 20% of included participants younger than 65 we would consider them potentially unrepresentative and would analyse them separately. In the event, the included studies reported median or mean age, and we decided that where this was 70 years or less, we would consider them potentially unrepresentative and would analyse them separately.