


## REVIEW ARTICLE

# A systematic review of the diagnostic accuracy of automated tests for cognitive impairment

Rabeeh W. Aslam<sup>1</sup>  | Vickie Bates<sup>1</sup> | Yenel Dundar<sup>2</sup> | Juliet Hounsome<sup>1</sup> | Marty Richardson<sup>1</sup> | Ashma Krishan<sup>1</sup> | Rumona Dickson<sup>1</sup> | Angela Boland<sup>1</sup> | Joanne Fisher<sup>1</sup> | Louise Robinson<sup>3</sup> | Sudip Sikdar<sup>2</sup>

<sup>1</sup>Health Services, University of Liverpool, Liverpool, UK

<sup>2</sup>Mersey Care NHS Foundation Trust, Liverpool, UK

<sup>3</sup>Institute of Health and Society, Newcastle University, Newcastle upon Tyne, UK

**Correspondence**

Dr R. W. Aslam, University of Liverpool, Whelan Building, The Quadrangle, Brownlow Hill, Liverpool, L69 3GB, UK.

Email: r.w.aslam@liverpool.ac.uk

**Funding information**

Health Technology Assessment Programme, Grant/Award Number: 15/67/01

**Objective:** The aim of this review is to determine whether automated computerised tests accurately identify patients with progressive cognitive impairment and, if so, to investigate their role in monitoring disease progression and/or response to treatment.

**Methods:** Six electronic databases (Medline, Embase, Cochrane, Institute for Scientific Information, PsycINFO, and ProQuest) were searched from January 2005 to August 2015 to identify papers for inclusion. Studies assessing the diagnostic accuracy of automated computerised tests for mild cognitive impairment (MCI) and early dementia against a reference standard were included. Where possible, sensitivity, specificity, positive predictive value, negative predictive value, and likelihood ratios were calculated. The Quality Assessment of Diagnostic Accuracy Studies tool was used to assess risk of bias.

**Results:** Sixteen studies assessing 11 diagnostic tools for MCI and early dementia were included. No studies were eligible for inclusion in the review of tools for monitoring progressive disease and response to treatment. The overall quality of the studies was good. However, the wide range of tests assessed and the non-standardised reporting of diagnostic accuracy outcomes meant that statistical analysis was not possible.

**Conclusion:** Some tests have shown promising results for identifying MCI and early dementia. However, concerns over small sample sizes, lack of replicability of studies, and lack of evidence available make it difficult to make recommendations on the clinical use of the computerised tests for diagnosing, monitoring progression, and treatment response for MCI and early dementia. Research is required to establish stable cut-off points for automated computerised tests used to diagnose patients with MCI or early dementia.

**KEYWORDS**

ageing, Alzheimer disease, automated tests, computerised tests, dementia, diagnosis, MCI, monitoring

## 1 | INTRODUCTION

Cognitive impairment in dementia is a growing public health concern.<sup>1</sup> It is a distinctive characteristic of all dementias, and its timely assessment is a crucial and essential element in the diagnosis of dementia.<sup>2</sup> This is because some causes of dementia are treatable and are fully

or partially reversible, including dementias caused by vitamin B<sub>12</sub> deficiency,<sup>3</sup> side effects of medications,<sup>4</sup> metabolic abnormality, and certain brain tumours.<sup>5</sup> There is evidence from the United States that early recognition and treatment of dementia may delay the subsequent need for nursing home care and may reduce the risk of misdiagnosis and inappropriate management and reduce responsibilities for carers.<sup>6</sup>

Obtaining accurate incidence and prevalence figures for MCI is difficult since people with cognitive impairment may go undiagnosed.

This research was conducted at the University of Liverpool.

**Study registration:** The study is registered as PROSPERO CRD42015025410.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2018 The Authors. International Journal of Geriatric Psychiatry Published by John Wiley & Sons Ltd.

These estimates also vary significantly depending on the definitions used in different studies. For example, a large population-based study of older-aged individuals in the United Kingdom<sup>7</sup> reported prevalence estimates of individuals not classified from current MCI definitions were variable (range, 2.5–41.0%). In addition, the rates of progression from MCI to dementia varied from 3.7% to 30.0%.<sup>7</sup>

Evidence from neuropathological and neuroimaging studies suggests that biological changes associated with dementia occur long before the onset of symptoms.<sup>8</sup> This has given rise to the concept of mild cognitive impairment (MCI), which is the state between the cognitive changes of normal ageing and early dementia.<sup>9–11</sup> Mild cognitive impairment refers to the clinical condition used to describe people whose cognitive function is below that of the normal population for their educational level and age but who do not have any loss of functional abilities or skills.<sup>11–14</sup> It is a heterogeneous state, with possible trajectories including Alzheimer disease (AD), Lewy body dementias, and even reversion to normal cognitive functioning.<sup>15</sup>

The difference between MCI and early dementia is based on the level of cognitive decline and pattern of change in mood and behaviour. Individuals diagnosed with early dementia present with multiple cognitive deficits, and their memory loss is sufficient to impact everyday social and occupational functioning. Among the 4 most common medical conditions causing dementia are AD, vascular conditions, frontotemporal atrophy, and Lewy body disease. Irrespective of the primary reason, the cognitive prognosis for people with most types of dementia is usually poor.<sup>16,17</sup>

There are a number of pen-and-paper-based tools as suitable tests for screening people for cognitive impairment, for example, the General Practitioner Assessment of Cognition, 6-item Cognitive Impairment Test, and Mini-cog assessment instrument.<sup>18,19</sup> There are different pen-and-paper tests used to aid diagnosis by specialists for MCI and early dementia, for example, the Dementia Toolkit for Effective Communication,<sup>20</sup> Montreal Cognitive Assessment,<sup>21</sup> and Saint Louis University Mental Status.<sup>22</sup> However, these specialist tests can be expensive and time-consuming.<sup>23</sup> More recently, several automated tests have been developed,<sup>24,25</sup> which may be uniquely suited to early detection of changes in cognition, by, for example, covering a wider range of ability to precisely record accuracy and speed of response with a level of sensitivity not possible in standard administrations.<sup>23</sup>

The rationale for this review is to determine whether automated computerised tests for cognitive impairment have the potential to contribute to early diagnosis and simplify the current method of monitoring progression and treatment response compared with standard clinical practice.

## 2 | METHODS

A systematic review was performed to describe the diagnostic accuracy of automated tests to detect MCI and early dementia as well as investigate their role in monitoring disease progression and response to treatment. The methodology and reporting of this review followed the guidance set out by the Cochrane Handbook for Diagnostic Test Accuracy Reviews.<sup>26</sup> See Appendix S1 found in the Supporting Information for an abbreviation list.

### Key points

- Timely diagnosis of mild cognitive impairment (MCI) and early dementia is important for good prognosis and effective management.
- A number of automated tests for diagnosing and monitoring progression of cognitive impairment have been developed, which need to be used in conjunction with clinical assessment.
- The overall quality and quantity of the available evidence are insufficient to make recommendations on the clinical use of these automated computerised tests.
- Further research is required to examine the cut-off points for different populations in automated tests for diagnosing and monitoring progression and treatment response of MCI and early dementia.

### 2.1 | Criteria for considering studies for this review

Any study assessing the diagnostic accuracy of automated computerised tests to diagnose or monitor MCI or early dementia against a reference standard was considered for inclusion. Case studies and qualitative studies were excluded. Studies or diagnostic tools published in a non-English language were also excluded.

#### 2.1.1 | Participants

Participants were people with MCI or early dementia diagnosed by any recognised diagnostic standard.

#### 2.1.2 | Index tests

The index tests considered for inclusion were automated computerised tests of cognitive impairment, which can either be self-administered or interviewer administered.

#### 2.1.3 | Reference standard

The reference standard for this review is the clinical diagnosis of MCI and early dementia using a diagnostic criteria, for example, the *International Classification of Diseases*<sup>2</sup> edition 10 and the *Diagnostic and Statistical Manual of Mental Disorders* editions 4 and 5 (DSM-IV and DSM-V, respectively).<sup>27</sup> It is recognised 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.

### 2.2 | Search methods for identification of studies

The following electronic databases were searched from January 2005 to August 2015 to identify studies for inclusion: Medline, Embase, Cochrane database, Institute for Scientific Information, PsycINFO, and ProQuest for dissertations and theses (see Appendix S2 found in the Supporting Information for search strategy in Medline). Through citation tracking, one study from 2001 was included since it reported on a computerised tests currently in use in clinical practise. The number of references retrieved from

different databases is provided in Appendix S3 found in the Supporting Information, and were managed in Endnote X7.

