

Table 2. Characteristics of included systematic reviews

Systematic review	Participants	Comparison (Index and reference test)	Follow-up	Outcome measures	Comments	Risk of bias (per outcome measure)*
<i>Index-1 MRI</i>						
Lombardi, 2020	<p>33 studies with 3935 participants with MCI referred to tertiary centres (in memory clinics or registry data).</p> <p>The median proportion of women was 53% (range 26% to 71%) and the mean age of participants ranged from 63 to 87 years (median 73 years).</p>	<p><i>Index test (cut-off point):</i></p> <p>Volume of:</p> <ul style="list-style-type: none"> - total hippocampus (22 studies, 2209 participants) - medial temporal lobe (7 studies, 1077 participants), - total entorhinal cortex (4 studies, 529 participants), - lateral ventricles (5 studies, 1077 participants), - whole brain (4 studies, 424 participants). <p>Volume measured with either quantitative manual or automated MRI technique.</p>	<p>The mean length of clinical follow-up ranged from 1 to 7.6 years (median 2 years).</p>	<p>Sensitivity and specificity was reported for MRI as an add-on test in addition to clinical diagnosis of MCI to improve early diagnosis of dementia due to Alzheimer's disease in people with MCI. In the absence of specified thresholds, the authors used a hierarchical summary ROC curve (HSROC) model to estimate pooled accuracy measures as well as to investigate relative diagnostic odds ratios (DORs) in subgroup analyses (assuming parallel ROC curves in logits).</p>	<p>The authors could not statistically evaluate the volumes of the lateral temporal lobe, amygdala, medial temporal gyrus, or cortical grey matter assessed in small individual studies.</p>	<p>Most studies were of poor methodological quality due to risk of bias for participant selection or the index test, or both.</p>

		<p>Cut-off values were not reported.</p> <p><i>Reference test (cut-off point):</i></p> <p>Clinical follow-up diagnosis of Alzheimer's disease dementia based on NINCDS-ADRDA or DSM criteria after two or three years of follow-up.</p>				
<i>Index-2 CSF</i>						
Ritchie, 2014	<p>14 studies with 1349 participants with MCI in secondary care, community care or mixed setting. More specifically, the setting in studies was secondary care - outpatient clinic (n=12), secondary care inpatients (n=1) or mixed (n=1).</p> <p>Four studies did not report age and gender, in the other 10 studies the proportion of men ranged from 33 to 67% (median 44%) and age</p>	<p><i>Index test (cut-off point):</i></p> <p>CSF Aβ42</p> <p>Lack of common thresholds and poor reporting of thresholds.</p> <p><i>Reference test (cut-off point(s)):</i></p> <p>Clinical follow-up diagnosis of Alzheimer's disease dementia based on NINCDS-ADRDA or DSM or ICD criteria for Alzheimer's disease dementia.</p>	<p>Four studies only reported maximum time of follow-up ranging from 1- 4 years. In the other 10 studies the mean length of follow-up ranged from 0,5 to 5,2 years (median 1,84 years).</p>	<p>Sensitivity and specificity was reported for CSF Aβ42 as an add-on test in addition to clinical diagnosis of MCI to improve</p> <p>Detection of patients with MCI who would convert to Alzheimer's disease dementia or other forms of dementia over time.</p>	<p>The studies defined their own cut-off points. Thus no overall choice was made and this could be a source for heterogeneity.</p>	<p>Most studies were of poor methodological quality due to risk of bias for participant selection, the index test and/or flow/timing.</p> <p>Patient selection was often poorly reported or there was no consecutive or random enrolment.</p> <p>The threshold used for the index test was not pre-specified and the optimal cut-off level was determined from ROC analyses; therefore, the</p>

Table 2. Characteristics of included systematic reviews
Richtlijn Mild Cognitive Impairment (MCI) 2025

	ranged from 64 to 75 years (median 70 years).					accuracy of the CSF A β biomarkers reported appeared to be overestimated. In terms of flow and timing, not all patients were accounted for in the analysis or the time interval between the index test and reference standard was not appropriate (duration of follow-up was shorter than one year).
Ritchie, 2017	11 studies with MCI in secondary care, community care or mixed setting. More specifically, the setting in studies was secondary care - outpatient clinic (n=9), secondary care inpatients (n=1) or mixed (n=1). The proportion of men ranged from 31 to 66% (median 51%) and age ranged from 59 to 75 years (median 70 years).	<i>Index test (cut-off point):</i> - CSF t-tau (7 studies, 291 cases and 418 non-cases) - CSF p-tau (6 studies, 164 cases and 328 non-cases), - CSF t-tau/A β ratio (2 studies, 251 participants), - CSF p-tau/A β ratio (5 studies, 140 cases and 293 non-cases). Lack of common thresholds and poor reporting of thresholds.	In the majority of studies follow-up was in the range of one to three years. Studies reported time of follow-up differently. Follow-up ranged from 1 to 3 years in one study. In two studies the median length of follow-up was	Sensitivity and specificity was reported for CSF markers as an add-on test in addition to clinical diagnosis of MCI to improve early diagnosis of dementia due to Alzheimer's disease in people with MCI.	The studies defined their own cut-off points. Thus, no overall choice was made and this could be a source for heterogeneity.	Most studies were of poor methodological quality due to risk of bias for participant selection, the index test and/or flow/timing. Patients were not consecutively or randomly enrolled or both the sampling procedure and exclusion criteria were not or poorly described. The threshold used for the index test was not pre-specified and the optimal cut-off level was determined from ROC analyses; therefore, the

