Received: 18 July 2025 | Revised: 3 September 2025 | Accepted: 19 September 2025

DOI: 10.1002/alz.70828

Alzheimer's & Dementia®
THE JOURNAL OF THE ALZHEIMER'S ASSOCIATION

SYSTEMATIC REVIEW

Blood-based biomarkers for detecting Alzheimer's disease pathology in cognitively impaired individuals within specialized care settings: A systematic review and meta-analysis

Sarah Pahlke[1] | **Lara A. Kahale**[2] | **Simin Mahinrad**[1] | **Bernardo Sousa-Pinto**[3,4,5] | **Rafael José Vieira**[3,4,5] | **Mary Beth McAteer**[6] | **Thomas Claessen**[7,8] | **José Contador**[9,10,11] | **Anna Hofmann**[12] | **Lindsey A. Kuchenbecker**[13] | **Luiza S. Machado**[14] | **Noëlle Warmenhoven**[15] | **Yara Yakoub**[16,17] | **Sebastian Palmqvist**[15] | **Suzanne E. Schindler**[18] | **Charlotte Teunissen**[7] | **Heather E. Whitson**[19,20] | **Henrik Zetterberg**[14,21,22,23] | **Rebecca Edelmayer**[1] | **Malavika P. Tampi**[1]

[1]Alzheimer's Association, Chicago, Illinois, USA

[2]Infectious Diseases Society of America, Arlington, Virginia, USA

[3]MEDCIDS–Department of Community Center, Information and Health Decision Sciences, Faculty of Medicine, University of Porto, Porto, Portugal

[4]CINTESIS@RISE, Centre for Health Technology and Services Research, Health Research Network, Faculty of Medicine, University of Porto, Porto, Portugal

[5]Department of Community Medicine, MEDCIDS, Information and Health Decision Sciences, Faculty of Medicine, University of Porto, Porto, Portugal

[6]Virginia Mason Medical Center, Seattle, Washington, USA

[7]Department of Laboratory Medicine, Neurochemistry Laboratory, Amsterdam UMC, Amsterdam, the Netherlands

[8]Department of Neurology, Alzheimer Center Amsterdam, Amsterdam UMC, Amsterdam, the Netherlands

[9]Barcelonaβeta Brain Research Center (BBRC), Pasqual Maragall Foundation, Barcelona, Spain

[10]Hospital del Mar Research Institute, Barcelona, Spain

[11]Servei de Neurologia, Hospital del Mar, Barcelona, Spain

[12]Department of Neurology and Knight Alzheimer Disease Research Center, Washington University, St. Louis, Missouri, USA

[13]Department of Neuroscience, Mayo Clinic, Jacksonville, Florida, USA

[14]Department of Psychiatry and Neurochemistry, The Sahlgrenska Academy, University of Gothenburg, Mölndal, Sweden

[15]Department of Clinical Sciences Malmö, Clinical Memory Research Unit, Lund University, Malmö, Sweden

[16]Douglas Mental Health University Institute, Verdun, Quebec, Canada

[17]Integrated Program in Neuroscience, Faculty of Medicine, McGill University, Montreal, Quebec, Canada

[18]Department of Neurology, Washington University in St. Louis, St. Louis, Missouri, USA

[19]Duke University School of Medicine, Durham, North Carolina, USA

[20]Durham VA Medical Center Geriatrics Research Education and Clinical Center, Durham, North Carolina, USA

[21]Clinical Neurochemistry Laboratory, Sahlgrenska University Hospital, Mölndal, Sweden

[22]Department of Pathology and Laboratory Medicine, University of Wisconsin School of Medicine and Public Health, Madison, Wisconsin, USA

[23]Wisconsin Alzheimer's Disease Research Center, University of Wisconsin School of Medicine and Public Health, University of Wisconsin–Madison, Madison, Wisconsin, USA

Correspondence
Sarah Pahlke, Alzheimer's Association, 225 North Michigan Avenue, Chicago, IL 60601, USA.
Email: scpahlke@alz.org

Abstract

BACKGROUND: This systematic review and meta-analysis aimed to assess the diagnostic test accuracy of blood-based biomarkers (BBMs) for detecting Alzheimer's disease (AD) pathology in cognitively impaired individuals in specialized care settings. The overarching goal is to inform the development of a clinical practice guideline, led by the Alzheimer's Association, for use in clinical practice.

METHODS: A systematic search of MEDLINE, Embase, and Cochrane Library was conducted from January 2019 to November 2024. Studies evaluating the diagnostic test accuracy of plasma phosphorylated tau (p-tau) and amyloid beta (Aβ) tests (p-tau217, %p-tau217, p-tau181, p-tau231, and Aβ42/Aβ40) compared to reference standard tests (cerebrospinal fluid [CSF] AD biomarkers, amyloid positron emission tomography [PET], or neuropathology) in individuals with cognitive impairment (mild cognitive impairment or dementia) in specialized care settings were included. Pooled diagnostic test accuracy measures were calculated, including sensitivity, specificity, and likelihood ratios. Across a range of pre-test probabilities, we evaluated how much a positive or a negative test result would lead to a change in the probability of having amyloid positivity (post-test probability). All analyses were conducted for each test within each biomarker. The Quality Assessment of Diagnostic Accuracy Studies 2 tool was used to evaluate the risk of bias, and the certainty of the evidence was assessed using the Grading of Recommendations, Assessment, Development, and Evaluation approach.

RESULTS: Across 49 observational studies meeting eligibility criteria, 31 different BBM tests were examined. When evaluated using a single cut-point, the diagnostic test accuracy varied considerably across tests: the pooled sensitivity ranged from 49.3% (95% confidence interval [CI]: 41.2–57.4) to 91.4% (95% CI: 86.6–94.6), and the pooled specificity ranged from 61.5% (95% CI: 45.6–75.3) to 96.7% (95% CI: 87.8–99.2). Differences in post-test probability based on a range of pre-test probabilities varied greatly across tests. Furthermore, the certainty of evidence across tests ranged from moderate to very low. Most included studies were judged to be at high risk of bias, particularly in domains related to patient selection, index test conduct, and reference standard.

