Cochrane Library

Cochrane Database of Systematic Reviews

Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE) for the early diagnosis of dementia across a variety of healthcare settings (Protocol)

Lees RA, Stott DJ, McShane R, Noel-Storr AH, Quinn TJ

Lees RA, Stott DJ, McShane R, Noel-Storr AH, Quinn TJ.

Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE) for the early diagnosis of dementia across a variety of healthcare settings.

Cochrane Database of Systematic Reviews 2014, Issue 10. Art. No.: CD011333.

DOI: 10.1002/14651858.CD011333.

www.cochranelibrary.com

TABLE OF CONTENTS

HEADER 1
ABSTRACT 1
BACKGROUND 1
OBJECTIVES 4
METHODS 4
REFERENCES 7
APPENDICES 8
DECLARATIONS OF INTEREST 19

[Diagnostic Test Accuracy Protocol]

Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE) for the early diagnosis of dementia across a variety of healthcare settings

Rosalind A Lees[1], David J Stott[2], Rupert McShane[3], Anna H Noel-Storr[3], Terry J Quinn[4]

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

Contact address: Terry J Quinn, Cardiovascular and Medical Sciences, University of Glasgow, Walton Building, Glasgow Royal Infirmary, Glasgow, G4 0SF, UK. Terry.Quinn@glasgow.ac.uk.

Editorial group: Cochrane Dementia and Cognitive Improvement Group.
Publication status and date: New, published in Issue 10, 2014.

Citation: Lees RA, Stott DJ, McShane R, Noel-Storr AH, Quinn TJ. Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE) for the early diagnosis of dementia across a variety of healthcare settings. *Cochrane Database of Systematic Reviews* 2014, Issue 10. Art. No.: CD011333. DOI: 10.1002/14651858.CD011333.

ABSTRACT

This is a protocol for a Cochrane Review (Diagnostic test accuracy). The objectives are as follows:

To determine the diagnostic accuracy of the informant based questionnaire IQCODE in a population free from dementia for the delayed diagnosis of dementia.

Where data are available, we will describe the following.

1. The delayed verification diagnostic accuracy of IQCODE at various thresholds. We recognise that various thresholds or 'cut-off' scores have been used to define IQCODE screen positive states, and thus various 'subthreshold' cut-points could be used to describe individuals with cognitive problems not diagnostic of dementia. We have not pre-specified IQCODE cut-points of interest, rather we will collect delayed verification test accuracy data for all cut-points described.

2. Effects of heterogeneity on the reported diagnostic accuracy of IQCODE for delayed verification dementia (see below).

Items of specific interest will include case-mix of population, IQCODE test format, time since index test and healthcare setting.

BACKGROUND

Dementia is a substantial and growing public health concern (Herbert 2013; Prince 2013). Depending on the case definition employed, contemporary estimates of dementia prevalence in the United States are in the range of 2.5 to 4.5 million individuals. Changes in population demographics will be accompanied by increases in global dementia incidence and prevalence. Although the magnitude of the increase in prevalent dementia is debated, there is no doubt that absolute numbers of older adults with dementia will increase substantially in the short to medium term future

(Ferri 2005).

A dementia diagnosis requires cognitive and functional decline. A syndrome of cognitive problems beyond those expected for age and education but not sufficient to impact on daily activities is also recognised. This possible intermediate state between normal cognitive ageing and pathological change is often labelled mild cognitive impairment (MCI) or cognitive impairment no dementia (CIND), although a variety of other terms are also used. For consistency we use the term MCI throughout this review. A proportion of individuals with MCI will develop a clinical dementia state over time (estimated at 10% to 15% of MCI individuals annually), while others will improve or remain stable. All definitions of this 'pre-dementia' state are based on key criteria of change in cognition (subjective or reported by an informant) with objective cognitive impairment but preserved functional ability.

