Received: 4 June 2018 | Accepted: 18 October 2018
DOI: 10.1002/gps.5016

REVIEW ARTICLE

WILEY International Journal of Geriatric Psychiatry

Cognitive tests for the detection of mild cognitive impairment (MCI), the prodromal stage of dementia: Meta-analysis of diagnostic accuracy studies

Alexandre Breton[1] | Daniel Casey[2] | Nikitas A. Arnaoutoglou[3,4]

[1] Medical School, Medical Sciences Division, University of Oxford, Oxford, UK

[2] Oxford Health NHS Foundation Trust, Oxford, UK

[3] Department of Psychiatry, University of Oxford, Oxford, UK

[4] Department of Psychiatry, University of Cambridge, Cambridge, UK

Correspondence
Dr Nikitas A. Arnaoutoglou, Department of Psychiatry, University of Cambridge, Level E4, Cambridge Biomedical Campus, Hills Road, Cambridge CB2 0QQ, UK.
Email: na488@cam.ac.uk

Funding information
Trinity College, University of Oxford

Introduction: Mild cognitive impairment (MCI) is regarded as a prodrome to dementia. Various cognitive tests can help with diagnosis; meta-analysis of diagnostic accuracy studies would assist clinicians in choosing optimal tests.

Methods: We searched online databases for "mild cognitive impairment" and "diagnosis" or "screening" from 01/01/1999 to 01/07/2017. Articles assessing the diagnostic accuracy of a cognitive test compared with standard diagnostic criteria were extracted. Risk of bias was assessed. Bivariate random-effects meta-analysis was used to evaluate sensitivity and specificity.

Results: Eight cognitive tests (ACE-R, CERAD, CDT-Sunderland, IQCODE, Memory Alteration Test, MMSE, MoCA, and Qmci) were considered for meta-analysis. ACE-R, CERAD, MoCA, and Qmci were found to have similar diagnostic accuracy, while the MMSE had lower sensitivity. Memory Alteration Test had the highest sensitivity and equivalent specificity to the other tests.

Discussion: Multiple cognitive tests have comparable diagnostic accuracy. The Memory Alteration Test is short and has the highest sensitivity. New cognitive tests for MCI diagnosis should not be compared with the MMSE.

KEYWORDS

assessment, diagnosis, mild cognitive impairment, screening, test

1 | INTRODUCTION

Mild cognitive impairment (MCI) is a potentially significant diagnosis. It is regarded as a prodrome of dementia, involving greater forgetfulness than one would expect for their age, yet retaining all or almost all of day-to-day independence and not meeting criteria for clinically probable dementia.[1] Around half of those diagnosed with MCI will develop dementia within 3 years, and from the point of MCI diagnosis, 6% to 15% of patients will convert to dementia per year.[2,3] Increasing interest is directed at early diagnosis of dementia; even though current drug treatments are not indicated for MCI,[4,5] early diagnosis has benefits: the future care needs of the patient can be to some degree anticipated, and arrangements can be made in good time, with the patient being involved in these decisions at a stage where their decision-making is relatively unimpaired. Furthermore, a recent meta-analysis has shown that there are some modifiable predictors of conversion to dementia, including untreated diabetes and low dietary folate (with the caveat that there is currently no prospective evidence proving that modifying these risk factors reduces the risk of progressing to dementia, with the possible exception of folate supplementation).[6,7] MCI is associated with depression, itself a potentially modifiable risk factor for progression to dementia.[8,9] Nonetheless, the modifiable risk factors identified by Cooper et al are pathologies in their own right, with effective treatments available in primary care.

Nonetheless, the diagnosis of MCI must be made carefully, and patients must be counselled that MCI has a variable natural history and prognosis, and that they will not necessarily develop dementia. A diagnosis of MCI can be both stigmatising and anxiety-provoking,[10] and we acknowledge that many clinicians may not wish to screen for MCI, given that it has a variable natural history and that treatment is

largely supportive. With this caveat in mind, screening for MCI, while not imperative, may clearly be useful in clinical practice, and current American Academy of Neurology consensus guidelines offer an up-to-date summary of the current clinical pathway.[11]

The diagnosis of MCI was first characterised in 1999 by the Petersen criteria.[12] The diagnosis of MCI is also codified in DSM-5.[13,14]

Petersen (2004) criteria for MCI diagnosis[15]:

1. Subjective cognitive complaint, ideally corroborated by an informant
2. Impaired cognition for age demonstrated on neuropsychological testing
3. Preserved general cognitive function
4. Intact activities of daily living
5. Not demented.

Mild cognitive impairment can be further categorised into amnestic, non-amnestic, and multi-domain impairment subtypes[15] (a-MCI, na-MCI, and md-MCI, respectively). However, the validity of some of these subtypes has been questioned.[16]

The prevalence of MCI in the community is difficult to determine accurately, since different studies have used slightly different definitions of MCI.[17] In community-dwelling adults over 65 years old, MCI prevalence is variously reported as between 3% and 25%,[18,19] with variability arising from factors such as the mean age of the sample.

The use of a standardised cognitive test is a typical starting point for the formal neuropsychological assessment of patients with a cognitive complaint and is recommended by current MCI consensus guidelines.[11] Several factors may influence the choice of test, such as clinicians' familiarity with the test, availability of translations, copyright, ease of administration, time constraints, evidence base of the test, and its perceived accuracy.[20] In the psychogeriatric clinic setting, the Mini-Mental State Examination[21] (MMSE) and Montreal Cognitive Assessment[22] (MoCA) are popular. It must be noted that each cognitive test arose in different contexts: the MMSE was devised in 1975 as an all-purpose bedside cognitive examination for psychiatric inpatients, not just dementia syndromes—the concept of MCI was not yet formally recognised—while the MoCA was devised specifically with more difficult items in mind to increase the sensitivity for MCI in the psychogeriatric clinic setting, but also included frontal and executive function items to increase sensitivity for atypical dementia syndromes. More recently, further brief tests have been developed specifically for the diagnosis of MCI, such as the Memory Alteration Test[23] (M@T) and Quick Screen for Mild Cognitive Impairment[24] (Qmci). Unlike older tests such as the Addenbrooke's Cognitive Examination Revised (ACE-R), Consortium to Establish a Registry for Alzheimer's Disease Battery total score[25] (CERAD) and MoCA, these newer instruments do not test as many different cognitive domains but rather aim to separate MCI and dementia from healthy individuals by testing only the domains believed to be more selectively affected. For instance, both the Qmci and M@T have a more extensive recall component, as this is believed to be affected early in the course of MCI.[26] In both tests, the orientation component aims to identify those with dementia. Both of these shorter tests can be administered in around 5 minutes, making them potentially suitable for use in general practice.