CONCLUSION: This systematic review provides a comprehensive synthesis of the current evidence on the diagnostic accuracy of BBMs for detecting AD pathology in cognitively impaired individuals in specialized care settings. The findings serve as a foundation for an accompanying clinical practice guideline that provides evidence-based recommendations for BBM use in the clinical diagnostic pathway. Given continuous developments in this rapidly evolving field, ongoing evaluation will be critical to ensure the synthesized evidence and clinical guidelines remain up to date and maintain clinical relevance.

KEYWORDS
Alzheimer's disease, blood-based biomarkers, cognitive impairment, secondary care, specialty care, systematic review, tertiary care

Highlights

- This is the first comprehensive systematic review and meta-analysis to evaluate the diagnostic accuracy of blood-based biomarker (BBM) tests specifically in individuals with objective cognitive impairment seen in specialized care settings.
- Across 49 studies, BBM test performance varied widely. Pooled sensitivity ranged from 49.3% to 91.4% and specificity from 61.5% to 96.7%, depending on the analyte and assay platform.
- This review followed the Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy, the Grading of Recommendations, Assessment, Development, and Evaluation approach to assess the certainty of evidence.
- This review served as the evidence base for the Alzheimer's Association's new clinical practice guidelines on BBMs, providing structured, evidence-based guidance for implementing BBMs in the diagnostic workup of suspected Alzheimer's disease.
- The review underscores that BBM test performance is assay- and platform specific and advises clinicians and laboratory directors to interpret results in the context of the specific test used and integrate the results with a comprehensive clinical assessment.

1 | INTRODUCTION

Clinical biomarker tests used to confirm Alzheimer's disease (AD) pathology in vivo have traditionally relied on positron emission tomography (PET) with radiotracers that bind to amyloid plaques or insoluble tau aggregates, or on cerebrospinal fluid (CSF) assays that measure biomarkers like amyloid beta (Aβ)42, Aβ40, total tau (t-tau), and phosphorylated tau (p-tau).[1] Despite the high diagnostic accuracy of these tools, limitations including high costs in some countries, low accessibility, and perceived invasiveness make them impractical for widespread use in clinical practice.[2] However, in vivo confirmation of AD pathology is essential for accurately diagnosing the etiology of cognitive impairment, informing appropriate treatment and care planning, offering prognostic insights,[3,4] and identifying eligibility for emerging amyloid-targeting therapies.[5,6]

Recently, blood-based biomarkers (BBMs) have emerged as promising alternatives for detecting AD pathology. Multiple BBMs—such as Aβ42/Aβ40 and tau phosphorylated at different sites (p-tau181, p-tau217, and p-tau231)—have been shown to strongly correlate with AD pathology. Compared to PET and CSF testing, BBMs are less costly, more accessible, and more acceptable for patients. Additionally, BBMs are uniquely positioned to address the growing diagnostic demands driven by the introduction of amyloid-targeting therapies.[7]

While several BBM tests are now commercially available, their diagnostic performance differs across assays,[8,9] and their use in clinical settings remains unstandardized. Prior expert efforts have provided preliminary recommendations for BBM implementation in different settings, such as the field's first Appropriate Use Recommendations for BBMs in clinical settings and clinical trials led by the Alzheimer's Association,[10] the "Revised Criteria for Diagnosis and Staging AD, Alzheimer's Association Workgroup,"[11] the Global CEO Initiative (CEOi) on Alzheimer's Disease BBM Workgroup recommendations outlining the minimum acceptable performance standards for BBMs[7] and the World Health Organization's Preferred Product Characteristics of BBMs in diagnostics of AD.[12]

While these efforts provide a valuable foundation, they are not based on a formal systematic review of the evidence. To our knowledge, no prior systematic review has comprehensively evaluated the diagnostic performance of BBMs for detecting AD pathology in cognitively impaired individuals across available assays, nor has a clinical practice guideline (CPG) been developed using a systematic, evidence-based approach. Therefore, evidence-based CPGs to guide decision making in specific clinical scenarios are urgently needed to enable consistent standards of care. To address this need, we conducted a systematic review and meta-analysis to synthesize the available evidence and inform the development of a CPG for BBM use in the diagnostic workup of suspected AD in specialty care settings.[13] Both the systematic review and the CPG were led by the Alzheimer's Association, and the CPG was developed using the Grading of Recommendations, Assessment, Development, and Evaluation (GRADE) approach.[14]

2 | METHODS

We adhered to the guidance from the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) checklist for Diagnostic Test Accuracy[15] to report this review and followed the Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy.[16] We have not published a protocol for this systematic review. Three

methodologists (S.Pahlke, L.A.Kahale, M.P.T.) oversaw all methodological aspects of the review.

2.1 | Clinical question

To facilitate a structured and comparative approach to evidence synthesis, the associated guideline panel formulated the clinical questions using the PICO (Population or Patient, Intervention or Exposure, Comparison, and Outcome) format. The PICO question of interest for this systematic review was: "Can a certain BBM test accurately detect amyloid pathology—as measured with amyloid PET, CSF AD biomarker analysis, or neuropathology when used in individuals with objective cognitive impairment (MCI [mild cognitive impairment] or dementia, determined by clinical assessment, and/or cognitive testing) presenting to specialized care for memory disorders?"