A key element of effective management in dementia is early, robust diagnosis. Recent guidelines place emphasis on very early diagnosis to facilitate improved management and to allow informed discussions and planning with patients and carers (Cordell 2013). An early or unprompted assessment paradigm needs to distinguish early pathological change from normal states. Diagnosis of early dementia or MCI is especially challenging. It is important to recognise those who will progress to dementia as identification of this group may allow for targeted intervention, however at present there is no accepted method for determining prognosis.

The ideal would be expert, multidisciplinary assessment informed by various supplementary investigations (neuropsychology, neuroimaging or other biomarkers). This approach is only really feasible in a specialist memory service and is not suited to population screening or case-finding.

In practice a two-stage process is often employed, with initial 'triage' assessments that are suitable for use by non-specialists used to select those patients who require further detailed assessment (Boustani 2003). Various tools for initial cognitive screening have been described (Brodaty 2002; Folstein 1975; Galvin 2005). Regardless of the methods employed, there is scope for improvement as observational work suggests that many patients with dementia are not diagnosed (Chodosh 2004; Valcour 2000).

Screening assessment often takes the form of brief, direct cognitive testing. Such an approach will only provide a 'snapshot' of cognitive function. However, a defining feature of dementia is cognitive or neuropsychological change over time. 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. These informant based interviews aim to retrospectively assess change in function.

An instrument that is prevalent in research and clinical practice, particularly in Europe, is the Informant Questionnaire for Cognitive Decline in the Elderly (IQCODE) with questionnaire based interviews. This screening or triage tool is the focus of this review (Jorm 2004).

Traditional screening tools for cognitive problems have defined threshold scores that differentiate individuals likely to have dementia from those with no dementia. As dementia is a progressive, neurodegenerative disease, a population with cognitive problems will have a range of test scores. Individuals with MCI, or indeed early dementia, may have screening test scores that although not at a threshold suggestive of dementia are still abnormal for age. It seems plausible that a subthreshold score on a screening test such as IQCODE could be predictive of future dementia states and so could be used to target those individuals who may need follow up or further investigation. This paradigm of using a screening test with delayed verification of a dementia state is commonly employed in studies of the diagnostic properties of dementia 'biomarkers' but can equally be applied to direct or informant based assessment scales.

This review will focus on the use of the IQCODE in individuals without a firm clinical dementia diagnosis and will assess the accuracy for delayed verification of a dementia diagnosis after prospective follow up.

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.

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

Dementia diagnosis can be made according to various internationally accepted diagnostic criteria, with exemplars being the World Health Organization International Classification of Diseases (ICD) and American Psychiatric Association Diagnostic and Statistical Manual of Mental Disorders (DSM) for all cause dementia and subtypes. The label of dementia encompasses varying pathologies of which Alzheimer's disease is the most common. Diagnostic criteria are available for specific dementia subtypes, that is National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer's Disease and Related Disorders Association (NINCDS-ADRDA) criteria for Alzheimer's dementia (McKhann 1984; McKhann 2011); McKeith criteria for Lewy Body dementia (McKeith 2005); Lund criteria for frontotemporal

dementias (McKhann 2001); and the NINDS-AIREN criteria for vascular dementia (Roman 1993).

We are examining delayed verification of dementia and so will describe the properties of a standard screening tool (the IQCODE) for detection of problems earlier in the disease journey than frank dementia. A proportion of participants included in relevant studies are likely to have MCI, that is cognitive problems beyond those expected for age and education but not sufficient to impact on daily activities. The usual research definition of MCI is that described by Petersen (Peterson 2004); and various subtypes have been proposed within the rubric of MCI. We will collate information on MCI described using any validated criteria, however the focus of the review is not IQCODE for the contemporaneous diagnosis of MCI but rather IQCODE for a future diagnosis of dementia. These two constructs are related but not synonymous as only a proportion of individuals with MCI will develop dementia.

Index test(s)

Our index test will be the Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE) (Jorm 1988).