Key points

- Summary sensitivity, but not specificity, of the MoCA, CERAD, M@T, and ACE-R is significantly higher than that of the MMSE for MCI.
- Comprehensive cognitive tests (ACE-R, CERAD, MoCA) have similar sensitivity and specificity for MCI.
- The Memory Alteration Test is the cognitive test with the highest sensitivity for MCI.
- Memory Alteration Test and Qmci are short tests with useful diagnostic accuracy for MCI.

Use of cognitive testing in UK general practice is mostly focussed on detecting dementia rather than MCI, and the most popular instruments are the 6-Item Cognitive Impairment Test, MMSE, Clock-Drawing Test (CDT), and Abbreviated Test of Mental Status, according to a small survey of general practitioners in Kent, England.[27] A review by Brodaty and colleagues[28] has found the Memory Impairment Screen,[29] General Practitioner Assessment of Cognition,[30] and Mini-Cog[31] to be the most suitable for dementia screening in primary care, based on brevity, ease of use, and validation in primary care settings.

The present meta-analysis aims to identify the cognitive tests with the most extensive evidence base for the diagnosis of MCI, evaluating both sensitivity and specificity. In particular, we believe that a brief screening test for MCI should favour sensitivity over specificity, such that diagnoses are not missed following an initial assessment. A false positive result can be rectified with more careful consideration of the patient's history, whereas a false negative result may result in a delay to the patient receiving the correct diagnosis or being lost to follow-up. On the other hand, the potential harms and costs arising from over-referral are not well studied, and in a research setting a high specificity is desired.

Additionally, we will review the published literature for (1) computer-based measures for the diagnosis of MCI and (2) informant-based measures for the diagnosis of MCI. Although a similar meta-analysis has recently been published on this subject,[32] our meta-analysis differs in several important respects: it excludes studies identifying MCI in the context of Parkinson's disease and other neurological conditions, and includes studies on new, brief, validated screening instruments (M@T and Qmci), comparing these to more established cognitive tests. Additionally, our meta-analysis includes a full presentation of the QUADAS-2 assessment of methodological quality.

2 | METHODS

2.1 | Literature search

We searched Medline, EMBASE, PsycINFO, and Scopus with the following search terms (MCI OR mild cognitive impairment) AND (diagnosis OR screening OR detection OR test OR assessment OR validation OR informant) and a list of exclusion terms, from 01/01/1999 to 01/07/2017. Fuller details of the search strategy are detailed in the Supporting Information.

We sought studies published in English where one or more neuropsychological tests (hereafter referred to as the "index test") were used, together with a specified numerical cut-off score, to distinguish individuals with MCI versus healthy controls or individuals with subjective memory complaints, and compared with a "reference standard" diagnosis (psychiatrist or multidisciplinary team consensus diagnosis based on published criteria, taking into account the patient's history and longer neuropsychological batteries). For the broadest possible range of studies, we included studies from both community and secondary care settings, using cutoffs at any numerical value, regardless of whether it was pre-specified or not. Studies were excluded if:

1. The reference standard diagnosis of MCI was not made according to published criteria (Petersen criteria,[12,15] Winblad criteria,[1] Gauthier criteria,[33] National Institute of Ageing—Alzheimer's Association criteria,[34] DSM-5 criteria for mild neurocognitive disorder[13]). Since the first Petersen criteria were published in 1999,[12] articles published before 1999 were excluded.
2. Patients with pre-existing medical, neurological, and/or psychiatric illness besides MCI or dementia were included in the study. This is because neurological diseases such as Parkinson's disease can present with atypical MCI syndromes,[35] including early functional impairment.
3. Sensitivity and specificity of the index test at a given cut-off for the diagnosis of MCI vs a control population were not presented in the paper.

2.2 | Selection of studies

Two independent reviewers (N.A. and A.B.) assessed studies for inclusion by brief sorting of titles, abstracts, and full texts. Disagreements were resolved by consensus.

2.3 | Data extraction

Two independent reviewers (N.A. and A.B.) extracted data from the studies into a standardised form. Different scoring systems of the same test were considered as distinct tests. This applies mostly to the different CDT scoring systems, which we have abbreviated as CDT-[author of scoring system]. The study design and population, MCI subtype identified, control group selection and characteristics, index test(s), and reference standard were extracted. Additionally, the total number of subjects with MCI and controls, alongside the reported sensitivity and specificity of the index test at a given cut-off for distinguishing between MCI and cognitively normal individuals, were extracted into Review Manager, and 2 × 2 tables were reconstructed. Where several sensitivity/specificity pairs were presented in the study, the one described as optimal by the authors was extracted, or the one with the greatest area under curve. For each test included in the meta-analysis, we extracted data on the number and type of items in the test, as well as the average administration time from the original citation of that cognitive test. This information is presented in Table 2.

2.4 | Assessment of methodological quality

The methodological quality of the selected studies was conducted with the QUADAS-2 tool independently by two raters (A.B. and D.C.). A summary table of the ratings was constructed with the two reviewers resolving disagreements by consensus. QUADAS-2 is a structured tool recommended by the Cochrane Collaboration, which evaluates studies in four domains: patient selection, index test, reference standard and flow and timing.[36] We added a further question, which was whether studies assessed functional impairment with a standardised scale. Results are presented in the Supporting Information.