2.2 | Eligibility criteria

The panel defined the eligibility criteria as follows. Types of studies included were: peer-reviewed observational studies, including prospective and retrospective cohort, cross-sectional, and case–control studies. Participants included individuals with cognitive impairment (dementia or MCI). Definitions of cognitive impairment were accepted as reported by the primary study authors when explicitly stated in the original publications. Studies that included both cognitively impaired and unimpaired individuals were only included if test accuracy data could be separately extracted for the cognitively impaired subgroup. This decision was made to maintain directness to the population of interest and to avoid the potential inflation of test performance in populations consisting primarily of individuals with either very low (cognitively unimpaired) or very high (AD-like dementia) amyloid levels. Interventions (index tests) included plasma p-tau and Aβ tests measuring the following analytes using any immunoassay or mass spectrometry: p-tau217, the ratio of p-tau217 to non-p-tau217 × 100 (expressed as percentage p-tau217 [%p-tau217]), p-tau181, p-tau231, and Aβ42/Aβ40. Amyloid PET imaging (using either visual read or quantitative cutoffs), neuropathological assessment, and/or CSF analysis of Aβ42/Aβ40 or combinations of Aβ42 and p-tau were included as comparators (reference tests). Diagnostic test accuracy outcomes or the raw data allowing for their computation included: sensitivity (Sn), specificity (Sp), area under the receiver operating characteristic curve (AUROC), accuracy, and likelihood ratios. Additionally, patient-important outcomes were of interest such as downstream consequences for patients including: progression of the disease, rate of decline (accomplished through early detection using accurate tests and/or routine monitoring), improvement of symptoms (symptom management), misattribution of cognitive symptoms entirely to AD (masking other underlying condition), increased anxiety related to false-positive results at the triage level, reduced anxiety when disease is ruled out, and level of pain or anxiety related to the invasiveness of the procedure. Settings included specialty care settings, such as memory clinics or research centers. Studies that assessed only the accuracy of combination of different types of analytes (e.g., p-tau217/Aβ42), rather than individual analytes or ratios of similar analytes (e.g., Aβ42/Aβ40), or those that only presented accuracy for models combining BBMs with other variables such as age, sex, and apolipoprotein E (*APOE*) genotype were excluded. Additionally, case reports, editorials, letters, comments, conference abstracts, and non-English studies were excluded.

2.3 | Database and search strategy

The panel consulted with a medical librarian (M.B.M.) to assist in drafting the search strategy for the various databases. The panel met to discuss keywords, Medical Subject Heading (MeSH) terms, and other terminology related to the identified PICO questions. A list of search strategy terms was generated by the librarian, reviewed, and edited by the panel of experts. Keywords and subject heading terms around AD and cognitive impairment, early detection, and current index testing were included. Gold-standard articles previously identified by the panel were used to verify the Sn and Sp of the final search strategy. These articles were also used to mine potential keywords and subject headings in the development of the search strategy. The search strategy is provided in the Supplementary Materials in supporting information.

MEDLINE (via PubMed), Embase (via Elsevier), and Cochrane Library databases were searched from January 1, 2019 to November 3, 2024. The start date was selected to coincide with the emergence of blood-based p-tau measurements. We checked the reference lists of any relevant systematic reviews for additional studies.

2.4 | Study selection

Pairs of reviewers (S.Pahlke, L.A.Kahale, J.C., T.C., A.H., L.A.Kuchenbecker, and N.W.) independently screened the titles and abstracts of references identified through the search strategy using Covidence.[17] Full-text articles included in this initial screening phase were then independently assessed in duplicate by the same reviewers against the eligibility criteria. Any discrepancy between the pairs of reviewers regarding study eligibility was resolved through consultation with third reviewers (S.Pahlke and L.A.Kahale).

2.5 | Data extraction

Data extraction was carried out independently by pairs of reviewers (S.Pahlke, L.A.Kahale, J.C., T.C., A.H., L.A.Kuchenbecker, N.W., Y.Y., and L.S.M.) using a piloted form in Microsoft Excel. To ensure accuracy, four additional reviewers (S.Pahlke, L.A.Kahale, L.S.M., and Y.Y.‘) conducted an independent verification of the extracted data. Disagreements were resolved through consensus discussions among the team.

The extracted variables from each study included conflicts of interest (COI), funding sources, study location, setting, study design,

number of participants, research cohort information, participant eligibility criteria, relevant risk factors (e.g., obesity), age, sex, cognitive status (MCI, dementia, or MCI and dementia combined), information about the BBM test analyte and assay, single analytical cutoff to determine test positivity, method used to determine cutoff (e.g., Youden index), reference test details, and prevalence of AD. Extracted data related to diagnostic test accuracy included Sn, Sp, accuracy, number of true positives (TPs), true negatives (TNs), false negatives (FNs), false positive (FP) results, AUROC, and likelihood ratio tests. For studies that did not report TP, TN, FP, FN, or sufficient data to construct a 2 × 2 contingency table—as well as details on BBM test cutoff values and methods, the review authors contacted the primary study authors to request the missing information. If the required data were not provided and the authors did not respond, Sn and Sp values corresponding to the Youden index were estimated from reported receiver operating characteristic (ROC) curves using WebPlotDigitizer, when available.[18]

2.5.1 | Assessment of risk of bias

We used the Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2) tool to evaluate the risk of bias (RoB) in the included diagnostic test accuracy studies.[19] QUADAS-2 assesses four domains: patient selection, index test, reference standard, and flow and timing (including the time between blood testing and reference testing). Two reviewers (Y.Y. and L.S.M.) independently applied the tool, with any disagreements resolved through discussion or consultation with a third reviewer (L.A. Kahale).