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. 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 the IQCODE is prevalent in both clinical practice and research. 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 (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; this modified IQCODE is common in clinical practice and has been recommended as the preferred IQCODE format (Jorm 2004). Further modifications to the IQCODE are described including fewer items and assessment over shorter time periods. Our analysis will include all versions of IQCODE but results for original and modified scales will not be pooled. In this review the term 'IQCODE' will refer to the original 26-item English language questionnaire as described by Jorm. Other versions of IQCODE will be described according to number of items and administration language (that is a 16-item IQCODE for Spanish speakers will be described as 'IQCODE-16 Spanish').

In the original IQCODE development and validation work normative data were described, with a total score of > 93 or an average score of > 3.31 indicative of cognitive impairment (Jorm 2004). There is no consensus on the optimal threshold and certainly no guidance on the use of subthreshold IQCODE scores for delayed verification. 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.

Clinical pathway

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

In secondary care settings, individuals will have been referred for expert input but not exclusively due to memory complaints. Opportunistic screening of adults presenting as unscheduled admissions to hospitals would be an exemplar secondary care pathway. The rubric of secondary care also includes those individuals referred to dementia and memory specific services. This population will have a high prevalence of cognitive disorders and mimics. More individuals will have had a greater degree of prior cognitive assessment than in other settings but cognitive testing is not always performed prior to memory service referral (Menon 2011).

In the general practice and primary care setting, the individual self presents to a non-specialist service because of subjective memory complaints. Previous cognitive testing is unlikely but prevalence will be reasonable high. Using IQCODE in this setting could be described as 'triage' or 'case-finding'. In the community setting, the cohort is largely unselected and the approach may be described as 'population screening'.

The IQCODE is not a diagnostic tool and the role of IQCODE in clinical practice is identifying those who may need further detailed assessment or follow up.

Alternative test(s)

Several other dementia screening and assessment tools have been described, for example Folstein's mini-mental state examination (Folstein 1975). These performance based measures for cognitive screening all rely on comparing single or multidomain cognitive testing against population-specific normative data.

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 changes in memory, problem solving, orientation and daily activities (Galvin 2005).

For this review we will focus on papers that describe IQCODE diagnostic properties, we will not consider other cognitive screening or assessment tools. Our IQCODE diagnostic test accuracy studies form part of a larger body of work by the Cochrane Dementia and Cognitive Improvement Group (Quinn 2014) describing test properties of all commonly used assessment tools (Appendix 1).

Rationale

There is no consensus on the optimal initial assessment for dementia and choice is currently dictated by experience with a particular instrument, time constraints and training. A better understanding of the diagnostic properties of various strategies would allow for an informed approach to testing. Critical evaluation of the evidence base for screening tests or other diagnostic markers is of major importance. Without a robust synthesis of the available information there is the risk that future research, clinical practice and policy will be built on erroneous assumptions about diagnostic validity.
This review will form part of a body of work describing the diagnostic properties of commonly used dementia tools. At present we are conducting single test reviews and meta-analyses. However, the intention is to then collate these data by performing an overview, allowing comparison of various test strategies.

OBJECTIVES

To determine the diagnostic accuracy of the informant based questionnaire IQCODE in a population free from dementia for the delayed diagnosis of dementia.

Secondary objectives

Where data are available, we will describe the following.
1. The delayed verification diagnostic accuracy of IQCODE at various thresholds. We recognise that various thresholds or 'cut-off' scores have been used to define IQCODE screen positive states, and thus various 'subthreshold' cut-points could be used to describe individuals with cognitive problems not diagnostic of dementia. We have not pre-specified IQCODE cut-points of interest, rather we will collect delayed verification test accuracy data for all cut-points described.
2. Effects of heterogeneity on the reported diagnostic accuracy of IQCODE for delayed verification dementia (see below).
Items of specific interest will include case-mix of population, IQCODE test format, time since index test and healthcare setting.