### 2.3 | Selection of studies

Two reviewers independently screened all relevant titles and abstracts and full-text articles for inclusion. Any disagreements were resolved by discussion with a third reviewer.

### 2.4 | Data extraction and management

Data extraction forms were developed and piloted in an Excel spreadsheet by using 2 of the included studies. Data on study design, population characteristics, and outcomes were extracted by one reviewer and independently checked for accuracy by a second reviewer, with disagreements resolved through discussion with a third reviewer when necessary. The extracted data included information on the reference standard, index test, cut-off points, and the measures of diagnostic test accuracy including sensitivity, specificity, receiver operating characteristic curve, and the area under the curve (AUC) for discriminating amongst MCI, early dementia, and cognitively healthy individuals.

### 2.5 | Assessment of methodological quality

The methodological quality of the included studies was assessed by one reviewer and independently checked for accuracy by a second reviewer using the Quality Assessment of Diagnostic Accuracy Studies tool,<sup>28</sup> which is recommended by the Cochrane Diagnostic Test Accuracy Reviews Guidelines.<sup>29</sup> This tool is designed to evaluate the risk of bias and applicability of primary diagnostic accuracy studies using signalling questions in 4 domains: patient selection, index test, reference standard, and flow and timing.

### 2.6 | Statistical analysis and data synthesis

An Excel spreadsheet was used to construct  $2 \times 2$  tables of index test performance. The measures of index test performance were recorded by the number of true-positive, true-negative, false-positive, and false-negative, sensitivity, and specificity values of MCI and early dementia. The sensitivity and specificity values with 95% confidence intervals, positive and negative predictive values (PPV and NPV, respectively), and positive and negative likelihood ratios (LR+ and LR-, respectively) were calculated when not reported in the studies. Out of authors of all the included studies approached with a request for specific sensitivity and specificity data, only 2 provided these data.

It was not possible to perform a meta-analysis because of non-comparable data; the study designs varied, the cut-off points for the primary outcome measure were heterogeneous, and the summary statistics were often inconsistently reported. A narrative synthesis of the results of the included studies was conducted.

### 2.7 | Patient and public involvement

An advisory group comprising clinicians and service users guided the team during the review. A call for participation was sent through frontline

groups, for example, Alzheimer's Society and Dementia UK, to identify people interested in giving feedback on the results of the review and on the final report. The review team took guidance from these agencies and INVOLVE<sup>30</sup> for planning and facilitating the meetings.

## 3 | RESULTS

### 3.1 | Results of the search

The electronic search was conducted in August 2015, and 18 796 records were retrieved, of which 399 articles were shortlisted for full-text assessment (Figure 1). The comprehensive search strategy was necessary because indexing of diagnostic accuracy studies is poor. In total, 16 studies met the inclusion criteria for detecting MCI and early dementia. No studies met the review inclusion criteria for monitoring progression or treatment response in MCI or early dementia, and therefore, there is no further mention of monitoring disease progression in the results section.

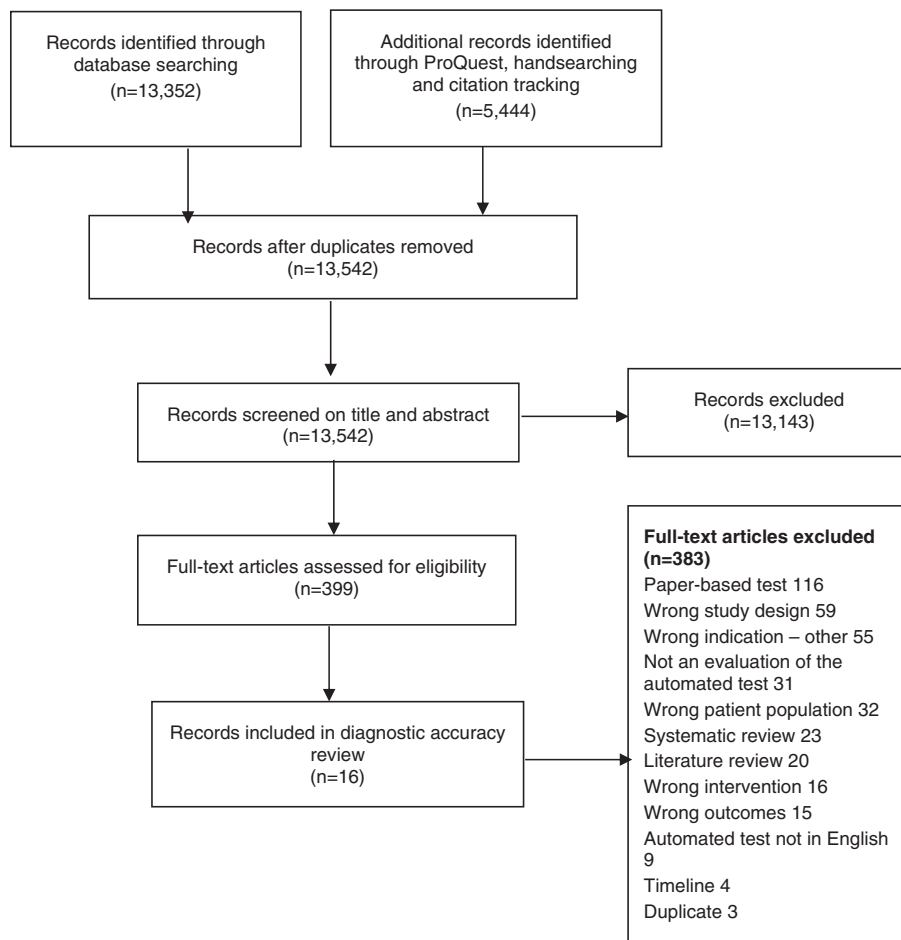
In addition to the 16 included studies, 4 trials were identified during hand searching (Appendix S4 found in the Supporting Information). The authors of these studies were approached by email and telephone for results, but no responses were received. The summary of the included 16 studies is presented in Table 1; there were 7 cohort studies, 7 case-control studies, and 2 cross-sectional studies.<sup>40,43</sup> Seven of the 16 included studies evaluated the use of automated computerised tests to detect MCI alone, 2 studies reported results for early dementia, 6 studies reported results for combined MCI/early dementia, and 1 study reported on cognitive impairment with a co-morbidity, eg, human immunodeficiency virus (HIV)-associated neurocognitive disorders (HANDs).<sup>43</sup> Two different reference standards were used for MCI in these studies, 9 studies used the Petersen criteria, and 4 studies used clinical diagnosis with a battery of neurocognitive tests. The reference standard for early dementia varied across different studies, 2 studies used National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer's Disease and Related Disorders Association Alzheimer's Criteria,<sup>42,46</sup> 2 studies used DSM-IV,<sup>33,34</sup> 1 study used the DSM-V criteria,<sup>39</sup> 2 studies used clinical diagnosis with neurocognitive tests,<sup>36,46</sup> and 1 study used the Clinical Dementia Rating score.<sup>41</sup>

#### 3.1.1 | Findings

The diagnostic accuracy of 11 automated computerised tests for the detection of MCI and/or early dementia without co-morbidities was evaluated in 15 studies and 1 study with co-morbidity.<sup>43</sup> The details of the index tests are summarised in Table 2. Pooling of data from these 16 studies was considered inappropriate since there were few studies evaluating the same index test in the same population, and it was only possible to extract  $2 \times 2$  data from 5 of the 16 studies.

### 3.2 | Studies reporting on diagnostic accuracy outcomes with a $2 \times 2$ table

There were 5 studies that reported diagnostic accuracy outcomes in a  $2 \times 2$  table as described in Table 3. Two studies reported the diagnostic



**FIGURE 1** Preferred Reporting Items for Systematic Reviews and Meta-Analyses flow diagram

accuracy outcomes for MCI, 3 studies reported outcomes for early dementia, and 1 study reported combined outcomes for both MCI and early dementia.

### 3.2.1 | Mild cognitive impairment

Juncos-Rabadan et al<sup>35</sup> evaluated 3 different visual episodic memory tests included in the Cambridge Neuropsychological Test Automated Battery (CANTAB); these memory tests were Pattern Recognition Memory, Delayed Matching to Sample, and Paired Associated Learning. The overall sensitivity and specificity for the 3 visual episodic memory tests were moderate at 79.7% and 76.3%, respectively. The overall AUC for the different visual episodic tests was not reported, but ranged from 0.623 (Delayed Matching to Sample) to 0.747 (Paired Associated Learning), showing poor ability to discriminate between the MCI group and the non-MCI group. This test had a high overall PPV of 71.4%; this means 71.4% of the people who tested positive for MCI with the index test actually had MCI according to the reference standard. Similarly, the overall NPV for this test was 83.3%, meaning that 83.3% of people who tested negative for MCI on the index test did not have MCI. This test had a low overall LR+ of 3.4, which shows a low likelihood of the test to establish the presence of disease. It also had a low overall LR- of 0.3, which shows a low likelihood of the test to establish the absence of disease.