2.6 | Data synthesis

Meta-analyses of diagnostic test accuracy were conducted using bivariate random effects logistic regression models. When bivariate models did not converge due to sparse data, univariate random effects logistic regression models were fitted to independently pool Sn and Sp, which are valid in situations with few studies.[20] Summary Sn, Sp, likelihood ratios, and accuracies were calculated in both cases. Based on the pooled likelihood ratios, we modeled the impact of positive and negative test results for each biomarker assay on pre-test probabilities (i.e., across a range of pre-test probabilities, we evaluated how much a positive or a negative test result would lead to a change in the probability of having the outcome). All analyses were conducted for each analyzed test within each biomarker.

In our main analysis, we considered the results obtained (Sn, Sp, or the raw data to calculate them) at the Youden index for each test, because this was the most commonly reported method for establishing the cutoff. In studies in which such information was not explicitly reported, we retrieved it from ROC curves. Additional sensitivity analyses were conducted to assess the robustness of the data: (1) using only information reported in the manuscript text or tables (i.e., using any cutoff for the index test and not graphically retrieving information from ROC curves), (2) using information reported in the manuscript text or tables (using any cutoff for the index test) plus, when such data were missing, information retrieved from curves at Youden index, and (3) fixing cutoffs at 75% Sp, (4) fixing cutoffs at 90% Sp, and (5) fixing cutoffs at 90% Sn. For the sensitivity analyses with fixed cutoffs, data were obtained either from the primary study (reported in the manuscript or provided by authors) or graphically retrieved from ROC curves.

2.7 | Certainty of the evidence

The certainty of the evidence was assessed using the GRADE approach.[21] The certainty of the evidence can be judged as "high," "moderate," "low," or "very low." For systematic reviews of diagnostic test accuracy, the certainty of the evidence starts at "high" for observational studies[22] and can then be rated down for concerns on RoB, inconsistency, indirectness, imprecision, and publication bias. Assessment of the RoB with the QUADAS-2 tool is described above. The sensitivity analysis helped determine whether the RoB should be further downgraded by evaluating whether the pooled Sn and Sp consistently met predefined accuracy thresholds across the various scenarios. If the estimates fell below these thresholds in any of the sensitivity analyses, this was considered evidence of potential bias, warranting further downgrading. Inconsistency and imprecision were assessed separately for Sn and Sp and were evaluated following a method proposed to disentangle them by comparing the results of the random effects and fixed effects meta-analyses.[23] Inconsistency was considered serious if decision thresholds for Sn or Sp were crossed by the 95% confidence interval (CI) of the meta-analytical measure in the random effects meta-analysis but not in the fixed effects meta-analysis. Sensitivity was judged to be imprecise if the 95% CI crossed a predefined decision threshold of 0.90 in the fixed effects meta-analysis, and Sp was judged to be imprecise if the 95% CI crossed a predefined threshold of 0.90 for confirmatory contexts, or 0.75 in triaging contexts.

3 | RESULTS

3.1 | Literature search

The search resulted in 1050 single citations, of which 356 full-text papers were assessed for eligibility. Of these 356 studies, 307 were excluded, and a total of 49 studies were included[8,9,24–70] (Figure 1). Common reasons for exclusion were: data not available for the population of interest (cognitively unimpaired only, cognitively impaired and unimpaired combined, or other population) or no diagnostic test accuracy data available. The list of excluded studies with reasons for exclusion are available in Table S1 in supporting information.

3.2 | Characteristics of included studies

The 49 eligible studies assessed the validity of plasma biomarkers for identifying AD pathology defined using PET or CSF analysis. Full details