METHODS

Criteria for considering studies for this review

Types of studies

In this review we are looking at the properties of IQCODE for diagnosis of the dementia state on prospective follow up, that is investigating whether a subthreshold score on IQCODE in a population free of dementia at baseline assessment is associated with development of dementia over a period of follow up. The implication is that at the time of testing the individual had a cognitive problem sufficient to be picked up on screening but not yet meeting dementia diagnostic criteria. We will describe this paradigm as 'delayed verification' diagnostic test accuracy. IQCODE for contemporaneous diagnosis of dementia is covered by other Cochrane reviews.
We anticipate that the majority of studies will be performed in secondary care settings. We will include test studies performed in other healthcare settings and classify these as: primary care or community.
Case-control studies are known to potentially overestimate properties of a test and such studies will not be included.
Case studies or samples with very small numbers (for the purposes of this review chosen as 10 participants) will not be included but will be described in the table of excluded studies.
There may be cases where settings are mixed, for example a population study 'enriched' with additional cases from primary care. We will consider separate data for patients from each setting, if available. If these data are not available we will treat these studies as case-control studies and not include them in this review.

Participants

All adults (aged over 18 years) and with no formal diagnosis of dementia will be eligible.
We have not predefined exclusion criteria relating to the case-mix of the population studied but will assess this aspect of the study as part of our assessment of heterogeneity. Where there is concern that the participants are not representative, this will be explored at study level using the risk of bias assessment framework outlined below.

Index tests

Studies must include (not necessarily exclusively) IQCODE used as an informant questionnaire for delayed verification.
IQCODE has been translated into various languages to allow international administration. The properties of a translated IQCODE in a cohort of non-English speakers may differ from properties of the original English language questionnaire. We will collect data on the principle language used for IQCODE assessment.
For this review we will not consider other cognitive screening or assessment tools. Where a paper describes the IQCODE with an

in-study comparison against another screening tool, we will include the IQCODE data only. Where IQCODE is used in combination with another cognitive screening tool we will include the IQCODE data only.

Target conditions

Any clinical diagnosis of all cause (unspecified) dementia will be included. Defining a particular dementia subtype is not required although where available these data will be recorded.

Reference standards

Our reference standard will be 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.

The primary analysis will be for clinical diagnosis to include all cause (unspecified) dementia, using any recognised diagnostic criteria (for example ICD-10, DSM-IV). Dementia diagnosis may specify a pathological subtype and all common dementia subtypes will be included (examples are NINCDS-ADRDA, Lund-Manchester, McKeith, NINCDS-AIREN). We have not defined preferred diagnostic criteria for rarer forms of dementia (for example alcohol related, HIV related, prion disease related) and these will be considered under our rubric of 'all cause' dementia and not separately.

Clinicians may use imaging, pathology or other data to aid diagnosis, however diagnosis based only on these data without a corresponding clinical assessment will not be included. 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 (for example individual clinician versus algorithm versus consensus determination); data on the method and application of dementia diagnosis will be collected for each study and potential effects will be explored as part of our assessment of heterogeneity. Use of other (brief) direct performance tests in isolation will not be an acceptable method for diagnosis.

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

Search methods for identification of studies

We will use a variety of information sources to ensure all relevant studies are included. Terms for electronic database searching will be devised in conjunction with the team at the Cochrane Dementia and Cognitive Improvement Group. As this IQCODE review forms part of a suite of reviews looking at informant scales we have created a comprehensive search strategy designed to pick up all cognitive assessment scales, we will complement this generic search with searches specific to IQCODE terminology.