The study by Saxton et al<sup>44</sup> evaluated the Computer Assessment of Memory and Cognitive Impairment (CAMCI) and reported good sensitivity (86%) and exceptional specificity (94%). The reported AUC (0.91) was also very high.

### 3.2.2 | Early dementia

The CANTAB Paired Associated Learning (CANTAB-PAL) was evaluated in 2 of the studies. Junkkila et al<sup>36</sup> reported high sensitivity (81.8%) and specificity (97.2%) and an AUC of exceptional discrimination (0.914) for early dementia.

The study by O'Connell et al<sup>42</sup> reported poor sensitivity (67.6%) and high specificity (100%) and an AUC of moderate discrimination (0.780) between the early-dementia group and non-early-dementia group.

Mundt et al<sup>41</sup> assessed the Computer Automated Telephone System and reported moderate sensitivity (79.17%) and high specificity (83.8%) for this test.

### 3.2.3 | MCI/early dementia

One study evaluated CANTAB-PAL. The authors reported high sensitivity (96.9%) and high specificity (80.8%) with an AUC of good discrimination (0.897) between the MCI/early-dementia group and non-MCI/early-dementia group.

TABLE 1 Study and participant characteristics

| Study                              | Condition   | Country, Setting  | N   | Mean age, years (SD, range)  | Gender (Male %)                             | Mean Education, y (SD, Range)   | Index Test Name                        | Reference Test   |
|------------------------------------|---|---|---|--|---|---|--|--|
| Ahmed et al <sup>31</sup>          | MCI   | United Kingdom<br>Primary care (Oxford OPTIMA study) <sup>a</sup> | 35 (control: 20, MCI: 15)                                     | Control: 77.4 (4)<br>MCI: 80.9 (7.2)   | Control: 55.0<br>MCI: 33.3                  | Control: 14.7 (2.9)<br>MCI: 13.1 (3)  | CANS-MCI                               | Clinical diagnosis using the Petersen criteria   |
| De Jager et al <sup>32</sup>       | MCI   | United Kingdom<br>Community                                       | 119 (control: 98, MCI: 21)                                    | Control: 77.18 (5.9)<br>MCI: 81.95 (5.4)   | NR  | Undeclared  | CogState                               | Clinical diagnosis using battery of neurocognitive tests   |
| Doniger et al <sup>33</sup>        | MCI<br>MCI/mild dementia                              | United States<br>Tertiary care<br>Memory clinic                   | 161 (control: 71, MCI: 58, mild AD: 32)                       | Entire group: 76.0 (8.2)   | Entire group: 37.5                          | Entire group: 13.3 (3.6)  | Mindstreams (abridged)                 | Clinical diagnosis using the Petersen criteria for MCI and DSM-IV for dementia                       |
| Dwolatky et al <sup>34</sup>       | MCI<br>Mild AD  | Canada/Israel<br>2 tertiary care<br>memory clinics                | 98 (control: 39, MCI: 30, mild AD: 29)                        | Control: 73.41 (8.00)<br>MCI: 77.15 (6.43)<br>Mild AD: 80.55 (4.91)                          | Control: 33.3<br>MCI: 56.7<br>Mild AD: 44.8 | Control: 14.95 (3.5)<br>MCI: 13.07 (2.86)<br>Mild AD: 11.31 (2.85)                    | Mindstreams                            | Clinical diagnosis using the Petersen criteria for MCI and DSM-IV for mild AD                        |
| Juncos-Rabadan et al <sup>35</sup> | aMCI  | Spain<br>Primary care   | 162 (control: 85, mda-MCI: 29, sda-MCI: 48)                   | Control: 62.25 (8.26, 50-82)<br>mda-MCI: 71.68 (7.74, 54-87)<br>sda-MCI: 68.02 (9.04, 50-84) | All participants: 36.4                      | Control: 10.83 (5, 2-21)<br>mda-MCI: 10.06 (3.99, 3-20)<br>sda-MCI: 9.83 (3.96, 2-20) | CANTAB-R (PRM, DMS, and PAL)           | Clinical diagnosis using neurocognitive tests and the Albert criteria and Peterson criteria for aMCI |
| Junkkila et al <sup>36</sup>       | aMCI/mild/probable dementia<br>Mild/probable dementia | Finland<br>Hospital   | 58 (control: 22, aMCI: 17, AD: 19)                            | Control: 70 (4.48, 65-80)<br>aMCI: 73 (6.3, 61-83)<br>AD: 73 (6.76, 61-83)                   | Control: 36.36<br>aMCI: 64.7<br>AD: 26.35   | Control: 10 (3.25)<br>aMCI: 8 (3)<br>AD: 8 (2.88)                                     | CANTAB-PAL                             | Clinical diagnosis using the Petersen criteria and neurocognitive tests                              |
| Kingsbury et al <sup>37</sup>      | MCI   | Australia<br>Community<br>Memory clinic                           | 140 (control: 95, MCI: 30, depressed: 15)                     | Control: 68.85 (7.96, 53-89)<br>MCI: 77.62 (7.45, 51-87)                                     | Control: 37<br>MCI: 43                      | Controls: 4.93 (1.71)<br>MCI: 3.07 (1.71)<br>Undeclared what is measured              | CogniScreen                            | Clinical diagnosis using the Petersen criteria   |
| Kluger et al <sup>38</sup>         | MCI<br>Early dementia                                 | United States<br>Memory clinic                                    | 101 (control: 39, MCI: 19, probable AD: 17, no diagnosis: 25) | Control: 64 (11)<br>MCI: 72 (10)<br>Probable AD: 78 (9)                                      | NR  | NR  | Computerised test (no name)            | Diagnosed by a consensus of at least 2 clinicians  |
| Lichtenberg et al <sup>39</sup>    | MCI/early dementia                                    | United States<br>Specialised<br>geriatric clinic                  | 102 (control: 55, MCI: 11, mild dementia: 36)                 | All participants: 79.3 (6.6)   | All participants: 46.1                      | All participants: 13.5 (2.9)  | CST                                    | Clinical diagnosis using the Petersen criteria; clinical diagnosis of dementia using DSM-V           |
| Maruff et al <sup>40</sup>         | MCI   | Australia<br>Primary care   | 766 (control: 659, aMCI: 107)                                 | Control: 69.5(6.6)<br>MCI: 75.7 (7.5)  | Control: 42.2<br>MCI: 49.5                  | Control: 12 <sup>a</sup> (9-15)<br>MCI: 12 <sup>a</sup> (9-15)                        | CBB                                    | Clinical diagnosis using the Petersen criteria   |
| Mundt et al <sup>41</sup>          | Dementia  | United States<br>Specialised<br>geriatric clinic                  | 116 (control: 74, mild dementia: 42)                          | All participants: 76.7 (7.0, 56-93)  | All participants: 36.7                      | All participants: 13.3 (3, 6-22)  | Computer-automated telephone screening | Clinical diagnosis using CDR score   |

(Continues)

TABLE 1 (Continued)