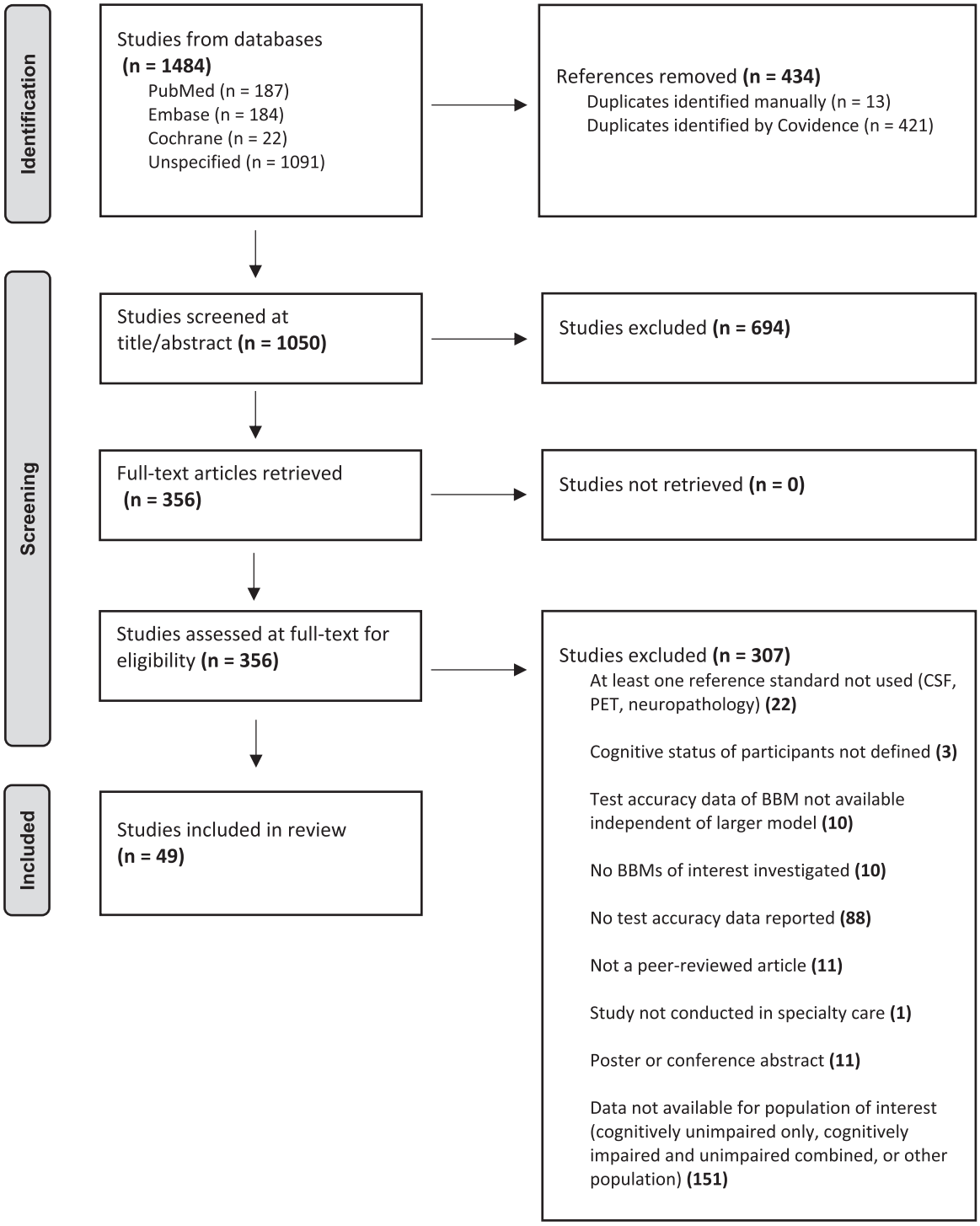


FIGURE 1 Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram of the screening and study selection process. BBM, blood-based biomarker; CSF, cerebrospinal fluid; PET, positron emission tomography.

of the characteristics of the included studies can be found in Table S4. All studies were either cross-sectional or cohort design. Several of the studies were based on data from research cohorts, with the most common ones being data from the Alzheimer's Disease Neuroimaging Initiative (ADNI; 12% of studies), the Australian Imaging Biomarkers and Lifestyle Study of Ageing (AIBL; 4%), the Advancing Research and Treatment in Frontotemporal Lobar Degeneration study (ARTFL; 4%), BioFINDER-1 (16%), BioFINDER-2 (24%), Mayo Clinic Study of Aging (MCSA; 6%), Mission AD (4%), PARIS subset of Imaging Dementia—Evidence for Amyloid Scanning study (IDEAS; 4%), and University of California San Francisco (UCSF; 4%). Reporting data from the same individuals more than once was not a concern because the same individual would not get the same index test more than once, as confirmed by the expert panelists.

In 14% of studies, demographic data (e.g., age and sex) were reported only for the cognitively impaired and unimpaired populations combined (not stratified). The mean sample size across studies was 560 participants (range: 70 to 2244). Mean age across studies ranged from 62.6 to 85.9 years, and the percentage of males ranged from 33.8% to 60.0%. Across the 32 studies that evaluated and reported *APOE ε*4 genotyping, carriers ranged from 27.1% to 56.2% of the population.

Table 1 lists the 31 BBM tests (assays and analyte combinations) that were assessed across all included studies.

3.3 | Assessment of RoB in included studies

As shown in the RoB traffic plot (Figure 2), the majority of included studies were judged to be at high RoB, particularly across domains related to patient selection, index test, and the reference standard. Commonly reported bias included insufficient information on patient enrollment (e.g., whether recruitment was consecutive or random), lack of clarity regarding blinding during the interpretation of index or reference results, and no report on the time of interval between tests. For the index test specifically, many studies did not disclose the threshold used. A smaller number of studies had unclear RoB, and only a few demonstrated low RoB across key domains. Applicability concerns were rated low for almost all studies. Most of these judgments were based on insufficient reporting of supporting evidence in the primary studies.

3.4 | Diagnostic test accuracy

Results of the meta-analyses and narrative synthesis for each BBM test can be found in Table S2 in supporting information, including pooled Sn, Sp, accuracy, and certainty of the evidence. Across all tests, pooled Sn ranged from 49.3% (95% CI: 41.2–57.4) to 91.4% (95% CI: 86.6–94.6), and pooled Sp ranged from 61.5% (95% CI: 45.6–75.3) to 96.7% (95% CI: 87.8–99.2; Figure 3). Overall, the certainty of the evidence for Sn was moderate for 19 tests, low for 9 tests, and very low for 3 tests. The certainty of the evidence for Sp was moderate for 20 tests, low for 8 tests, and very low for 3 tests.

Figures S1.1-S1.31 in supporting information display the forest plots for pooled Sn and Sp for each BBM test. Figures S2.1-S2.31 in supporting information show the relationship between pre-test probability and changes in post-test probability for each BBM test after a positive or negative test result. An interpretation guide and example are available in Box 1.

For Aβ42/Aβ40, we identified 11 different test platforms as noted in Table 1. The number of studies contributing data to each analysis ranged from one study (for immunoprecipitation mass spectrometry [IP-MS; UGOT], Immunoassay [Simoa, Quanterix single plexes], and Immunoassay [Simoa, Quanterix Neuro 3-plex A kit]) to seven studies for Immunoassay (Lumipulse, Fujirebio; Table 1). Across test platforms, the pooled Sn of Aβ42/Aβ40 varied, ranging from 59.3% (95% CI: 45.8%–71.5%) for the IP-MS (UGOT) test to 90.1% (95% CI: 71.0%–97.1%) for the Immunoassay (HISCL, Sysmex). Similarly, the pooled Sp ranged from 61.5% (95% CI: 45.6%–75.3%) for the Immunoassay (Simoa, Quanterix single plexes) to 83.3% (95% CI: 77.4%–87.9%) for the Immunoassay (HISCL, Sysmex). Overall, IP-MS (UGOT) was the test displaying the lowest accuracy (63.7%; 95% CI: 54.5%–72.0%), while Immunoassay (HISCL, Sysmex) presented the highest accuracy (85.8%; 95% CI: 78.4%–90.9%). Immunoassay (HISCL, Sysmex) was also the test in which negative or positive results could potentially lead to the highest changes in post-test probabilities (> 40 percentage points; Figure S2.11).