Electronic searches

We will search the specialised register of the Cochrane Dementia and Cognitive Improvement Group, ALOIS (which includes both intervention and diagnostic accuracy studies), MEDLINE (OvidSP), EMBASE (OvidSP), BIOSIS (ISI Web of Knowledge), Science Citation Index (ISI Web of Knowledge), PsycINFO (OvidSP), CINAHL (EBSCOhost) and LILACS (BIREME). See Appendix 2 and Appendix 3 for the strategy we will run in MEDLINE (OvidSP) along with a narrative describing how the strategy was developed and validated. Similarly structured search strategies will be designed using search terms appropriate for each database. MeSH terms and other controlled vocabulary will be used where appropriate.

We will also search sources specific to diagnostic accuracy or to systematic reviews:

- MEDION database (Meta-analyses van Diagnostisch Onderzoek at www.mediondatabase.nl);
- DARE (Database of Abstracts of Reviews of Effects at www.york.ac.uk/inst/crd/crddatabases.html);
- HTA Database (Health Technology Assessments Database in *The Cochrane Library*);
- ARIF database (Aggressive Research Intelligence Facility at www.arif.bham.ac.uk).

No language or date restrictions will be applied to the electronic searches. Translation services will be used as necessary.

Initial searches will be run by the Cochrane Dementia and Cognitive Impairment Group Search Co-ordinator.

Searching other resources

Grey literature and proceedings: chosen electronic databases include assessments of conference proceedings. We will aim to access theses or PhD abstracts from institutions known to be involved in prospective dementia studies.

Handsearching: we will not perform handsearching as there is little published evidence of the benefits of handsearching for diagnostic studies (Glanville 2012) .

Reference lists: we will check the reference lists of all relevant studies and reviews in the field for further possible titles and the process will be repeated until no new titles are found (Greenhalgh 1997).

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

Relevant studies will be used in PubMed to search for additional studies with the related article feature. Key studies will be examined in citation databases such as Science Citation Index and Scopus to ascertain any further relevant studies.

Data collection and analysis

Selection of studies

Two review authors will independently screen all titles generated by the electronic database searches for relevance. Abstracts of selected titles will be reviewed by the two review authors and all potentially eligible studies will be selected for full paper review. Two review authors will independently assess full manuscripts against the inclusion criteria. Disagreement will be resolved by discussion or by involving an arbitrator if necessary.

Where a study may include useable data but these are not presented in the published manuscript, we will contact the authors directly to request further information. If the same data are presented in more than one paper we will include the primary paper only.

The study selection process will be detailed in a PRISMA flow diagram.

Data extraction and management

Data will be extracted to a study-specific pro forma that includes clinical and demographic details of the participants, details of the setting, details of IQCODE administration, and details of the dementia diagnosis process.

Test accuracy data will be extracted to a standard two by two table. Data extraction will be performed independently by two blinded review authors. Disagreement in data extraction will be resolved by discussion, with the use of an arbitrator if necessary.

For each included paper, the flow of patients (numbers recruited, included, assessed) will be detailed in a flow diagram.

Assessment of methodological quality

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

Quality of study reporting will be assessed using the STARD checklist (Bossuyt 2003) (Appendix 4). If it becomes available during the course of the review we will use the proposed dementia-specific extension to the STARD tool, STARDdem (http://stard-dem.org/). STARD data will be tabulated and presented as an appendix to the review.

We will assess the methodological quality of each study using the QUADAS-2 tool (http://www.bris.ac.uk/quadas/quadas-2) (Appendix 5). This tool incorporates domains specific to patient selection, index test, reference standard and patient flow. Each domain is assessed for risk of bias and the first three domains are also assessed for applicability. Certain key areas that are important for quality assessment are participant selection, blinding and missing data. Following a group meeting of review authors we created guidance for the application of QUADAS-2 to dementia screening assessments, specifically developing anchoring statements for QUADAS based assessment that are suited to dementia test accuracy studies. This QUADAS guidance was created through a multidisciplinary working group and has been extensively piloted (Davis 2013). The process and resulting statements for assessment are described (Appendix 6).