		<p><i>Reference test (cut-off point(s)):</i></p> <p>Clinical follow-up diagnosis of Alzheimer's disease dementia based on NINCDS-ADRDA or DSM or ICD criteria for Alzheimer's disease dementia.</p>	<p>1,5 to 5,2 years.</p> <p>Mean follow-up as reported by eight studies ranged from 0,5 to 5,9 years (median 2,3 years).</p>			<p>accuracy of the CSF biomarkers reported appeared to be overestimated.</p> <p>In terms of flow and timing, not all patients were accounted for in the analysis, the time interval between the index test and reference standard was not appropriate (duration of follow-up was shorter than one year) or reporting for this domain was poor.</p>
<i>Index-3 PET</i>						
Cotta Ramusino, 2024	The authors do not report characteristics such as age, sex and setting.	<p><i>Index test (cut-off point):</i></p> <p>- 18F-florbetapir or 18F-flutemetamol amyloid PET (8 studies, 1806 participants). Four studies used semi-quantitative assessment, and the other four studies combined visual and semi-quantitative assessment.</p> <p>- 18F-FDG PET (25 studies, 6803 participants). Assessments were primarily semi-</p>	Follow-up varied between 18 to 120 months but was mostly within the 24 to 70 months range.	Sensitivity and specificity was reported for amyloid PET or 18F-FDG PET as an add-on test in addition to clinical diagnosis of MCI to improve early diagnosis of dementia due to Alzheimer's disease in people with MCI.		<p>For 18F-FDG PET there is a high risk of bias in participant selection. Most studies enrolled participants non-consecutively. Also, four studies used a case-control design and likely tend to overestimate diagnostic accuracy.</p> <p>For amyloid PET there is some risk of bias in the index test, three studies did not report whether the reference standard was interpreted without knowledge of the index test result.</p>

		<p>quantitative, although 4 studies combined visual and semi-quantitative assessment.</p> <p><i>Reference test (cut-off point):</i></p> <ul style="list-style-type: none"> - Pathology (n=1 for FDG-PET, n=0 for amyloid-PET) - Clinical follow-up diagnosis of Alzheimer's disease dementia (n=22 for FDG-PET, n=7 for amyloid-PET). - Biomarker-based diagnosis (n=2 for FDG-PET, n=1 for amyloid-PET). 				
Martinez, 2017	<p>Two studies about 18F-florbetapir amyloid PET:</p> <p>1. 401 participants with a follow-up of 1,6 years and mean age of 72 years (Schreiber, 2015). A multicentre study that did not report specific data regarding setting.</p>	<p><i>Index test (cut-off point):</i></p> <p>Visual:</p> <p>a) Increased tracer uptake reduced or absent white matter/gray matter contrast in at least one cortical (frontal, parietal, temporal, occipital) region detectable on more than two adjacent scan slices (n = 1)</p>	<p>The studies reported data on 401 participants with 1.6 years of follow-up and in 47 participants with three years of follow-up. Sixty-one</p>	<p>Sensitivity and specificity was reported for 18F-florbetapir PET as an add-on test in addition to clinical diagnosis of MCI to improve early diagnosis of dementia due to Alzheimer's disease in people with MCI.</p>	<p>For Doraiswamy (2014) the manufacturer of the tracer provided financial support for the study and six authors were employees.</p>	<p>Most studies were of poor methodological quality due to unclear reporting and/or risk of bias for participant selection or the reference test, or both.</p>

Table 2. Characteristics of included systematic reviews
Richtlijn Mild Cognitive Impairment (MCI) 2025

	<p>2. 47 participants with a follow-up of 3 years, and mean age of 72 years (Doraiswamy, 2014). Participants from 21 sites in the United States of America, no data regarding the specific setting were reported.</p>	<p>b) Amyloid burden based on successive levels of florbetapir retention from 0 (no amyloid) to 4 (high levels of cortical amyloid). The median of the three visual scores was used to dichotomize participants into Aβ (-) (score, 0 to 1 point) and Aβ (+) (score, 2 to 4 points)</p> <p>(n = 2)</p> <p>Quantitative: SUVR (Standardized Uptake Volume ratio) > 1.11</p> <p><i>Reference test (cut-off point):</i></p> <p>Clinical follow-up diagnosis of Alzheimer's disease dementia based on NINCDS-ADRDA or DSM or ICD criteria for Alzheimer's disease dementia.</p> <p>Although Doraiswamy (2014) did not explicitly state usage of NINCDS-ADRDA criteria, these</p>	<p>(15.2%) participants converted at 1.6 years follow-up; nine (19.1%) participants converted at three years of follow-up.</p>			
--	---	---	--	--	--	--