For p-tau181, we identified eight different test platforms (Table 1). The number of studies contributing data to each analysis ranged from 1 study (for Immunoassay [Simoa, Quanterix 4plexE] and Immunoassay [Simoa ADx Neurosciences]) to 12 studies (for Immunoassay [Simoa, Quanterix p-tau181 Advantage Kit]; Table 1). Across test platforms, the pooled Sn of p-tau181 varied, ranging from 66.9% (95% CI: 51.0%–79.7%) for the Immunoassay (S-PLEX, MSD) to 86.1% (95% CI: 71.0%–94.1%) for the Immunoassay (Lilly, MSD). The latter was also the test in which a negative result could potentially lead to the highest decrease in the post-test probability (> 40 percentage points for pre-test probabilities ≈ 70%; Figure S2.12). The pooled Sp also varied across test platforms, ranging from 67.5% (95% CI: 62.6%–72.1%) for the Immunoassay (Simoa, Quanterix UGOT) to 89.1% (95% CI: 78.8%–94.7%) for the Immunoassay (Simoa, ADx Neurosciences). The latter was also the test in which a positive result could potentially lead to the highest increase in the post-test probability (> 40 percentage points for pre-test probabilities between 20% and 40%; Figure S2.18). Overall, Immunoassay (MSD S-PLEX) was the test displaying the lowest accuracy (67.9%; 95% CI: 61.3%–73.9%), while Immunoassay (MSD Lilly) presented the highest accuracy (84.7%; 95% CI: 75.5%–90.8%).

We identified one test (Immunoassay [Simoa, Quanterix UGOT]) on p-tau231, which was assessed by two studies included in the meta-analysis (Table 1). The pooled Sn was 82.3% (95% CI: 64.4%–92.3%), the Sp was 82.1% (95% CI: 71.7%–89.2%), and the accuracy was 82.5% (95% CI: 69.9%–90.5%). For some pre-test probability values, either a negative (46.9%–85.7% of pre-test probability) or a positive result (15.2%–54.2%) was able to lead to changes of > 30 percentage points in the post-test probability (Figure S2.20).

For p-tau217, we identified seven different test platforms (Table 1). The number of studies contributing data to each analysis ranged from 1 study (for Immunoassay [S-PLEX, MSD] and IP-MS [WashU]) to 12 studies (for Immunoassay [Lilly assay, MSD]; Table 1). Across test platforms, the pooled Sn of p-tau217 varied, ranging from 49.3% (95% CI: 41.2%–57.4%) for the Immunoassay (Elecsys prototype, Roche [mid-domain]) to 91.4% (95% CI: 86.6%–94.6%) for the IP-MS (Precivity, C_2N Diagnostics). The latter was also the test in which a negative result could potentially lead to the highest decrease in the post-test probability (> 50 percentage points for pre-test probabilities of ≈ 80%; Figure S2.22). The pooled Sp ranged from 75.0% (95% CI: 64.9%–82.9%) for Immunoassay (Elecsys prototype, Roche [N-terminal]) to 96.7% (95% CI: 87.8%–99.2%) for the Immunoassay (S-PLEX, MSD). The latter was also the test in which a positive result could potentially lead

TABLE 1 List of BBM tests and studies included for each test.

Analyte	Assay	Number of eligible studies (# of participants)
Aβ42/Aβ40	Immunoprecipitation mass spectrometry (IP-MS) (Washington University [WashU])	Three studies (860) Ashton 2022 Benedet 2022 Tosun 2021
	IP-MS (Amyloid MS, Shimadzu)	One study; two cohorts (206) Niimi 2024 (two cohorts)
	IP-MS (Precivity, C_2N Diagnostics)	Four studies; five cohorts (1818) Hu 2022 (two cohorts) Meyer 2024 Schindler 2024 Devanarayan 2024
	IP-MS (University of Gothenburg [UGOT])	One study (68) Benedet 2022
	High-performance liquid chromatography differential mobility spectrometry tandem mass spectrometry (HPLC-DMS-MS/MS) (Araclon Biotech)	Two studies; four cohorts (650) Janelidze 2022 (two cohorts) Jang 2021 (MCI and dementia separate)
	Immunoassay (Simoa, Quanterix 4plexE)	Five studies (583) Benedet 2022 Bermudez 2023 Bucci 2023 Chatterjee 2023 Schindler 2024
	Immunoassay (Simoa, Quanterix single plexes)	Two studies (192) Fowler 2022 Dakterzada 2024
	Immunoassay (Simoa, Quanterix Neuro 3-plex A kit)	Two studies (206) Ni 2023 Quaresima 2024
	Immunoassay (Lumipulse, Fujirebio)	Seven studies; eight cohorts (1177) Giuffre 2024 Cecchetti 2024 Figdore 2024 Dakterzada 2024 Schindler 2024 Arranz 2024 Quaresima 2024
	Immunoassay (Elecsys, Roche)	Three studies (1151) Palmqvist 2019 Palmqvist 2023 Schindler 2024
	Immunoassay (HISCL, Sysmex)	One study; two cohorts (397) Yamashita 2022 (two cohorts)
p-tau181	Immunoassay (Lilly assay, Meso Scale Discovery [MSD])	Four studies (295) Ashton 2022 Mielke 2021 Thijssen 2020 Thijssen 2021
	Immunoassay (S-PLEX, MSD)	Two studies (243) Janelidze 2023 Dyer 2024

(Continues)

TABLE 1 (Continued)