2.5 | Meta-analysis

Quantitative meta-analysis of diagnostic accuracy was considered for index tests whose sensitivity and specificity had been evaluated in at least five published articles meeting the criteria above. The studies considered for quantitative meta-analysis are presented in Table 1 of the Supporting Information. Studies comparing more than one index test to a reference standard diagnosis have each index test comparison entered separately. Bivariate random-effect meta-analysis was conducted according to the method described by Reitsma et al,[37] generating a summary receiver-operated curve (SROC) with calculated area under curve and a summary estimate of sensitivity and specificity, with confidence intervals (CI) creating a 95% confidence region ellipse on the SROC. We did not calculate the I^2 statistic, as this is a univariate measure not recommended for diagnostic accuracy reviews.[38] All statistics were conducted in R using the mada package,[39] which does not calculate the 95% prediction region.

3 | RESULTS

3.1 | Literature search

Our literature search revealed 9694 primary studies, of which 66 were considered for the meta-analysis. The results of the search are summarised in the flow diagram (Figure 1). Cognitive tests evaluated in less than five published articles are as follows: 6-Item Cognitive Impairment Test (n = 1), ABCS 135 (n = 2), ACE (n = 4), ACE-III (n = 2) AD8 (n = 4), CAMCOG (n = 4), CANS-MCI (n = 2), CANTAB (n = 1), CDT-Babins (n = 3), CDT-Cohen (n = 2), CDT-Lin (n = 1), CDT-Mendez (n = 1), CDT-Rouleau (n = 4), CDT-Shulman (n = 3), CDT-Wolf-Klein (n = 2), Cogstate (n = 1), DemTect (n = 3), DRS-2 (n = 2), HVLT (n = 1), LST (n = 1), Mini-ACE (n = 1) Memory Impairment Screen (n = 2), SAGE (n = 1) TYM (n = 4), and VFT (n = 4). The only computer-based tests in the above list are the CANTAB, CANS-MCI, and Cogstate.

Two informant measures were identified: the AD8 questionnaire[40] and Informant Questionnaire for Cognitive Decline in Elderly[41] (IQCODE). Of these, the IQCODE's diagnostic accuracy was evaluated in five studies meeting our inclusion criteria. Other tests included in the meta-analysis were as follows: ACE-R (n = 6 studies), CDT-Sunderland[42] (n = 7 studies), CERAD (n = 5 studies), M@T (n = 5 studies), MoCA (n = 24 studies), MMSE (n = 46 studies), and Qmci (n = 5 studies). Forest plots for all the studies, grouped by test, are included in the Supporting Information. The results of the literature search are

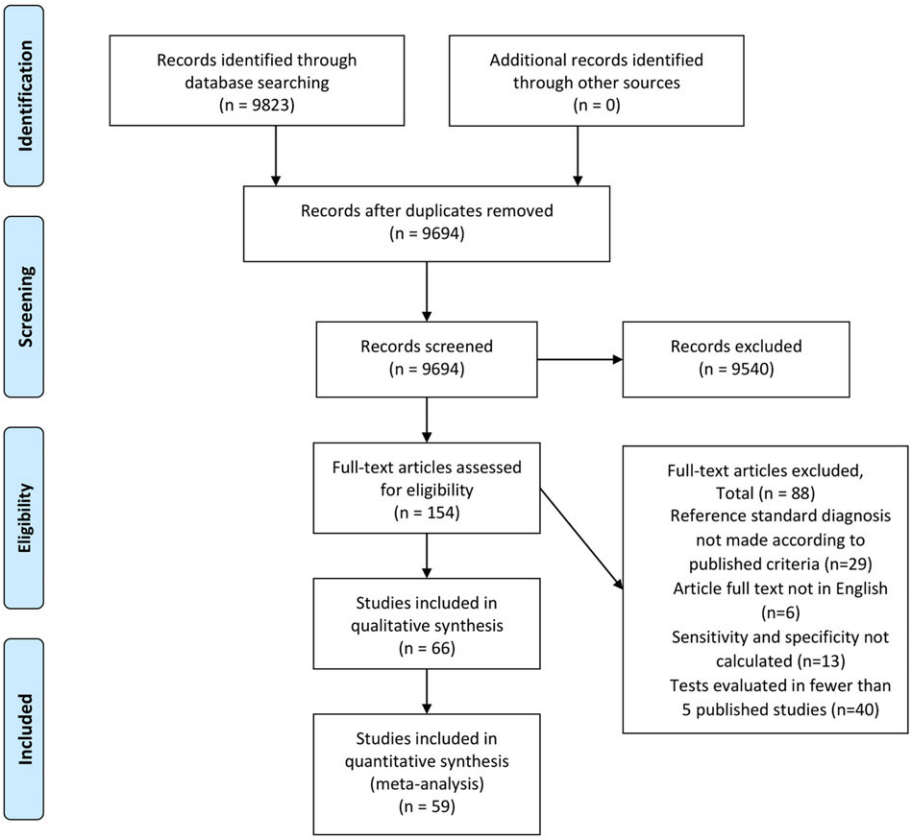

FIGURE 1 Flow diagram summarising the literature search and study selection [Colour figure can be viewed at wileyonlinelibrary.com]

summarised in Table 1. With the exception of the IQCODE, which is an informant questionnaire, the components of the cognitive tests meeting our inclusion criteria are summarised in Table 2. Characteristics of the included studies are described in the Supporting Information.

While considering tests for meta-analysis, studies of the CDT-Sunderland and IQCODE diagnostic accuracy were found to be too heterogeneous for quantitative meta-analysis (visual inspection of scatterplots, Figure 2). We therefore decided to give a narrative overview of the studies evaluating the CDT and IQCODE.

3.2 | CDT has unclear diagnostic accuracy for MCI diagnosis

The diagnostic accuracy of the CDT-Sunderland was evaluated in seven studies. Four studies found high specificity (>0.85), and three found lower specificity (0.57-0.70). Of the four studies reporting high specificity, two reported very low sensitivity (ca. 0.3) while two reported low sensitivity (0.60). The remaining three studies reported moderate sensitivity (0.58 to 0.70). This heterogeneity in results is likely due to considerable heterogeneity in the underlying study populations (community samples vs tertiary care samples) (Figure 2, and Forest plot in the Supporting Information).