| Study                         | Condition                 | Country, Setting  | N                                       | Mean age, years (SD, range)  | Gender (Male %)  | Mean Education, y (SD, Range)   | Index Test Name | Reference Test   |
|-------------------------------|---------------------------|---|---|--|--|---|-----------------|--|
| O'Connell et al <sup>42</sup> | Probable AD               | Ireland<br>Memory clinic  | 50 (control: 16, probable AD: 34)       | Control: 72.6 (7.7)<br>Probable AD: 73 (5.9)   | Control: 12.5<br>Probable AD: 32.4   | NR  | CANTAB-PAL      | Clinical diagnosis using the NINCDS-ADRDA criteria   |
| Rosenthal et al <sup>43</sup> | HAND                      | United States<br>General clinical research clinic                     | 55 (HIV+ controls: 16, HAD: 39)         | HIV+ controls: 45.4 (6)<br>HAD: 48.3 (6.3)   | HIV+ controls: 75.0<br>HAD: 71.8   | HIV+ controls: 12.3 (1.8)<br>HAD: 12.6 (2.1)  | CAMCI modified  | HAND category using the Frascati criteria  |
| Saxton et al <sup>44</sup>    | MCI                       | United States<br>Primary care and community                           | 524 (control: 296, MCI: 228)            | Control: 71.84 (5.95)<br>MCI: 75.18 (6.76)   | MCI: 37.7<br>Control: 32.8   | Control: 13.74 (2.69)<br>MCI: 13.10 (2.61)  | CAMCI           | Clinical diagnosis by consensus using battery of neurocognitive tests and functional and medical information |
| Tierney et al <sup>45</sup>   | MCI                       | Canada<br>Tertiary care   | 263<br>NR                               | Completed without assistance: 78.7 (6.9)<br>Completed with assistance: 81.8 (6.5)  | All participants: 41.4   | Completed without assistance: 15.2 (3.2)<br>Completed with assistance: 13.9 (4.0)   | CAMCI           | Clinical diagnosis using battery of neurocognitive tests   |
| Vacante et al <sup>46</sup>   | MCI<br><br>Early dementia | United Kingdom<br><br>Primary care (Oxford OPTIMA study) <sup>a</sup> | 78 (control: 40, MCI: 20, early AD: 18) | Traditional version Control: 74.7 (7.78)<br>MCI: 78.3 (8.4)<br>Early AD: 73.67 (6.28)<br>Novel version Control: 73.67 (7.14)<br>MCI: 79.7 (6.07)<br>Early AD: 77.22 (4.94) | Traditional version Control: 50<br>MCI: 60<br>Early AD: 66.7<br>Novel version Control: 45<br>MCI: 60<br>Early AD: 77.8 | Traditional version Control: 15.85 (3.36)<br>MCI: 15.9 (3.32)<br>Early AD: 15 (3.04)<br>Novel version Control: 16.35 (3.18)<br>MCI: 15 (2.66)<br>Early AD: 16.11 (2.97) | TPT             | Clinical diagnosis using the Petersen criteria   |

Abbreviations: AD, Alzheimer disease; aMCI, amnesic mild cognitive impairment; CAMCI, Computer Assessment of Mild Cognitive Impairment; CANS-MCI, Computer-Administered Neuropsychological Screen for Mild Cognitive Impairment; CANTAB, Cambridge Neuropsychological Test Automated Battery; CANTAB-PAL, Cambridge Neuropsychological Test Automated Battery Paired Associated Learning; CBB, CogState Brief Battery; CDR, Clinical Dementia Rating Scale; CST, Computerised Self-Test; DMS, Delayed Matching to Sample; DSM-IV, *Diagnostic and Statistical Manual of Mental Disorders* edition 4; HAD, HIV-associated dementia; HAND, HIV-associated neurocognitive disorder; HIV+, human immunodeficiency virus; NR, not reported; MCI, mild cognitive impairment; mda-MCI, multiple-domain amnesic mild cognitive impairment; NINCDS-ADRDA, National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer's Disease and Related Disorders Association; OPTIMA, Oxford Project to Investigate Memory and Ageing; PAL, Paired Associated Learning; PRM, Pattern Recognition Memory; sda-MCI, single-domain amnesic mild cognitive impairment; TPT, The Placing Test.

<sup>a</sup>It is unclear as to whether these cohorts were independent to each other.

<sup>b</sup>Median.



TABLE 2 Index test details

| Study                               | Test Name                              | Cognitive Domains Tested  | Details of Test Platform Used   | Time (min)       | Method of Administration   |
|-------------------------------------|--|---|---|------------------|--|
| Ahmed et al <sup>31</sup>           | CANS-MCI                               | Memory<br>Language<br>Visuospatial<br>Executive function                      | Desktop computer, a touch screen system with both oral (loud speakers) and on screen instructions               | 30               | Self-administered<br>Researcher in room  |
| De Jager et al <sup>32</sup>        | CogState                               | Memory<br>Executive function<br>Attention<br>Processing speed                 | Internet  | Approximately 20 | Self-administered<br>Practice session with a psychologist                                      |
| Doniger et al <sup>32</sup>         | Mindstreams (abridged)                 | Memory<br>Executive function<br>Visuospatial<br>Motor skills                  | Computer and mouse  | 30               | Self-administered<br>Practice session  |
| Dwolatzky et al <sup>34</sup>       | Mindstreams                            | Memory<br>Executive function<br>Visuospatial                                  | Designed for use with older people. Mouse with the number pad on the keyboard (similar to the telephone keypad) | 45               | Self-administered<br>Practice session with feedback prior to testing<br>Research assistant     |
| Juncos-Rabadian et al <sup>35</sup> | CANTAB-R (PRM, DMS, and PAL)           | Verbal<br>Attention<br>Information processing<br>Motor skills                 | Touch screen computer   | NR               | Self-administered<br>Researcher present  |
| Junkkila et al <sup>36</sup>        | CANTAB-PAL                             | Memory  | Touch screen computer   | NR               | Self-administered  |
| Kingsbury et al <sup>37</sup>       | CogniScreen                            | Memory  | Laptop, headset with microphone   | 20-40            | Self-administered<br>Experimenter in room  |
| Kluger et al <sup>38</sup>          | Computerised test (no name)            | Memory<br>Praxis<br>Naming<br>Executive function                              | Laptop  | 12-15            | Self-administered<br>Screening test for computer competency                                    |
| Lichtenberg et al <sup>39</sup>     | CST                                    | Learning<br>Memory<br>Executive function                                      | Internet based, interface with both written and oral instructions   | 15               | Self-administered<br>Keyboard proficiency test<br>Administered by graduate psychology students |
| Maruff et al <sup>40</sup>          | CBB                                    | Memory  | Desktop computer, yes/no button attached through USB port   | 10               | Self-administered<br>Verbal instructions by supervisor<br>Practice session                     |
| Mundt et al <sup>41</sup>           | Computer-automated telephone screening | Memory<br>Spatial (auditory)<br>Executive function<br>orientation<br>Language | Standard touch tone telephones  | 11-15            | Self-administered<br>Researcher provided assistance in dialling the number                     |
| O'Connell et al <sup>42</sup>       | CANTAB-PAL                             | Memory  | Touch screen computer   | 10               | NR   |

(Continues)

TABLE 2 (Continued)

| Study                         | Test Name      | Cognitive Domains Tested                                      | Details of Test Platform Used | Time (min)       | Method of Administration                                     |
|-------------------------------|----------------|---|-------------------------------|------------------|--|
| Rosenthal et al <sup>43</sup> | CAMCI modified | Memory<br>Attention<br>Executive function<br>processing speed | Tablet with stylus            | 25               | Self-administered  |
| Saxton et al <sup>44</sup>    | CAMCI          | Memory<br>Attention<br>Executive function<br>processing speed | Desktop computer              | Approximately 20 | Self-administered  |
| Tierney et al <sup>45</sup>   | CAMCI          | Memory<br>Attention<br>Executive function<br>processing speed | Tablet computer               | 30               | Self-administered,<br>some required<br>researcher assistance |
| Vacante et al <sup>46</sup>   | TPT            | Memory  | Computer                      | 20               | Self-administered<br>Including practice pages                |

Abbreviations: CAMCI, Computer Assessment of Mild Cognitive Impairment; CANS-MCI, Computer-Administered Neuropsychological Screen for Mild Cognitive Impairment; CANTAB, Cambridge Neuropsychological Test Automated Battery; CANTAB-PAL, Cambridge Neuropsychological Test Automated Battery Paired Associated Learning; CBB, CogState Brief Battery; CST, Computerised Self-Test; DMS, Delayed Matching to Sample; NR, not reported; PAL, Paired Associated Learning; PRM, Pattern Recognition Memory; TPT, The Placing Test.

### 3.3 | Studies reporting on diagnostic accuracy outcomes without a 2 × 2 table

The authors of 11 studies reported diagnostic accuracy outcomes for 9 different index tests without using 2 × 2 data as tabulated in Table 4. Instead, they calculated optimal sensitivity and specificity values using receiver operating characteristic curve analysis.

#### 3.3.1 | Mild cognitive impairment

Eight studies reported the diagnostic accuracy outcomes for MCI. Ahmed et al evaluated Computer-Administered Neuropsychological Screen for Mild Cognitive Impairment and reported high sensitivity (89.0%) and moderate specificity (73.0%) with an AUC of 0.867, which shows a good ability to discriminate between the MCI group and the non-MCI group. Tierney et al evaluated the CAMCI test and reported a high sensitivity (80.0%) and a moderate specificity (74.0%); the authors did not report AUC values. Maruff et al evaluated the CogState Brief Battery (CBB). The CogState Brief Battery has 2 composite scores for 4 tasks: psychomotor function, attention function, learning memory, and working memory. The psychomotor/attention function had poor discrimination since its AUC was 0.67. It also had poor sensitivity (41.1%) but high specificity (85.7%). The AUC for the learning/working memory was 0.91, which shows exceptional ability to discriminate between the MCI group and the non-MCI group. It also had high sensitivity (80.4%) and high specificity (84.7%). The overall sensitivity, specificity, and AUC were not reported.