Analyte	Assay	Number of eligible studies (# of participants)
	Immunoassay (Simoa, Quanterix p-Tau-181 Advantage Kit)	Fifteen studies (2130) Bermudez 2023 Chatterjee 2023 Kwon 2023 Mielke 2021 Ni 2023 Yang 2023 Fowler 2022 Lehmann 2023 Dakterzada 2024 Quispialaya 2024 Giacomucci 2024 Cousins 2024 Parvizi 2024 Gildengers 2024 Quaresima 2024
	Immunoassay (Simoa, Quanterix 4plexE)	One study (192) Schindler 2024
	Immunoassay (Simoa, Quanterix UGOT)	Eight studies (2311) Benedet 2022 Bucci 2023 Janelidze 2023 Karikari 2021 Karikari 2020 (two cohorts) Moscoso 2022 Palmqvist 2020 Tosun 2021
	Immunoassay (Lumipulse, Fujirebio)	Six studies; seven cohorts (693) Giuffre 2024 Janelidze 2023 Cecchetti 2024 Dakterzada 2024 Arranz 2024 Quaresima 2024
	Immunoassay (Simoa, ADx Neurosciences)	One study (135) Janelidze 2023
	Immunoassay (ElecsysTM, Roche)	Two studies (619) Palmqvist 2023 Schindler 2024
p-tau231	Immunoassay (Simoa, Quanterix UGOT)	Five studies (455) Ashton 2022 Bucci 2023 Janelidze 2023 Mielke 2021 Fowler 2022
p-tau217	IP-MS (WashU)	One study (484) Warmenhoven 2024
	IP-MS (Precivity, C_2N Diagnostics)	Two studies (775) Meyer 2024 Schindler 2024

(Continues)

TABLE 1 (Continued)

Analyte	Assay	Number of eligible studies (# of participants)
	Immunoassay (Lilly assay, MSD)	Twelve studies (2055) Ashton 2022 Brum 2023 Groot 2022 Janelidze 2022 (two cohorts) Janelidze 2023 Mattsson-Carlgren 2024 Mielke 2021 Palmqvist 2020 Thijssen 2021 Warmenhoven 2024 Niimi 2024 (two cohorts) Howe 2024
	Immunoassay (S-PLEX, MSD)	One study (108) Dyer 2024
	Immunoassay (Simoa, Quanterix Janssen)	Four studies (958) Janelidze 2023 Groot 2022 Warmenhoven 2024 Schindler 2024
	Immunoassay (Simoa, ALZpath)	Four studies (1177) Warmenhoven 2024 Figdore 2024 Schindler 2024 Gildengers 2024
	Immunoassay (Elecsys prototype, Roche [N-terminal])	One study (427) Palmqvist 2023
	Immunoassay (Elecsys prototype, Roche [mid-domain])	One study (427) Palmqvist 2023
	Immunoassay (Lumipulse, Fujirebio)	Five studies; six cohorts (1173) Arranz 2024 Cecchetti 2024 Figdore 2024 Schindler 2024 Feizpour 2024
%p-tau217	IP-MS (WashU)	Three studies; four cohorts (1371) Barthelemy 2024 Janelidze 2023 Warmenhoven 2024
	IP-MS (Precivity, C_2N Diagnostics)	Four studies (2153) Meyer 2024 Palmqvist 2024 Devanarayan 2024 Schindler 2024

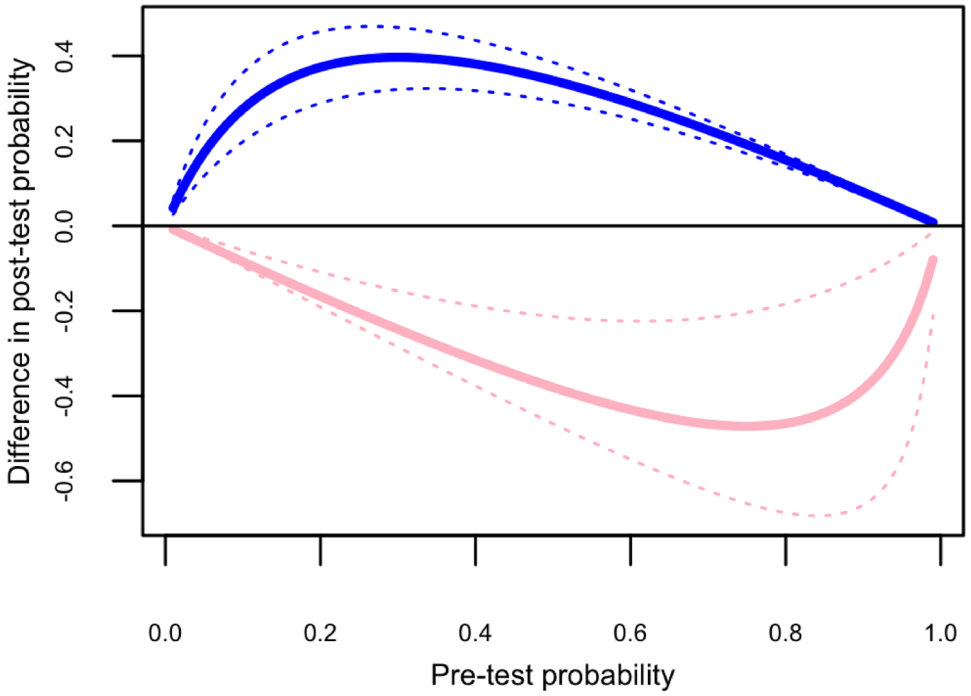
Abbreviations: Aβ, amyloid beta; BBM, blood-based biomarker; p-tau, phosphorylated tau.

to the highest increase in the post-test probability (> 60 percentage points for pre-test probabilities of ≈ 20%; Figure S2.24). Among these tests, Immunoassay (Elecsys prototype, Roche [mid-domain]) was the test displaying the lowest accuracy (63.4%; 95% CI: 57.0%–69.3%), while IP-MS (WashU) presented the highest accuracy (89.3%; 95% CI: 84.4%–92.8%).