TABLE 1 Summary of literature search results

Test	Number of Studies	Total Participants	Total Participants with MCI
ACE-R	6	563	271
CDT-Sunderland	7	4263	867
CERAD	5	4076	706
IQCODE	5	1372	849
Memory alteration test	5	1485	427
MoCA	24	4095	1573
MMSE	46	17 749	7493
Qmci	5	1206	395

3.3 | IQCODE has unclear diagnostic accuracy for MCI diagnosis

The largest study of the IQCODE's diagnostic accuracy found very high sensitivity (0.98) and moderate specificity (0.72). Three other studies found more moderate sensitivities (0.74-0.82) and specificities (0.54-0.72). One very small study found poor sensitivity (0.46) and high specificity (0.89). The relatively small number of primary studies with large heterogeneity and highly unequal sample sizes made fitting the bivariate model impossible (Figure 2, and Forest plot in the Supporting Information).

3.4 | Meta-analysis

3.4.1 | Comprehensive cognitive tests have similar diagnostic accuracy for MCI diagnosis

The three comprehensive tests in the meta-analysis (ACE-R, CERAD, MoCA) have extremely similar summary sensitivity and specificity with 95% CI for both parameters which overlap (Figure 3). Area under the SROC was 0.839 for ACE-R, 0.856 for CERAD, and 0.847 for MoCA (see Supporting Information for full results).

TABLE 2 Characteristics of the cognitive tests examined in our meta-analysis, including components and administration time. Note that only ACE-R, CERAD, and MoCA test all of the domains listed

Test Name	Orientation (Number of Items)	Memory/Learning (Number of Items)	Semantic Memory	Attention, Calculation	Fluency, Abstraction	Visuospatial, Constructional	Total Score	Administration Time (min)
ACE-R[43]	Time (5), place (5)	Word recall (3), name and address (7)	Item naming, semantic probe questions, irregular word reading, sentence and word repetition, sentence reading and writing	Backward spelling or serial subtraction	"P" letter and "animal" category	Copy Necker cube, overlapping pentagons. Clock drawing	100	12-20
CDT-Sunderland[42]	-	-	-	Hand and number placement	-	Clock drawing	10	5
CERAD[25]	Time (5), place (5)	Word registration and recall, word list	Item naming, sentence repetition, command following, sentence reading/writing	Backward spelling or serial subtraction	Fluency: "Animal" category	Copy: Overlapping pentagons, 2D and 3D figures	100	30
M@T[23]	Time (5)	Sentence registration (5) with free (10) and cued (10) recall, word registration (5) and recall (5)	General knowledge questions (15)	-	-	-	50	5
MoCA[22]	Time (4), place (5)	Word registration and recall (3)	Item naming, sentence repetition	Short TMT B, digit span forward and backward, tapping task, serial subtraction	Abstraction: Similarities	Copy: cube. Drawing: clock	30	10-12
MMSE[21]	Time (5), place (5)	Word registration (3), word recall (3)	Item naming, command following, sentence reading, sentence writing	Backward spelling	-	Pentagon copying	30	10
Qmci[24]	Time (4), place (1)	Word registration (5), word recall (5), sentence registration with free recall	-	-	Fluency: "Animal" category	Clock drawing	100	5

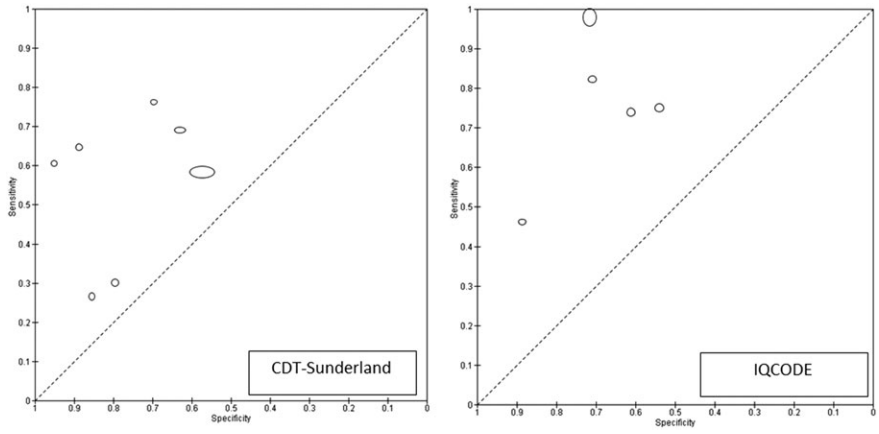

FIGURE 2 ROC plot of CDT-Sunderland studies (left) and IQCODE studies (right). Larger ovals represent studies with larger sample sizes

3.4.2 | MMSE is less sensitive for MCI diagnosis

The meta-analysis of the MMSE's summary sensitivity (0.664, 95% CI 0.605-0.718) does not overlap with that of the MoCA (0.812, 95% CI 0.771-0.847), strongly suggesting that the MMSE is a less sensitive instrument than the MoCA for MCI diagnosis. Specificity is however equivalent (0.735, 95% CI 0.686-0.778) (Figure 3).

3.4.3 | Qmci is an effective non-comprehensive test for MCI diagnosis

Quick Screen for Mild Cognitive Impairment had a similar diagnostic accuracy for MCI as the comprehensive cognitive tests (area under SROC 0.836, sensitivity 0.770 [95% CI 0.712-0.820], specificity 0.789 [95% CI 0.711-0.851]). Sensitivity and specificity estimates are not significantly different from those of the comprehensive cognitive tests (Figure 3).

3.4.4 | Memory Alteration Test (M@T) is the test with the highest sensitivity

Summary ROC curves for MMSE, MoCA, ACE-R and Memory Alteration Test are compared in Figure 4. The confidence region of the M@T does not overlap with those of the remaining cognitive tests on the vertical axis (sensitivity). The M@T's summary estimate of sensitivity is 0.951 (95% CI 0.892-0.978) compared with 0.812 (95% CI 0.771-0.847) for the MoCA. M@T also has good specificity, equivalent to the MoCA (0.84, 95% CI 0.667-0.932). Area under SROC was 0.961 for M@T.