QUADAS-2 data will not be used to form a summary quality score, rather there will be a narrative summary describing the numbers of studies that found high, low or unclear risk of bias or concerns regarding applicability with corresponding tabular and graphical displays.

Both assessments will be performed by paired independent raters who are blinded to each other's scores. Disagreement will be resolved by further review and discussion with recourse to a third party arbitrator where necessary.

Statistical analysis and data synthesis

We are interested in test accuracy of IQCODE for the delayed diagnosis of dementia using a dichotomous variable, 'dementia' or 'no dementia'. Thus, we will apply the current DTA framework for analysis of a single test and fit the extracted data to a standard two by two data table showing binary test results cross-classified with the binary reference standard. This process will be repeated for each IQCODE threshold score described in the source papers. We will repeat the process for each assessment if the reference standard is assessed at more than one follow up, as well as exploring the effect of time since index test in our assessment of heterogeneity.

We will use RevMan 5 to calculate sensitivity, specificity and 95% confidence intervals (CIs) from the two by two tables abstracted from the included studies. We will present individual study results graphically by plotting estimates of sensitivities and specificities as forest plots and in receiver operating characteristic (ROC) space.

To allow for pooled analysis, we will use software additional to RevMan (SAS release 9.1). As we expect a common threshold, we will use the bivariate approach in the first instance. We will describe summary metrics of sensitivity, specificity and positive and negative likelihood ratios all with corresponding 95% CIs. If data allow we will use the HSROC method to explore differing thresholds across studies.

We plan analysis across all studies; this will be for information only and we will be cautious in how we interpret these data. Final decisions on whether pooling data for meta-analysis is appropriate will be made by review author consensus.

The 'delayed verification' nature of the included studies adds a further level of complexity as a proportion of individuals recruited at baseline may be 'lost' to subsequent review. In the first instance we will apply the usual DTA framework ignoring any censoring that might have occurred. We acknowledge that such a reduction in the data may represent a significant oversimplification. We will therefore adopt an intention to diagnose (ITD) approach if data

allow. As a sensitivity analysis we will present what the result would be if all dropouts would have developed dementia and if all dropouts would not have developed dementia. We may also need to assume that the proportion of positive and negative test results is the same in the unknown as with the known participants in order to do this.

Investigations of heterogeneity

Heterogeneity is expected in DTA reviews and 'traditional' measures of heterogeneity that are used in meta-analysis are not appropriate to DTA reviews.

We will include IQCODE studies that span various settings. We will offer a narrative review of all studies. We will perform pooled analysis across all studies for information but primary analyses will be restricted to the various predefined healthcare settings.

The properties of a tool describe the behaviour of the instrument under particular circumstances. Thus, for our assessment of potential sources of heterogeneity (where data allow) we will collect data on the following.

1. Included patients (age and case mix).

In the first instance we will explore age, taking age over 65 years as a reference point. We suspect that the majority of included participants in eligible studies will be aged over 65 years. IQCODE may have different properties in younger cohorts and so we will look at age ranges within studies, and studies that have greater than 20% of included participants younger than 65 years will be graded as potentially unrepresentative and analysed separately.

We anticipate that most studies will be of unselected adults, however if the study if of a specific population, for example stroke survivors, these data will be pooled and analysed separately.

2. Clinical criteria used to reach dementia diagnosis.

We will record the classification used (for example ICD-10, DSM-IV) and the methodology used to reach dementia diagnosis (for example individual assessment, group (consensus) assessment).

3. Technical features of the testing strategy.

Our focus will be the language of assessment. In the first instance we will classify the assessments as English language and non-English language tests. Summary estimates will be compared for subgroups of interest: all language IQCODE and then English language IQCODE versus non-English language IQCODE.

4. Factors specific to the delayed verification analysis.

We will assess test accuracy at various follow-up time points if available. We will record any interventions administered during follow up that may influence the outcome (for example cholinesterase inhibitors).