Table 2. Characteristics of included systematic reviews
Richtlijn Mild Cognitive Impairment (MCI) 2025

		were used for diagnosis at baseline.				
Smailagic, 2015	<p>14 studies with MCI in university departments, clinics or research centre setting. More specifically, four studies in an outpatient setting (including a memory clinic), three studies in memory clinics, one geriatric hospital, one cognitive disorder clinic, one department of neurology, one research unit, one medical centres and two studies did not report the setting.</p> <p>Five studies with a mean age of less than 70 years. The youngest sample was 55 to 73 years, while the oldest sample was 71 to 86 years.</p> <p>The percentage of participants with</p>	<p><i>Index test (cut-off point):</i></p> <p>12 studies combined visual analysis (qualitative) and rCGMr estimations (quantitative). Four other studies performed qualitative-only analysis.</p> <p>A range of different cut-off values were used. Different brain regions were examined, and different scaling was used.</p> <p><i>Reference test (cut-off point):</i></p> <p>Clinical follow-up diagnosis of Alzheimer's disease dementia based on NINCDS-ADRDA or DSM or ICD criteria for Alzheimer's disease dementia with at least one year delay in the verification of the diagnosis.</p>	Follow-up varied substantially but was mostly within the 18 to 36 months range.	Sensitivity and specificity was reported for 18F-florbetapir PET as an add-on test in addition to clinical diagnosis of MCI to improve early diagnosis of dementia due to Alzheimer's disease in people with MCI.	<p>Sensitivity analyses indicated no impact of the interpretation of PET scan (combination of visual inspection and quantitative rCGMr evaluation versus visual interpretation</p> <p>only) and use of a prespecified threshold.</p>	<p>Very high risk of bias due to lack of blinding or lack of reporting on the index test or reference test as well as the variation of methods used.</p> <p>10 studies did not pre-specify the threshold at baseline.</p>

Table 2. Characteristics of included systematic reviews
Richtlijn Mild Cognitive Impairment (MCI) 2025

	<p>positive 18F-FDG PET scans at</p> <p>baseline ranged in the included studies from 10.5% to 74% and the percentage of those participants who converted to Alzheimer's disease</p> <p>dementia over a period of time ranged from 22% to 50%.</p>					
Zhang, 2014	<p>Nine studies with 274 participants (sample sizes ranged from 15 to 67 per study) with MCI in secondary care (n=6), tertiary care (n=1), mixed (n=1) or unclear (n=1) setting.</p> <p>Mean age ranged from 64 years to 73 years, with four studies reporting a mean age of over 70 years and the other five studies a mean age of below 70 years.</p>	<p><i>Index test (cut-off point):</i></p> <p>The studies varied markedly in how the PIB scans were done and interpreted.</p> <p>Different analytic approaches were used from quantitative binding potentials (BPs) or distribution volume ratios (DVRs) using invasive arterial sampling to semi-quantitative standardized uptake value ratios (SUVR).</p>	<p>Follow-up varied substantially but was mostly within the 18 to 36 months range.</p>	<p>Sensitivity and specificity was reported for 11C- PIB-PET scan as an add-on test in addition to clinical diagnosis of MCI to improve early diagnosis of dementia due to Alzheimer's disease in people with MCI.</p>	<p>The studies varied markedly in how the PIB scans were done and interpreted.</p> <p>Sensitivity analyses were performed to evaluate the influence of reference standard (excluding the one study with unclear reference standard) and without pre-set thresholds (one study). There was no impact on the authors' findings.</p>	<p>Very high risk of bias due to lack of blinding or lack of reporting on the index test or reference test as well as substantial variation in methods used.</p>

	<p>Gender and other patient characteristics such as APOE ε4 carriers and years of education were poorly reported.</p>	<p>Of the nine included studies:</p> <ul style="list-style-type: none"> - two did not report the threshold used and the seven others used seven different thresholds. - Six different measures for PIB retention and six different ROIs were used. - Timing of the scans and data acquisition (dynamic versus static) also varied greatly. <p><i>Reference test (cut-off point):</i></p> <p>Clinical follow-up diagnosis of Alzheimer's disease dementia based on NINCDS-ADRDA or DSM or ICD criteria for Alzheimer's disease dementia.</p>				
--	---	--	--	--	--	--

* For further details, see risk of bias table in the appendix