3.4.5 | Meta-regression

The bivariate model of diagnostic accuracy meta-analysis used thus far allows covariates to be entered into the model at the study level to determine whether they are associated with statistically significant differences in sensitivity, specificity, or both[38] (meta-regression). Using only the MoCA dataset, we performed meta-regression using type of controls (subjective memory complaints or cognitively healthy older adults) as a covariate. While most of the studies of the MoCA used individuals without cognitive complaints as controls (n = 21 studies), others used patients with subjective memory complaints presenting to memory clinic, but not meeting criteria for MCI or dementia, as controls (n = 3 studies). Studies using patients with subjective memory complaints as controls reported lower sensitivity and specificity, but the effect did not reach statistical significance (P = 0.216 for sensitivity, P = 0.134 for specificity).

Given that the M@T showed the highest diagnostic accuracy but was used for the diagnosis of aMCI in four out of the five studies analysed, we performed a meta-regression to see whether the MoCA's diagnostic accuracy was higher in studies identifying aMCI only (n = 4 studies) compared with studies identifying all MCI subtypes (n = 20 studies). In studies identifying aMCI only, the MoCA showed higher specificity, but this was not statistically significant (P = 0.894 for sensitivity and P = 0.211 for specificity).

4 | DISCUSSION

The results of the meta-analysis presented can contribute to the literature on early detection of MCI. Despite having different inclusion criteria, our results largely concur with those of the meta-analysis by Tsoi and colleagues.[32] With respect to the MoCA and MMSE, our summary estimates of sensitivity and specificity were almost identical. While Tsoi and colleagues considered a variety of recall tests together, we only considered single tests, demonstrating that the M@T (which has a more significant recall component than the other tests in this meta-analysis) to have a similar summary sensitivity and specificity to the recall tests as presented by Tsoi and colleagues. We have shown that two brief instruments (Qmci and M@T) have useful diagnostic accuracy for MCI detection and an administration time of around 5 minutes, making them potentially suitable for the evaluation of patients presenting with cognitive complaints in primary care. Although current consensus guidelines on MCI recommend that patients with cognitive complaints are assessed with any standardised test,[11] implying that all tests have similar diagnostic accuracy, our meta-analysis shows that although several tests have largely similar accuracy, the MMSE has clearly inferior sensitivity to the other tests evaluated in this meta-analysis and the CDT-Sunderland has an unclear sensitivity and specificity, while the IQCODE as an informant measure is clearly not suitable for all patients, and also has an unclear sensitivity and specificity for MCI. Our assessment of the CDT is in agreement with the systematic review by Ehreke and colleagues.[44]

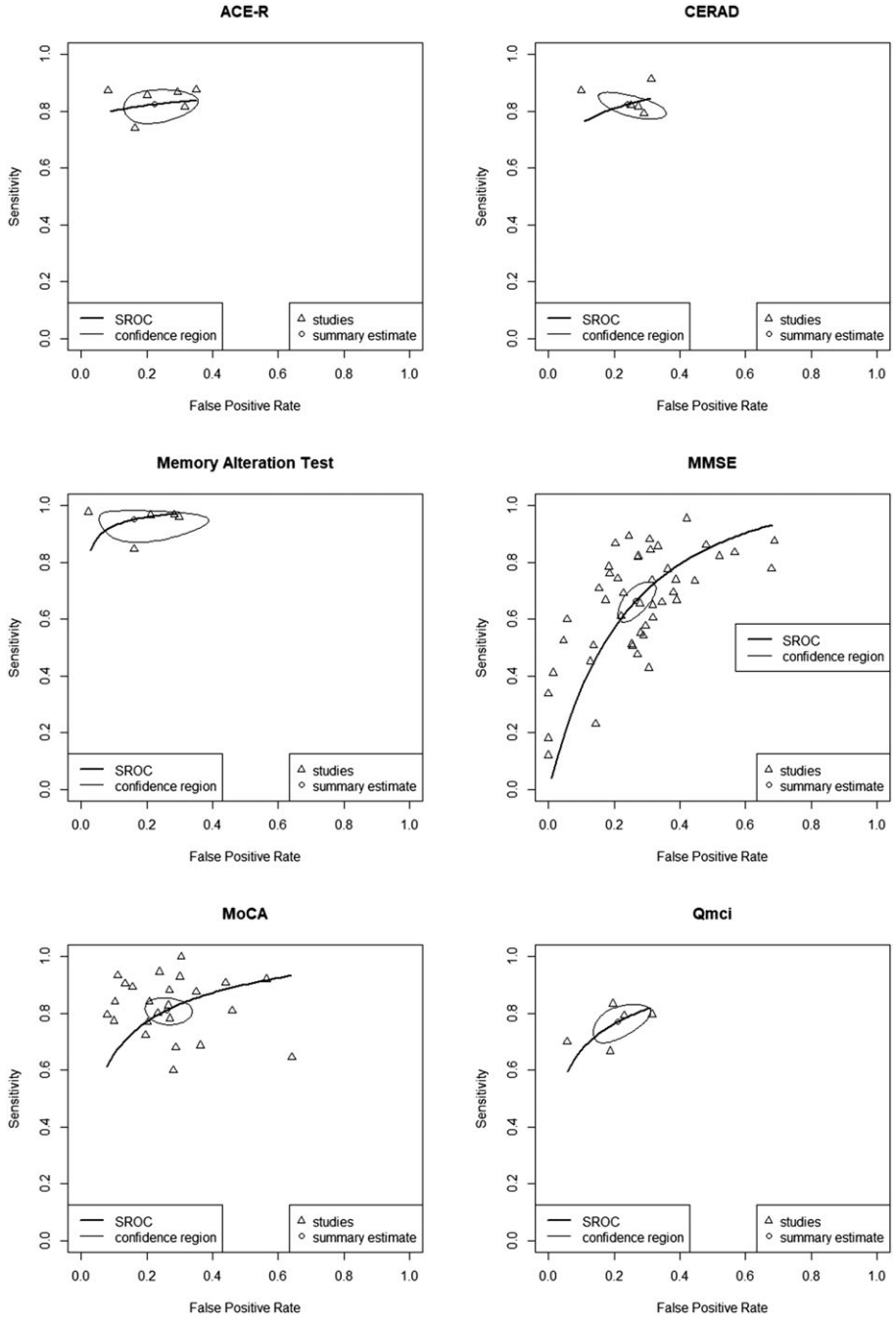

FIGURE 3 SROC plots (bivariate model) of diagnostic test performance for ACE-R, CERAD, Memory Alteration Test, MMSE, MoCA, and Qmci with SROC curve (black line), summary estimate (circle), 95% confidence regions (ellipse), and individual studies (triangles). SROC curves are not extrapolated beyond the range of the original data. False positive rate = (1 − specificity)