Sensitivity analyses

Where appropriate (that is if not already explored in our analyses of heterogeneity), and as data allow, we will explore the sensitivity of any summary accuracy estimates to aspects of study quality such as nature of blinding and loss to follow up guided by the anchoring statements developed in our QUADAS-2 exercise. Primary analysis will include all eligible studies, sensitivity analysis will exclude studies of low quality (high likelihood of bias) to determine if the results are influenced by inclusion of the lower quality studies. Due to the potential for bias, we have pre-specified that case-control data will not be included.

Assessment of reporting bias

Reporting bias will not be investigated because of current uncertainty about how it operates in test accuracy studies and in the interpretation of existing analytical tools such as funnel plot.

REFERENCES

Additional references

Bossuyt 2003

Bossuyt PM, Reitsma JB, Bruns DE Gatsonis CA, Glasziou PP, Irwig LM, et al. for the STARD steering group. Towards complete and accurate reporting of studies of diagnostic accuracy: The STARD initiative. *BMJ* 2003;**326**: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.

Brodaty 2002

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

Chodosh 2004

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

Cordell 2013

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

Davis 2013

Davis DHJ, Creavin ST, Noel-Storr AH, 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* March 2013, Issue CD010460. [DOI: 10.1002/14651858.CD010460]

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

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

Galvin 2005
Galvin JE. The AD8: A brief informant interview to detect dementia. *Neurology* 2005;**65**(4):559–64.

Glanville 2012
Glanville J, Cikalo M, Crawford F, Dozier M, McIntosh H. Handsearching did not yield additional FDG-PET diagnostic test accuracy studies compared with electronic searches: a preliminary investigation. *Research Synthesis Methods* 2012;**3**:202–13.

Greenhalgh 1997
Greenhalgh T. Papers that report diagnostic or screening tests. *BMJ* 1997;**315**:540–3.

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

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

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

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.

Menon 2011
Menon R, Larner AJ. USe of cognitive screening instruments in primary care: the impact of national dementia directives (NICE/SCIE National Dementia Strategy). *Family Practice* 2011;**28**:272–6.

Peterson 2004
Petersen RC. Mild Cognitive Impairment as a diagnostic entity. *Journal of Internal Medicine* 2004;**256**:183–94.

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* April 2014, Issue CD010079.pub2. [DOI: 10.1002/14651858.CD010079.pub2]

Roman 1993
Roman GC, Tatemichi TK, Erkinjuntti T, et al. for the Vascular Dementia: Diagnostic Criteria for Research Studies work group. Report of the NINDS-AIREN International Workshop. *Neurology* 1993;**43**(2):250–60.

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.

* *Indicates the major publication for the study*

APPENDICES

Appendix 1. Commonly used cognitive assessment or screening tools

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

For each test, the planned review will encompass diagnostic test accuracy in community; primary and secondary care settings. As well as standard diagnosis, where applicable reviews will also describe delayed verification design trials.

Appendix 2. Search strategy for use with MEDLINE electronic database

MEDLINE In-process and other non-indexed citations and MEDLINE 1950 to present (OvidSP)	1. IQCODE.ti,ab. 2. "informant questionnaire on cognitive decline".ti,ab. 3. "Informant Questionnaire for Cognitive Decline in the Elderly".ti,ab 4. ("informant* questionnair*" adj3 (dement* or screening)).ti,ab 5. "informant* questionnair*".ti,ab. AND exp *Dementia/ 6. "screening test*".ti,ab. 7. (dement* or alzheimer* or "cognit* impair*").ti,ab. 8. exp Dementia/ 9. or/6,7 10. 5 AND 8 11. or/1-5 12. or/10,11

Appendix 3. Search strategy (MEDLINE OvidSP) run for specialised register (ALOIS)