#### 3.3.2 | Early dementia

Dwolatzky et al<sup>34</sup> and Doniger et al<sup>33</sup> both assessed the Mindstreams computerised cognitive testing. Only Doniger et al reported results relating to early dementia. They evaluated an abridged version of Mindstreams with an overall AUC of 0.886, which showed a good ability to discriminate between the early-dementia group and the non-early-dementia group.

#### 3.3.3 | MCI/early dementia

Kluger et al evaluated an automated computerised test, which did not have a specific name. The authors reported an AUC of 0.97, which shows exceptional ability to discriminate between early dementia and healthy controls.

Doniger et al reported an overall AUC of 0.823, which showed a good ability to discriminate between the cognitively healthy group and the cognitive unhealthy group. The AUC values for individual test results ranged from 0.671 to 0.773.

Lichtenberg et al<sup>39</sup> reported sensitivity and specificity values (80.0% and 87.0%, respectively), PPV (88.0%), and NPV (79.0%).

#### 3.3.4 | HIV-associated neurocognitive disorders

One study<sup>43</sup> evaluated diagnostic accuracy of an automated computerised test that included people with cognitive impairment with co-morbidities. This study examined the HAND and used the automated test CAMCI. The CAMCI test assessed multiple domains with different tasks. The study examined a range of diagnostic accuracy outcomes but did not report the values for all of them.



**TABLE 3** Diagnostic accuracy outcomes with 2 × 2 table

| Study                              | Index Test                          | Cut-off                     | Sensitivity, %     | Specificity, %    | AUC                | TP  | FN | TN  | FP | PPV, %            | NPV, %            | LR +              | LR -               |
|------------------------------------|-------------------------------------|-----------------------------|--------------------|-------------------|--------------------|-----|----|-----|----|-------------------|-------------------|-------------------|--------------------|
| MCI                                |                                     |                             |                    |                   |                    |     |    |     |    |                   |                   |                   |                    |
| Juncos-Rabadan et al <sup>35</sup> | CANTAB Overall <sup>a</sup>         |                             | 79.7               | 76.3              | NR                 | 55  | 14 | 71  | 22 | 71.4              | 83.3              | 3.4               | 0.3                |
|                                    | PRM                                 | 1.5 SD below controls       | 45.5 <sup>b</sup>  | 92.9 <sup>b</sup> | 0.704 <sup>b</sup> | 35  | 42 | 79  | 6  | 85.4 <sup>b</sup> | 65.3 <sup>b</sup> | 6.44 <sup>b</sup> | 0.59 <sup>b</sup>  |
|                                    | DMS                                 | 1.5 SD below controls       | 23.4 <sup>b</sup>  | 97.6 <sup>b</sup> | 0.623 <sup>b</sup> | 18  | 59 | 83  | 2  | 90.0 <sup>b</sup> | 58.5 <sup>b</sup> | 9.94 <sup>b</sup> | 0.78 <sup>b</sup>  |
|                                    | PAL                                 | 1.5 SD below controls       | 58.4 <sup>b</sup>  | 89.4 <sup>b</sup> | 0.747 <sup>b</sup> | 45  | 32 | 76  | 9  | 83.3 <sup>b</sup> | 70.4 <sup>b</sup> | 5.52 <sup>b</sup> | 0.46 <sup>b</sup>  |
| Saxton et al <sup>44</sup>         | CAMCI                               | Final tree model            | 86                 | 94                | 0.91 <sup>b</sup>  | 201 | 27 | 277 | 19 | 91.4 <sup>b</sup> | 91.1 <sup>b</sup> | 13.7 <sup>b</sup> | 0.127 <sup>b</sup> |
| Early dementia                     |                                     |                             |                    |                   |                    |     |    |     |    |                   |                   |                   |                    |
| Junkkila et al <sup>36</sup>       | CANTAB-PAL                          | NR                          | 81.8 <sup>b</sup>  | 97.2 <sup>b</sup> | 0.914 <sup>b</sup> | 18  | 4  | 35  | 1  | 94.7 <sup>b</sup> | 89.7 <sup>b</sup> | 5.35 <sup>b</sup> | 0.03 <sup>b</sup>  |
| Mundt et al <sup>41</sup>          | Computer-automated telephone system | A derived scoring algorithm | 79.17 <sup>b</sup> | 83.8 <sup>b</sup> | 0.819 <sup>b</sup> | 38  | 10 | 62  | 12 | 76.0 <sup>b</sup> | 86.1 <sup>b</sup> | 4.88 <sup>b</sup> | 0.249 <sup>b</sup> |
| O'Connell et al <sup>42</sup>      | CANTAB-PAL                          | 32 errors                   | 67.6               | 100               | 0.780              | 23  | 11 | 16  | 0  | 100               | 59.3 <sup>b</sup> |                   | 0.324              |
| MCI/early dementia                 |                                     |                             |                    |                   |                    |     |    |     |    |                   |                   |                   |                    |
| Junkkila et al <sup>36</sup>       | CANTAB-PAL                          | NR                          | 96.9 <sup>b</sup>  | 80.8 <sup>b</sup> | 0.897 <sup>b</sup> | 31  | 1  | 21  | 5  | 86.1 <sup>b</sup> | 95.5 <sup>b</sup> | 5.04 <sup>b</sup> | 0.04 <sup>b</sup>  |

Abbreviations: AUC, area under curve; CAMCI, Computer Assessment of Mild Cognitive Impairment; CANTAB, Cambridge Neuropsychological Test Automated Battery; CANTAB-PAL, Cambridge Neuropsychological Test Automated Battery Paired Associated Learning; DMS, Delayed Matching to Sample; FN, false negative; FP, false positive; LR-, negative likelihood ratio; LR+, positive likelihood ratio; MCI, mild cognitive impairment; NPV, negative predictive value; NR, not reported; PAL, Paired Associated Learning; PPV, positive predictive value; PRM, Pattern Recognition Memory; TN, true negative; TP, true positive.

<sup>a</sup>The study details were provided by the primary author.

<sup>b</sup>Calculated by the research team.