With respect to the MMSE, our meta-analysis joins an increasing literature calling on clinicians not to use the MMSE if MCI is suspected,[45] especially since the MoCA is both more sensitive to impairment in a variety of clinical settings[46,47] (including MCI, as demonstrated in the present meta-analysis) and available free of copyright in 35 languages. This is to be expected, given that the MoCA was developed with the specific intention of being sensitive for MCI. We note that a recent systematic review of RCTs on aMCI found that MMSE was the most commonly used instrument for defining the "preserved general cognitive function" criterion of the Petersen criteria.[48] Also of note is that many electro-convulsive therapy services in Europe still specify the use of the MMSE as screening for cognitive impairment among their patients,[49] and many general practitioners continue to use the MMSE in their practice.[27] We hypothesise that under-diagnosis of dementia may be in part due to the ongoing widespread use of cognitive tests with lower sensitivity for the early stages of the disease (MCI), and our meta-analysis demonstrates that brief and sensitive cognitive tests are available free of copyright. The Memory Alteration Test seems to be the most suited of all the tests reviewed in this article for use in primary care, given that it had the highest sensitivity, has been evaluated in two studies with primary care populations and does not require the patient to write, making it suitable for patients with lower educational level. We suspect however that more validation studies in naturalistic general practice

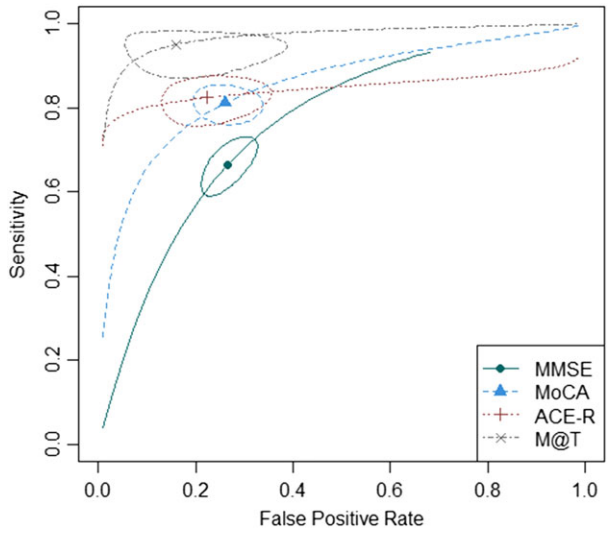

FIGURE 4 SROC plot (bivariate model) comparing diagnostic test accuracy of MMSE, MoCA, ACE-R, and M@T. Points represent the summary estimate of sensitivity and specificity as determined by bivariate meta-analysis. Ellipses show 95% confidence intervals for each test's diagnostic accuracy. SROC curves have been extrapolated beyond the range of the original data. False positive rate = (1 − specificity) [Colour figure can be viewed at wileyonlinelibrary.com]

samples would be required. Moreover, four out of five studies evaluating its diagnostic accuracy were conducted with aMCI patients only, while the Qmci's evidence base included all MCI subtypes. The M@T is available free of copyright for clinical use (but licensed for research and commercial use) in English, Spanish, and Portuguese, and the Qmci is available for unrestricted non-commercial use under Creative Commons license in English, Dutch, and Turkish.

In the tertiary care setting, comprehensive cognitive tests (ACE-R, CERAD, MoCA) were found to have very similar sensitivity and specificity. Even though brief tests were shown to have comparable sensitivity and specificity for MCI diagnosis, we believe that the use of a comprehensive test is important in the tertiary care setting, as its use may uncover deficits in the visuospatial or executive domains, which are not tested by the M@T and minimally tested in the Qmci. Given that several underdiagnosed subtypes of dementia (dementia with Lewy bodies, Parkinson's disease dementia, together making up 10%-15% of all dementia diagnoses) may present with prominent visuospatial and/or executive deficits[50] as well as memory impairment, use of comprehensive cognitive tests remains desirable.

Our meta-analysis also has some limitations. Most of the included studies were case-control studies of patients with pre-existing diagnoses and healthy controls taken as samples of convenience or recruited from the community, which can inflate the diagnostic accuracy of a test by including phenotypic extremes. Few studies used consecutive memory clinic referrals as their study population. We note that our meta-regression found that in the case of the MoCA, studies using patients with subjective memory complaints as controls (n = 4) reported lower sensitivity and specificity, although the low number of studies meant that our analysis may have been underpowered to show statistical significance. Additionally, patients with psychiatric comorbidity, pre-existing medical illness, or sensory impairments were not included. We therefore believe that the diagnostic accuracies reported in the primary studies are likely to be higher than would be encountered in daily clinical practice. Another significant source of bias comes from the fact that most studies reported sensitivity and specificity pairs at the cut-off which was found to be optimal after results were analysed, again inflating estimates of diagnostic accuracy.

In the QUADAS-2 summary results, more than half (60%-100%) of all the studies for each test showed "high" or "unclear" risk of bias in the "patient selection" domain, largely for the reasons noted above. The exception was the Qmci, where only 40% of studies had "high" or "unclear" risk of bias in this domain. For all the tests, minority of studies (10%-40%) showed "high" or "unclear" risk of bias in the "index test" and "reference standard" domains, while only a smaller minority (>10%) showed concerns regarding applicability. Therefore, with the exception of the Qmci, it appears that the biases noted above do not affect one cognitive test significantly more than another, with the caveat that QUADAS-2 is a purely qualitative assessment of risk of bias. Regarding the Qmci, it is possible that reduced bias from patient selection may have reduced the summary sensitivity and/or specificities calculated in this meta-analysis, when compared with the other tests.