MEDLINE In-process and other non-indexed citations and MEDLINE 1950 to present (OvidSP)	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. “HVLT”.ti,ab.
67. “time and change test”.ti,ab.
68. “modified world test”.ti,ab.
69. “symptoms of dementia screener”.ti,ab.
70. “dementia questionnaire”.ti,ab.
71. “7MS”.ti,ab.
72. (“concord informant dementia scale” or CIDS).ti,ab.
73. (SAPH or “dementia screening and perceived harm*”).ti,ab
74. or/1-73
75. exp Dementia/
76. Delirium, Dementia, Amnestic, Cognitive Disorders/
77. dement*.ti,ab.
78. alzheimer*.ti,ab.
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 (lines 1-73)

B General terms (both free text and MeSH) for tests/testing/screening (lines 111-122)

C Outcome: dementia diagnosis (unfocused MeSH with diagnostic subheadings) (lines 107-109)

D Condition of interest: Dementia (general dementia terms both free text and MeSH - exploded and unfocused) (75-83)

E Methodological filter: not used to limit all search (85-105)

The concept combinations are:

1. (A OR B) AND C

	2. (A OR B) AND D AND E 3. A AND E **Search strategy (MEDLINE OvidSP) run for specialised register (ALOIS)** Search narrative: The search in Appendix 2 is largely based on a single concept: 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 OvidSP). 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

Appendix 4. 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		
Participants	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)?
Test methods	7	The reference standard and its rationale.

(*Continued*)

	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
Statistical methods	12	Methods for calculating or comparing measures of diagnostic accuracy, and the statistical methods used to quantify uncertainty (e.g. 95% confidence intervals)
	13	Methods for calculating test reproducibility, if done.
RESULTS		
Participants	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)
Test results	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
Estimates	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

(*Continued*)

	24	Estimates of test reproducibility, if done.
DISCUSSION	25	Discuss the clinical applicability of the study findings.

Appendix 5. 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	Are there concerns that the target condition as defined by the reference standard does not match	

(*Continued*)

		question?	the review question?	

Appendix 6. 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 QUADAS2 has been operationlised for use with dementia reference stand rad studies below. In these descriptors dominant items are labelled as 'hIgh risk'.

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

Patient 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 of patients 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 will be assessed as a potential source of heterogeneity.

High risk of bias (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.

High risk of 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' 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. For a community sample we would expect relatively few exclusions.

Post hoc exclusions will be labelled 'high risk' of bias.

Low risk

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' 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.

High risk

Were IQCODE thresholds pre-specified?

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 not applicable.
Low risk
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.
Low risk

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 is not familiar to the review authors and the Cochrane Dementia and Cognitive Improvement group this item should be classified as 'high risk of bias'.
High risk
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' 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 pre-specified that dementia diagnosis that explicitly uses IQCDODE will be classified as high risk of bias.
High risk
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).
Low risk

Patient flow

Was there an appropriate interval between IQCODE and clinical dementia assessment.
For a study looking at delayed verification there is no agreement on how long the interval should be between index test and first/last assessment for dementia. An interval of less than six months is unlikely to be sufficient time for progression.
Low risk of bias
Did all patients get the same assessment for dementia regardless of IQCODE result?
There may be scenarios where only those patients who score 'test positive' on IQCODE have a more detailed assessment. Where dementia assessment (or other reference standard) differs between patients this should be classified as high risk of bias.
High risk of bias
Were all patients who received IQCODE assessment included in the final analysis?
If dropouts these should be accounted for; a maximum proportion of dropouts to remain low risk of bias has been specified as 20%.
Low risk of bias
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.
Low risk of bias

Applicability

Were included patients representative of the general population of interest?

The included patients should match the intended population as described in the review question. If not already specified in the review inclusion criteria, setting will be particularly important - the review authors should consider population in terms of symptoms; pre-testing; potential disease prevalence. Studies that use very selected patients 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 of patients with disease than usual clinical practice. In this instance the item should be rated poor applicability.

DECLARATIONS OF INTEREST

The authors have no relevant conflicts of interest or disclosures.