TABLE 4 Diagnostic accuracy outcomes without 2 × 2 table

| Study                         | Index Test   | Cut-off    | Sensitivity, % | Specificity, %      | AUC (95% CI)                        | PPV, %   | NPV, %   | LR+      | LR-      |
|-------------------------------|--|------------|----------------|---------------------|-------------------------------------|----------|----------|----------|----------|
| MCI                           |  |            |                |                     |                                     |          |          |          |          |
| Ahmed et al <sup>31</sup>     | CANS-MCI   | 0.5        | 89.0           | 73.0                | 0.867 (0.743-0.990)                 | 60       | 84       | NR       | NR       |
| De Jager et al <sup>32</sup>  | CogState   | 82.6       | 78.0           | 90.0                | 0.86 (NR)                           | NR       | NR       | NR       | NR       |
|                               | Accuracy<br>Accuracy speed ratio                       | 3.54       | 76.0           | 79.0                | 0.84 (NR)                           | NR       | NR       | NR       | NR       |
| Dwolatzy et al <sup>34</sup>  | Mindstreams computerised cognitive testing             | NA for AUC | NR             | NR                  | 0.84 (NR)                           | NR       | NR       | NR       | NR       |
| Kingsbury et al <sup>37</sup> | CogniScreen  |            |                |                     |                                     |          |          |          |          |
|                               | Pair recognition                                       | 0.47       | 76.0           | 60.0                | 0.72 (0.62-0.83)                    | NR       | NR       | NR       | NR       |
|                               | Cued recall  | 0.305      | 82.1           | 76.7                | 0.87 (0.80-0.95)                    | NR       | NR       | NR       | NR       |
|                               | Immediate and delayed serial recall                    | 0.385      | 92.6           | 80.0                | 0.89 (0.81-0.97)                    | NR       | NR       | NR       | NR       |
| Kluger et al <sup>38</sup>    | Computerised test (no name)                            | NR         | NR             | NR                  | 0.89                                | NR       | NR       | NR       | NR       |
| Maruff et al <sup>40</sup>    | CBB  |            |                |                     |                                     |          |          |          |          |
|                               | Psychomotor/attention<br>Learning/working memory       | 90<br>90   | 41.1<br>80.4   | 85.7<br>84.7        | 0.67 (0.6-0.73)<br>0.91 (0.87-0.94) | NR<br>NR | NR<br>NR | NR<br>NR | NR<br>NR |
| Tierney et al <sup>45</sup>   | CAMCI  | 2          | 80.0           | 74.0                | NR                                  | NR       | NR       | NR       | NR       |
| Vacante et al <sup>46</sup>   | Computerised total (novel and traditional)             | 19.5       | 70.0           | 76.2                | NR                                  | NR       | NR       | NR       | NR       |
|                               | Computerised objects and faces (novel and traditional) | 12.5       | 50             | 64.3                | NR                                  | NR       | NR       | NR       | NR       |
|                               | Computerised objects and faces (novel and traditional) | 13.5       | 75             | 52.4                | NR                                  | NR       | NR       | NR       | NR       |
| Early dementia                |  |            |                |                     |                                     |          |          |          |          |
| Doniger et al <sup>33</sup>   | Mindstreams (abridged)                                 | NA         | NR             | NR                  | 0.886                               | NR       | NR       | NR       | NR       |
|                               | Overall  |            |                |                     |                                     |          |          |          |          |
|                               | Memory   |            |                |                     |                                     |          |          |          |          |
|                               | Verbal memory  |            |                |                     | 0.830 (0.762-0.898)                 | NR       | NR       | NR       | NR       |
|                               | Nonverbal memory                                       |            |                |                     | 0.825 (0.756-0.893)                 | NR       | NR       | NR       | NR       |
|                               | Executive function                                     |            |                |                     |                                     |          |          |          |          |
|                               | Go-No Go   |            |                |                     | 0.733 (0.640-0.826)                 | NR       | NR       | NR       | NR       |
|                               | Stoop interference                                     |            |                |                     | 0.790 (0.690-0.890)                 | NR       | NR       | NR       | NR       |
|                               | Catch game   |            |                |                     | 0.748 (0.670-0.827)                 | NR       | NR       | NR       | NR       |
|                               | Visual spatial   |            |                |                     |                                     |          |          |          |          |
| Visual spatial imagery        |  |            |                | 0.678 (0.567-0.789) | NR                                  | NR       | NR       | NR       |          |
| Dwolatzy et al <sup>34</sup>  | Mindstreams computerised cognitive testing             | NR         | NR             | NR                  | NR                                  | NR       | NR       | NR       | NR       |
|                               | Computerised test (no name)                            | NR         | NR             | NR                  | 0.97                                | NR       | NR       | NR       | NR       |
| Kluger et al <sup>38</sup>    | TPT  |            |                |                     |                                     |          |          |          |          |
|                               | Computerised total (novel and traditional)             | 15.5       | 88.9           | 92.9                | NR                                  | NR       | NR       | NR       | NR       |
|                               | Computerised objects and faces (novel and traditional) | 11.5       | 94.4           | 78.6                | NR                                  | NR       | NR       | NR       | NR       |
|                               | Computerised objects and faces (novel and traditional) | 13.5       | 94.4           | 52.4                | NR                                  | NR       | NR       | NR       | NR       |
| Vacante et al <sup>46</sup>   |  |            |                |                     |                                     |          |          |          |          |
|                               |  |            |                |                     |                                     |          |          |          |          |

(Continues)

TABLE 4 (Continued)

| Study   | Index Test   | Cut-off    | Sensitivity, % | Specificity, % | AUC (95% CI)  | PPV, % | NPV, % | LR+ | LR- |
|---|--|------------|----------------|----------------|---|--------|--------|-----|-----|
| MCI/early dementia<br>Doniger et al <sup>33</sup> | Mindstreams (abridged)<br>Overall<br>Memory<br>Verbal memory<br>Nonverbal memory<br>Executive function<br>Go-No Go<br>Stoop interference<br>Catch game<br>Visual spatial<br>Visual spatial imagery | NA for AUC | NR             | NR             | 0.823 (0.757-0.888)<br>0.773(0.697-0.849)<br>0.767 (0.690-0.844)<br>0.719 (0.639-0.800)<br>0.671 (0.575-0.766)<br>0.685 (0.595-0.776) | NR     | NR     | NR  | NR  |
| Lichtenberg et al <sup>39</sup>                   | CST  | 1.5        | 80.0           | 87.0           | 0.721 (0.638-0.803)<br>NR   | 88.0   | 79.0   | NR  | NR  |

Abbreviations: AUC, area under curve; CAMCI, Computer Assessment of Mild Cognitive Impairment; CANS-MCI, Computer-Administered Neuropsychological Screen for Mild Cognitive Impairment; CBB, CogState Brief Battery; CST, Computerised Self-Test; LR+, positive likelihood ratio; LR-, negative likelihood ratio; MCI, mild cognitive impairment; NA, not applicable; NPV, negative predictive value; NR, not reported; PPV, positive predictive value; TPT, The Placing Test.

3.3.5 | Methodological quality

The methodological quality of the included studies was assessed using the Quality Assessment of Diagnostic Accuracy Studies tool as summarised in Figure 2.

The risk-of-bias criterion for patient selection was high for 7 studies because a case-control study design had not been avoided (see Appendix S6 found in the Supporting Information). Seven studies were judged to be at unclear risk in the index test criteria for risk of bias since the threshold values for the index tests were not prespecified. There was high concern regarding the applicability of the index test for all of the studies because the interpretation of the index test was different from the review question, since it is not possible to establish diagnosis of MCI and early dementia using automated computerised tests in isolation; specialist expertise is necessary to establish a diagnosis.

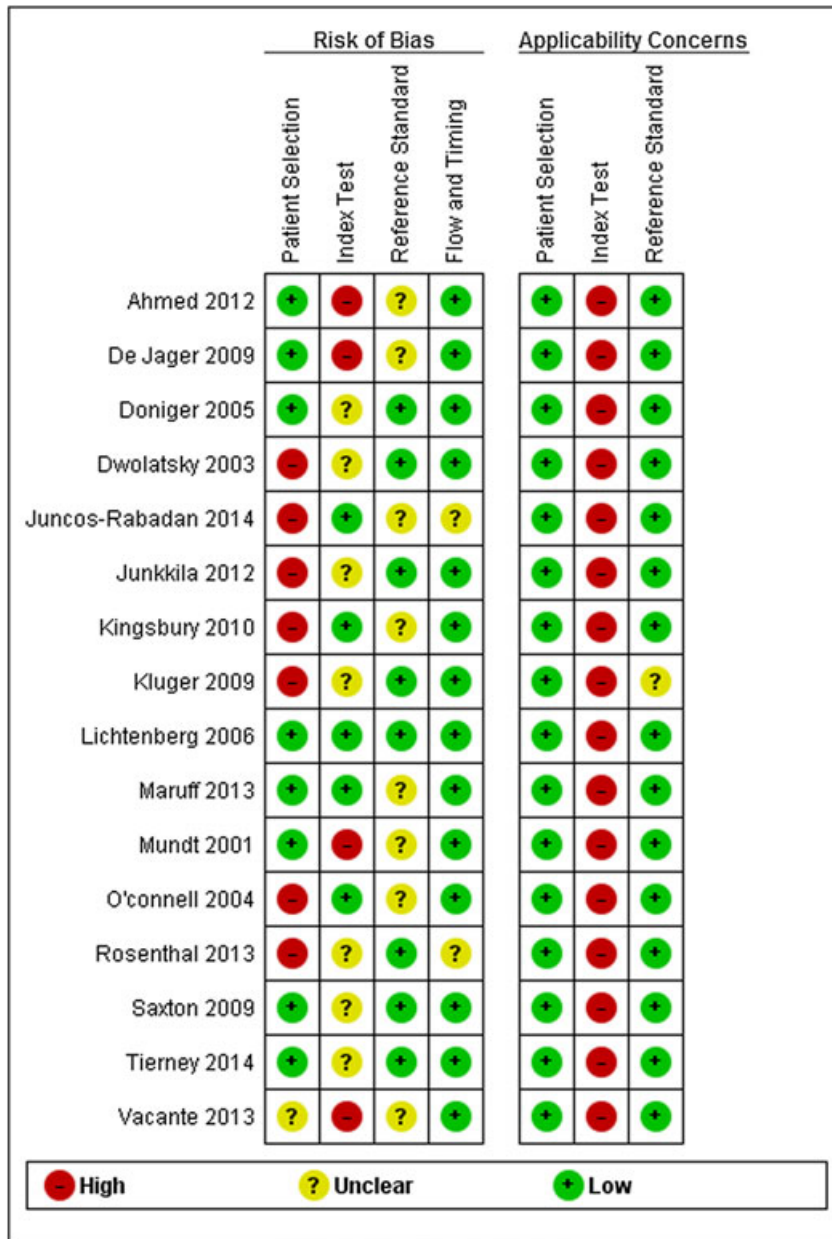
The reference standard domain for the risk of bias was unclear in 8 studies since it was not possible to ascertain whether reference standard results were interpreted without knowledge of the results of the index tests. All but one study<sup>38</sup> were judged to have low concern for applicability regarding the reference standard since it used a consensus of 2 clinicians' opinions as the reference standard. In the flow and timing domain for the risk of bias, a judgement of unclear risk of bias was given to 2 studies<sup>35,43</sup> since attrition or timing was not described in the papers. However, 14 studies were assessed as being at low risk because all patients had received the same reference standard and all patients were included in the analysis. There was a high concern in the domains of applicability for 16 studies. Of the 16 studies, only 1 was judged to be at low of risk of bias across the 4 domains examined<sup>39</sup>; despite this, the overall quality of the included studies was considered to be good.