Considerable heterogeneity was found among the included studies, particularly for the IQCODE, CDT-Sunderland, MoCA, and less for CERAD and ACE-R. This is partly due to the heterogeneity of included patients and controls (we note that the MoCA has been evaluated across a broader range of countries), and the fact that MCI remains a clinical diagnosis with imprecise definitions. We attempted to minimise the latter factor by only including studies making the MCI diagnosis with explicit reference to published criteria, but the interpretation of the criteria may nonetheless vary between clinicians and patients, especially since different patients (and different cultures) may have varying expectations regarding their cognitive function as they age.

ACKNOWLEDGEMENT

We would like to thank Prof. Klaus P. Ebmeier for his support, advice on this project and reviewing the manuscript, Dr Ronan O'Caoimh, and Dr Lorena Rami for supplying additional calculations.

During this study Dr N.A was a H. Clinical Research Fellow at the Department of Psychiatry, University of Oxford and now is a visiting researcher at the Department of Psychiatry, University of Cambridge. Dr A.B was a student at the Medical School of the University of Oxford during this study and is now a junior doctor at Derriford Hospital, Plymouth, UK.

FUNDING

Dr A.B received a grant from Trinity College, University of Oxford to assist in the purchase of a reference management software.

DECLARATIONS OF INTERESTS

None.

ORCID

Nikitas A. Arnaoutoglou http://orcid.org/0000-0003-0840-271X

REFERENCES

1. Winblad B, Palmer K, Kivipelto M, et al. Mild cognitive impairment—beyond controversies, towards a consensus: Report of the International Working Group on Mild Cognitive Impairment. *J Intern Med*. 2004;256(3):240-246.
2. Tierney MC, Szalai JP, Snow WG, et al. Prediction of probable Alzheimer's disease in memory-impaired patients: a prospective longitudinal study. *Neurol*. 1996;46(3):661-665.
3. Mitchell AJ, Shiri-Feshki M. Rate of progression of mild cognitive impairment to dementia--meta-analysis of 41 robust inception cohort studies. *Acta Psychiatr Scand*. 2009;119(4):252-265.
4. Tricco AC, Soobiah C, Berliner S, et al. Efficacy and safety of cognitive enhancers for patients with mild cognitive impairment: a systematic review and meta-analysis. *CMAJ*. 2013;185(16):1393-1401.
5. Petersen RC, Thomas RG, Grundman M, et al. Vitamin E and donepezil for the treatment of mild cognitive impairment. *N Engl J Med*. 2005;352(23):2379-2388.
6. Cooper C, Sommerlad A, Lyketsos CG, Livingston G. Modifiable predictors of dementia in mild cognitive impairment: a systematic review and meta-analysis. *Am J Psychiatry*. 2015;172(4):323-334.
7. Malouf R, Grimley EJ. Folic acid with or without vitamin B12 for the prevention and treatment of healthy elderly and demented people. *Cochrane Database Syst Rev*. 2008;(4):CD004514.
8. Ismail Z, Elbayoumi H, Fischer CE, et al. Prevalence of depression in patients with mild cognitive impairment: a systematic review and meta-analysis. *JAMA Psychiat*. 2017;74(1):58-67.
9. Polyakova M, Sonnabend N, Sander C, et al. Prevalence of minor depression in elderly persons with and without mild cognitive impairment: a systematic review. *J Affect Disord*. 2014;152–154:28-38.
10. Garand L, Lingler JH, Conner KO, Dew MA. Diagnostic labels, stigma, and participation in research related to dementia and mild cognitive impairment. *Res Gerontol Nurs*. 2009;2(2):112-121.
11. Petersen RC, Lopez O, Armstrong MJ, et al. Practice guideline update summary: mild cognitive impairment: Report of the Guideline Development, Dissemination, and Implementation Subcommittee of the American Academy of Neurology. *Neurol*. 2018;90(3):126-135.
12. Petersen RC, Smith GE, Waring SC, Ivnik RJ, Tangalos EG, Kokmen E. Mild cognitive impairment: clinical characterization and outcome. *Arch Neurol*. 1999;56(3):303-308.
13. Association AP. *Diagnostic and Statistical Manual of Mental Disorders*. 5th Edition ed. Arlington, VA: American Psychiatric Publishing; 2013.
14. Sachs-Ericsson N, Blazer DG. The new DSM-5 diagnosis of mild neurocognitive disorder and its relation to research in mild cognitive impairment. *Aging Ment Health*. 2015;19(1):2-12.
15. Petersen RC. Mild cognitive impairment as a diagnostic entity. *J Intern Med*. 2004;256(3):183-194.
16. Klekociuk SZ, Summers MJ. Exploring the validity of mild cognitive impairment (MCI) subtypes: multiple-domain amnestic MCI is the only identifiable subtype at longitudinal follow-up. *J Clin Exp Neuropsychol*. 2014;36(3):290-301.
17. Petersen RC. Challenges of epidemiological studies of mild cognitive impairment. *Alzheimer Dis Assoc Disord*. 2004;18(1):1-2.
18. Nie H, Xu Y, Liu B, et al. The prevalence of mild cognitive impairment about elderly population in China: a meta-analysis. *Int J Geriatr Psychiatry*. 2011;26(6):558-563.
19. Sachdev PS, Lipnicki DM, Kochan NA, et al. The prevalence of mild cognitive impairment in diverse geographical and ethnocultural regions: the COSMIC collaboration. *PLoS One*. 2015;10(11):e0142388.
20. Newman JC. Copyright and bedside cognitive testing: why we need alternatives to the mini-mental state examination. *JAMA Intern Med*. 2015;175(9):1459-1460.
21. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state"A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res*. 1975;12(3):189-198.
22. Nasreddine ZS, Phillips NA, Bedirian V, et al. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *J Am Geriatr Soc*. 2005;53(4):695-699.
23. Rami L, Molinuevo JL, Sanchez-Valle R, Bosch B, Villar A. Screening for amnestic mild cognitive impairment and early Alzheimer's disease with M@T (Memory Alteration Test) in the primary care population. *Int J Geriatr Psychiatry*. 2007;22(4):294-304.
24. O'Caoimh R, Gao Y, McGlade C, et al. Comparison of the quick mild cognitive impairment (Qmci) screen and the SMMSE in screening for mild cognitive impairment. *Age Ageing*. 2012;41(5):624-629.
25. Chandler MJ, Lacritz LH, Hynan LS, et al. A total score for the CERAD neuropsychological battery. *Neurol*. 2005;65(1):102-106.
26. Ribeiro F, Guerreiro M, De Mendonca A. Verbal learning and memory deficits in mild cognitive impairment. *J Clin Exp Neuropsychol*. 2007;29(2):187-197.
27. Milne A, Culverwell A, Guss R, Tuppen J, Whelton R. Screening for dementia in primary care: a review of the use, efficacy and quality of measures. *Int Psychogeriatr*. 2008;20(5):911-926.
28. Brodaty H, Low LF, Gibson L, Burns K. What is the best dementia screening instrument for general practitioners to use? *Am J Geriatr Psychiatry*. 2006;14(5):391-400.
29. Buschke H, Kuslansky G, Katz M, et al. Screening for dementia with the memory impairment screen. *Neurol*. 1999;52(2):231-238.
30. Brodaty H, Pond D, Kemp NM, et al. The GPCOG: a new screening test for dementia designed for general practice. *J Am Geriatr Soc*. 2002;50(3):530-534.
31. Borson S, Scanlan JM, Chen P, Ganguli M. The mini-cog as a screen for dementia: Validation in a population-based sample. *J Am Geriatr Soc*. 2003;51(10):1451-1454.
32. Tsoi KKF, Chan JYC, Hirai HW, et al. Recall tests are effective to detect mild cognitive impairment: a systematic review and meta-analysis of 108 diagnostic studies. *J Am Med Dir Assoc*. 2017;18(9):807 e817-807 e829.
33. Gauthier S, Reisberg B, Zaudig M, et al. Mild cognitive impairment. *The Lancet*. 2006;367(9518):1262-1270.
34. Albert MS, DeKosky ST, Dickson D, et al. The diagnosis of mild cognitive impairment due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement*. 2011;7(3):270-279.
35. Caviness JN, Driver-Dunckley E, Connor DJ, et al. Defining mild cognitive impairment in Parkinson's disease. *Mov Disord*. 2007;22(9):1272-1277.
36. Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med*. 2011;155(8):529-536.
37. 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. *J Clin Epidemiol*. 2005;58(10):982-990.
38. Takwoingi Y, Riley RD, Deeks JJ. Meta-analysis of diagnostic accuracy studies in mental health. *Evid Based Ment Health*. 2015;18(4):103-109.
39. Doebler P. Meta-analysis of diagnostic accuracy with mada. 2012; https://cran.r-project.org/web/packages/mada/vignettes/mada.pdf. Accessed 18/04/2018, 2018.
40. Galvin JE, Roe CM, Powlishta KK, et al. The AD8: a brief informant interview to detect dementia. *Neurol*. 2005;65(4):559-564.
41. Jorm AF, Jacomb PA. The Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE): socio-demographic correlates, reliability, validity and some norms. *Psychol Med*. 1989;19(4):1015-1022.
42. Sunderland T, Hill JL, Mellow AM, et al. Clock drawing in Alzheimer's disease. A novel measure of dementia severity. *J Am Geriatr Soc*. 1989;37(8):725-729.