3.4 | Patient and public involvement

Data from the included studies were presented and discussed with a service user. The structure of the meeting is described in Appendix S5 found in the Supporting Information. The service user thought that all of the index text domains needed to be tested to enable a comprehensive overview of any suspected cognitive impairment. His view was that more information on key domains would help clinicians and patients address the challenges faced by patients with MCI or early dementia. The service user raised concerns about the age of the study participants since there were no tests that assessed cognitive impairment in people over the age of 90 years. Another concern was the effect of little or no education on the ability to perform well on the test. The importance of the index tests being user-friendly and acceptable to patients was also highlighted. He also stated a preference for desktop computers over touch screen test, in case a patient had tremors. He also highlighted the importance of ensuring that the colour palette in visual components of the tests had a sharp contrast because it is likely that older people will have problems with their eyesight. He also stated that some people might become frustrated with tests that lasted longer than 40 minutes.

4 | DISCUSSION

In assessing the diagnostic accuracy of a test, an index test with high specificity is preferable for diagnosis, and high sensitivity is preferred



**FIGURE 2** Risk of bias and applicability concerns summary [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

for screening.<sup>47</sup> When patients are diagnosed with MCI or early dementia, an index test with both high sensitivity and specificity is needed to be able to appreciate a distinctive pattern of cognitive impairment in MCI and early dementia. This distinctive pattern of cognitive impairment distinguishes the cognitive impairment caused by another disease process, eg, cognitive impairment as presented in depression or HIV.

A number of studies included in this review were not conducted in samples representative of the usual clinical population in which these tests might be used (eg, patients visiting the memory clinics with a mix of MCI and dementia of various aetiologies and the “worried well” and depressed patients) but were conducted in convenience samples of patients with limited diagnoses (mostly MCI and AD). This, along with the lack of reliable evidence to support one test over the other, makes it difficult to draw a clear picture of the diagnostic accuracy of the index tests in this review.

There was some disparity in how the studies were reported; for example, all of the index tests, except 4, were used as screening tests, yet the authors reported outcomes for diagnostic accuracy. It is also not clear from reviewing the included studies whether these computerised tests ought to be used in primary or secondary care. In the United Kingdom, some primary care practices take part in “case finding” for dementia, for example, targeting “high-risk” groups (eg, older adults or patients with high vascular risk, learning disability, or Parkinson disease), and hospital staff undertake brief cognitive assessments during all acute admissions for older adults.

The pen-and-paper tests currently used in clinical practice not only help clinicians differentiate between normal cognition, MCI, and dementia<sup>20-22</sup> but also assist in staging severity of illness. The CANTAB test was the only automated test that could stage severity.<sup>35,36,42</sup> But 2 of the 3 CANTAB-PAL studies<sup>36,42</sup> had very small sample sizes

(58 and 50, respectively), and the slightly larger study<sup>35</sup> only tested the domain of visual episodic memory. The time taken to complete these computerised tests is not clear in the case of CANTAB-PAL and depending on the version of Mindstreams, ranged from 30 to 45 minutes.<sup>33</sup> In contrast, the paper-based tests range from 7 to 10 minutes in their application.<sup>20-22</sup> Concern for the time it takes to complete the tests was raised in the service user feedback; the user pointed out the possibility of people becoming frustrated with tests that lasted for more than 40 minutes, especially if they are not familiar with using technology. The data in the included papers also did not describe the time needed for training the assessor and the need for a specialist for scoring.

An important point to consider is that current diagnosis of patients with MCI and early dementia is based on clinical judgement and medical history as well as on the results of paper-based cognitive tests. Automated tests cannot be used in isolation or substituted for clinical judgement. Even with prespecified cut-off values for a particular population, any cognitive testing measure alone is insufficient to render a diagnostic classification.

None of the previously conducted relevant reviews in this area conducted a diagnostic accuracy review.<sup>23,48,49</sup> They were narrative reviews that provided a summary of the battery of tests used and rated this evidence on validity and reliability, comprehensiveness, and usability. This review focused on computerised tests that were self-administered and had a minimum level of involvement from professionals. In line with the findings of this review, the authors of the other reviews concluded that there is significant difference in automated computerised tests, and hence, they must be judged on a case-by-case basis.<sup>23</sup>

More research is required to establish stable cut-off points for each automated test used to diagnose patients with MCI or early dementia. An important consideration is testing the cut-off points in specific patient populations, for example, in patients of different age groups or education levels and from different geographical regions.

Another area for future research is providing more information on the costs of automated tests and include time for training, administration, and scoring of the different tests, as these are important factors for their use in routine clinical practice. This information is currently absent in the published studies describing automated tests used to diagnose or monitor people with MCI or early dementia. No studies reporting on outcomes relating to monitoring progression of disease could be identified, which highlights a difficulty in the current method of monitoring progression and treatment response compared with standard clinical practice.

#### 4.1 | Strengths of this review

The search strategy for this review was extensive. The methodological rigour of the review process was enhanced by the use of 2 assessors to perform citation screening, quality assessment, and data extraction/checking. All of the primary study authors were contacted and asked to fill in the contingency tables. A patient and public involvement exercise was also conducted.

#### 4.2 | Weaknesses of the review

This review is limited in part by the number of included studies for the same automated computerised test. Because of noncomparable data relating to the index test, it was not appropriate to pool the data. Another limitation with the studies is the lack of comparative results across the different domains being examined.

### 5 | CONCLUSIONS

It is difficult to draw a clear picture of the diagnostic accuracy of automated computerised tests to establish a diagnosis of MCI or early dementia in this review because there is currently insufficient evidence to support the use of one test over the other. Further research is required to examine the cut-off points for the diagnosis of MCI and early dementia when using automated tests. These test scores do not always relate with medical history and more importantly with functioning. The suitability of these tests also depends on their cost, time needed for training the assessor, time needed for the administration of the test, and the need for a specialist for scoring.

#### ACKNOWLEDGEMENTS

The authors are pleased to acknowledge the contributions of Mr Robert Little, who was the representative for public and patient involvement for this report. We are also appreciative of Prof John O'Brien (University of Cambridge) who reviewed the final draft of the report. We would also like to thank Dr Lewis Haddow (University College London) for his advice on the role of HIV in cognitive impairment.

The views and opinions expressed are those of the authors and do not necessarily reflect those of the UK Department of Health.

#### CONFLICT OF INTEREST

None declared.

#### FUNDING

The NIHR Health Technology Assessment Programme commissioned this report with project reference number 15/67/01.

#### ORCID

Rabeeh W. Aslam  <http://orcid.org/0000-0002-0916-9641>

#### REFERENCES

1. Prince M, Bryce R, Albanese E, Wimo A, Ribeiro W, Ferri CP. The global prevalence of dementia: a systematic review and meta-analysis. *Alzheimers Dement*. 2013;9(1):63-75. e62
2. World Health Organization. International Classification of Disease—classification of diseases, functioning, and disability. 2010; <http://www.who.int/classifications/icd/en/>. Accessed: 16.07.2015
3. O'Neill D, Barber RD. Reversible dementia caused by vitamin B12 deficiency. *J Am Geriatr Soc*. 1993;41(2):192-193.
4. Meador KJ. Cognitive side effects of medications. *Neurol Clin*. 1998; 16(1):141-155.



5. Muangpaisan W, Petcharat C, Srinonprasert V. Prevalence of potentially reversible conditions in dementia and mild cognitive impairment in a geriatric clinic. *Geriatr Gerontol Int*. 2012;12(1):59-64.
6. Chang CY, Silverman DH. Accuracy of early diagnosis and its impact on the management and course of Alzheimer's disease. *Expert Rev Mol Diagn*. 2004;4(1):63-69.
7. Stephan BC, Brayne C, McKeith IG, Bond J, Matthews FE; Medical Research Council Cognitive Function and Ageing Study. Mild cognitive impairment in the older population: Who is missed and does it matter? *Int J Geriatric Psychiatry*. 2008;23(8):863-871.
8. Braak H, Braak E. Evolution of neuronal changes in the course of Alzheimer's disease. In: Jellinger K, Fazekas F, Windisch M, eds. *Ageing and Dementia*. Vol. 53 Springer: Vienna; 1998:127-140.
9. Bruscoli M, Lovestone S. Is MCI really just early dementia? A systematic review of conversion studies. *Int Psychogeriatr*. 2004;16(2):129-140.
10. Busse A, Angermeyer MC, Riedel-Heller SG. Progression of mild cognitive impairment to dementia: a challenge to current thinking. *Br J Psychiatry*. 2006;189:399-404.
11. 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.
12. Gauthier S, Reisberg B, Zaudig M, et al. Mild cognitive impairment. *Lancet*. 2006;367(9518):1262-1270.
13. Feldman HH, Jacova C. Mild cognitive impairment. *Am J Geriatr Psychiatry*. 2005;13(8):645-655.
14. Forlenza OV, Diniz BS, Stella F, Teixeira AL, Gattaz WF. Mild cognitive impairment. Part 1: clinical characteristics and predictors of dementia. *Rev Bras Psiquiatr*. 2013;35(2):178-185.
15. Petersen RC, Caracciolo B, Brayne C, Gauthier S, Jelic V, Fratiglioni L. Mild cognitive impairment: a concept in evolution. *J Intern Med*. 2014;275(3):214-228.
16. Patient. Dementia. 2015; <http://patient.info/doctor/dementia-pro>. Accessed: 27.07.2015.
17. NICE. Dementia—clinical knowledge summary. 2015; <http://cks.nice.org.uk/dementia>.
18. Borson S, Scanlan J, Brush M, Vitaliano P, Dokmak A. The mini-cog: a cognitive 'vital signs' measure for dementia screening in multi-lingual elderly. *Int J Geriatr Psychiatry*. 2000;15(11):1021-1027.
19. Robinson L, Tang E, Taylor J-P. Dementia: timely diagnosis and early intervention. *BMJ: British Medical Journal*. 2015;350:
20. Kalbe E, Kessler J, Calabrese P, et al. DemTect: a new, sensitive cognitive screening test to support the diagnosis of mild cognitive impairment and early dementia. *Int J Geriatr Psychiatry*. 2004;19(2):136-143.
21. Velayudhan L, Ryu SH, Raczek M, et al. Review of brief cognitive tests for patients with suspected dementia. *Int Psychogeriatr*. 2014;26(8):1247-1262.
22. Tariq SH, Tumosa N, Chibnall JT, Perry MH 3rd, Morley JE. Comparison of the Saint Louis University mental status examination and the minimal state examination for detecting dementia and mild neurocognitive disorder—a pilot study. *Am J Geriatr Psychiatry*. 2006;14(11):900-910.
23. Wild K, Howieson D, Webbe F, Seelye A, Kaye J. Status of computerized cognitive testing in aging: a systematic review. *Alzheimers Dement*. 2008;4(6):428-437.
24. Darby D, Maruff P, Collie A, McStephen M. Mild cognitive impairment can be detected by multiple assessments in a single day. *Neurology*. 2002;59(7):1042-1046.
25. Tornatore JB, Hill E, Laboff JA, McGann ME. Self-administered screening for mild cognitive impairment: initial validation of a computerized test battery. *J Neuropsychiatry Clin Neurosci*. 2005;17(1):98-105.
26. Leeflang MM, Deeks JJ, Gatsonis C, Bossuyt PM, Cochrane Diagnostic Test Accuracy Working Group. Systematic reviews of diagnostic test accuracy. *Ann Intern Med*. 2008;149(12):889-897.
27. Kendell R, Jablensky A. Distinguishing between the validity and utility of psychiatric diagnoses. *Am J Psychiatry*. 2003;160(1):4-12.
28. 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.
29. Reitsma JB RA, Whiting P, Vlassov VV, Leeflang MMG, Deeks JJ. Assessing methodological quality. Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy 2009; Version 1.0: <http://srdta.cochrane.org/>
30. Cares AL FA, Roberts L, Turner K, Ward D Williamson T. INVOLVE: briefing notes for researchers. 2012; <http://www.invo.org.uk/resource-centre/resource-for-researchers/>. Accessed: 22.11.2015
31. Ahmed S, de Jager C, Wilcock G. A comparison of screening tools for the assessment of mild cognitive impairment: preliminary findings. *Neurocase*. 2012;18(4):336-351.
32. de Jager CA, Schrijnemaekers AC, Honey TE, Budge MM. Detection of MCI in the clinic: evaluation of the sensitivity and specificity of a computerised test battery, the Hopkins Verbal Learning Test and the MMSE. *Age Ageing*. 2009;38(4):455-460.
33. Doniger GM, Zucker DM, Schweiger A, et al. Towards practical cognitive assessment for detection of early dementia: a 30-minute computerized battery discriminates as well as longer testing. *Curr Alzheimer Res*. 2005;2(2):117-124.
34. Dwolatzky T, Whitehead V, Doniger GM, et al. Validity of a novel computerized cognitive battery for mild cognitive impairment. *BMC Geriatr*. 2003;3:4
35. Juncos-Rabadan O, Pereiro AX, Facal D, Reboredo A, Lojo-Seoane C. Do the Cambridge Neuropsychological Test Automated Battery episodic memory measures discriminate amnesic mild cognitive impairment? *Int J Geriatr Psychiatry*. 2014;29(6):602-609.
36. Junkkila J, Oja S, Laine M, Karrasch M. Applicability of the CANTAB-PAL computerized memory test in identifying amnesic mild cognitive impairment and Alzheimer's disease. *Dement Geriatr Cogn Disord*. 2012;34(2):83-89.
37. Kingsbury RP, NA, Humphreys M, Tehan G, Byrne GJA. Utility of a computerised cognitive screen in MCI and depression in an older population. *Aust J Rehabil Counsell*. 2010;16(1):14-26.
38. Kluger BM, Saunders LV, Hou W, et al. A brief computerized self-screen for dementia. *J Clin Exp Neuropsychol*. 2009;31(2):234-244.
39. Lichtenberg PJ, AS, Erlanger DM, Kaushik T, Maddens ME, Imam K. Enhancing cognitive screening in geriatric care: Use of an internet-based system. *Int J Healthcare Inf Sysinformatics*. 2006;1:47-57.
40. Maruff P, Lim YY, Darby D, et al. Clinical utility of the CogState Brief Battery in identifying cognitive impairment in mild cognitive impairment and Alzheimer's disease. *BMC Psychol*. 2013;1(1):30
41. Mundt JC, Ferber KL, Rizzo M, Greist JH. Computer-automated dementia screening using a touch-tone telephone. *Arch Intern Med*. 2001;161(20):2481-2487.
42. O'Connell H, Coen R, Kidd N, Warsi M, Chin AV, Lawlor BA. Early detection of Alzheimer's disease (AD) using the CANTAB Paired Associates Learning Test. *Int J Geriatr Psychiatry*. 2004;19(12):1207-1208.
43. Rosenthal LS, Skolasky RL, Moxley IRT, et al. A novel computerized functional assessment for human immunodeficiency virus-associated neurocognitive disorder. *J NeuroVirol*. 2013;19(5):432-441.
44. Saxton J, Morrow L, Eschman A, Archer G, Luther J, Zuccolotto A. Computer assessment of mild cognitive impairment. *Postgrad Med*. 2009;121(2):177-185.
45. Tierney MC, Naglie G, Upshur R, Moinuddin R, Charles J, Jaakkimainen RL. Feasibility and validity of the self-administered computerized assessment of mild cognitive impairment with older primary care patients. *Alzheimer Dis Assoc Disord*. 2014;28(4):311-319.



46. Vacante M, Wilcock GK, de Jager CA. Computerized adaptation of the Placing Test for early detection of both mild cognitive impairment and Alzheimer's disease. *J Clin Exp Neuropsychol*. 2013; 35(8):846-856.
47. Gilbert R, Logan S, Moyer VA, Elliott EJ. Assessing diagnostic and screening tests: part 1. *Concepts West J of Med*. 2001;174(6):405-409.
48. Tierney MC, Lerner MA. Computerized cognitive assessment in primary care to identify patients with suspected cognitive impairment. *J Alzheimers Dis*. 2010;20(3):823-832.
49. Zygouris S, Tsolaki M. Computerized cognitive testing for older adults: a review. *Am J Alzheimers Dis Other Demen*. 2015;30(1):13-28.

## SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

**How to cite this article:** Aslam R'h W, Bates V, Dundar Y, et al. A systematic review of the diagnostic accuracy of automated tests for cognitive impairment. *Int J Geriatr Psychiatry*. 2018;33:561-575. <https://doi.org/10.1002/gps.4852>