43. Mioshi E, Dawson K, Mitchell J, Arnold R, Hodges JR. The Addenbrooke's Cognitive Examination Revised (ACE-R): A brief cognitive test battery for dementia screening. *Int J Geriatr Psychiatry*. 2006;21(11):1078-1085.
44. Ehreke L, Luppa M, Konig HH, Riedel-Heller SG. Is the clock drawing test a screening tool for the diagnosis of mild cognitive impairment? A systematic review. *Int Psychogeriatr*. 2010;22(1):56-63.
45. Carnero-Pardo C. Should the mini-mental state examination be retired? *Neurologia*. 2014;29(8):473-481.
46. Xu Q, Cao WW, Mi JH, Yu L, Lin Y, Li YS. Brief screening for mild cognitive impairment in subcortical ischemic vascular disease: a comparison study of the Montreal cognitive assessment with the mini-mental state examination. *Eur Neurol*. 2014;71(3–4):106-114.
47. Pendlebury ST, Cuthbertson FC, Welch SJ, Mehta Z, Rothwell PM. Underestimation of cognitive impairment by mini-mental state examination versus the Montreal cognitive assessment in patients with transient ischemic attack and stroke: a population-based study. *Stroke*. 2010;41(6):1290-1293.
48. Christa Maree Stephan B, Minett T, Pagett E, Siervo M, Brayne C, McKeith IG. Diagnosing mild cognitive impairment (MCI) in clinical trials: a systematic review. *BMJ open*. 2013;3(2).
49. Moirand R, Galvao F, Lecompte M, Poulet E, Haesebaert F, Brunelin J. Usefulness of the Montreal cognitive assessment (MoCA) to monitor cognitive impairments in depressed patients receiving electroconvulsive therapy. *Psychiatry Res*. 2018;259:476-481.
50. McKeith IG, Boeve BF, Dickson DW, et al. Diagnosis and management of dementia with Lewy bodies: fourth consensus report of the DLB consortium. *Neurol*. 2017;89(1):88-100.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Breton A, Casey D, Arnaoutoglou NA. Cognitive tests for the detection of mild cognitive impairment (MCI), the prodromal stage of dementia: Meta-analysis of diagnostic accuracy studies. *Int J Geriatr Psychiatry*. 2018;1–10. https://doi.org/10.1002/gps.5016