For %p-tau217, we identified two different test platforms—IP-MS (WashU) and IP-MS (Precivity, C_2N Diagnostics)—which were evaluated in three and four studies, respectively (Table 1). The two tests displayed similar pooled sensitivities (89.5% [95% CI: 86.67%–91.79%] for IP-MS [C_2N] and 91.4% [95% CI: 88.2%–93.8%] for IP-MS [WashU]). However, pooled specificities (86.4% [95% CI: 82.1%–89.8%] for IP-MS [C_2N] and 92.2% [95% CI: 88.7%–94.7%] for IP-MS [WashU]), and accuracies (87.9% [95% CI: 84.9%–90.4%] for IP-MS [C_2N] and 91.8% [95% CI: 89.5–93.6] for IP-MS [WashU]) were less comparable. The IP-MS (WashU) test showed the largest variations in

Box 1. Interpretation guide and example of the relationship between pre-test probability and changes in post-test probability

An example plot describing the relationship between post-test probability and pre-test probability for the p-tau181 Immunoassay Lilly assay, Meso Scale Discovery (MSD) is displayed below (also available in Figure S2.12 in supporting information). The same context is applicable to all the other results presented in Figures S1.1-S1.31 in supporting information.



The horizontal axis represents the range of pre-test probabilities. The pre-test probability corresponds to the probability of an individual having amyloid pathology before being subject to the assay (i.e., based on all previously collected information available at the time of evaluation).

The vertical axis displays the absolute change in probability of having amyloid pathology if the assay result is positive or negative (post-test probability). This absolute change, computed based on likelihood ratios, is not uniform but varies depending on the pre-test probability.

For example, consider two individuals: (1) individual A, with a pre-test probability of 20% of having amyloid pathology, and (2) individual B with a pre-test probability of 80%.

For individual A, a positive result would increase the probability of amyloid pathology by 38% resulting in a post-test probability of 58% (20% + 38%). Conversely, a negative test result decreases the probability of amyloid pathology by 16%, yielding a post-test probability of 4% (20%−16%).

For individual B, a positive result increases the probability of having amyloid pathology by 16%, giving a post-test probability of 96% (80% + 16%). On the other hand, a negative result decreases the probability of having amyloid pathology by 47%, leading to a post-test probability of 33%.

post-test probabilities after either a negative or positive result (Figure S2.30). Nevertheless, for both tests, a positive or negative result led to changes of > 40 percentage points in post-test probability across certain ranges of pre-test probabilities. For IP-MS (C_2N), the values of pre-test probabilities that were more subject to change ranged from 16.4% to 43.0% (before a positive result) and 51.5% to 86.0% (before a negative result; Figure S2.31). For IP-MS (WashU), these values were 6.7% and 52.4% (before a positive result) and 47.3% to 92.1% (before a negative result; Figure S2.30).

3.5 | Sensitivity analysis

The results of our sensitivity analyses for all evaluated tests are displayed in Table S3 in supporting information. Overall, similar results to those of our main analyses were obtained when considering reported data in the manuscript text or tables at any cutoff (complemented or not by the graphical retrieval of information from ROC curves).

Fixing the Sn at 90%, we identified one test reaching a Sp of > 90% (IP-MS [WashU] for %p-tau217)—as well as six tests reaching a Sp

FIGURE 2 Risk of bias summary plot for all included studies.

between 75% and 90%. These included three tests for p-tau217 (IP-MS [Precivity, C_2N Diagnostics], immunoassay [ALZpath], and Immunoassay [Lumipulse, Fujirebio]), one test for %p-tau217 (IP-MS [C_2N]), one test for p-tau181 (Immunoassay [Lilly, MSD]), and one test for Aβ42/Aβ40 (Immunoassay [HISCL, Sysmex]).

Fixing the Sp at 75%, we identified seven tests reaching a Sn of at least 90%, including three for p-tau217 (IP-MS [Precivity, C_2N Diagnostics], Immunoassay [Lumipulse, Fujirebio], and Immunoassay [S-PLEX, MSD]), two for %p-tau217 (IP-MS [Precivity, C_2N Diagnostics] and IP-MS [WashU]), one for p-tau181 (Immunoassay [Lilly, MSD]), and one for Aβ42/Aβ40 (Immunoassay [HISCL, Sysmex]). On the other hand, there was only one test with a Sn of at least 90% when fixing the Sp at 90% (IP-MS [WashU] for %p-tau217).

The results of this systematic review will be periodically updated with the emergence of new evidence. The most up to date results will be available at [https://app.magicapp.org/#/guideline/nyO1Yj].

4 | DISCUSSION

In this systematic review, we applied a rigorous GRADE methodology to evaluate the diagnostic performance of BBMs for detecting AD pathology in cognitively impaired individuals within specialized care settings. Across 49 eligible observational studies, a total of 31 distinct BBM tests were assessed. When analyzed using a single cut-point per test, diagnostic accuracy showed substantial variability: pooled Sn ranged from 49.3% (95% CI: 41.2–57.4) to 91.4% (95% CI: 86.6–94.6), while pooled Sp ranged from 61.5% (95% CI: 45.6–75.3) to 96.7% (95% CI: 87.8–99.1). The findings of this systematic review directly inform the accompanying CPG document,[13] which provides evidence-based recommendations for the use of BBM tests in the diagnostic workup of suspected AD within specialized care settings. While the current evidence base is limited and the field is evolving, this work highlights several promising BBMs with sufficient preliminary data to support cautious clinical use.

To our knowledge, this is the first systematic review to focus on the diagnostic test accuracy of various BBMs in detecting AD pathology among cognitively impaired individuals in specialized care settings. A recent systematic review by Therriault et al. similarly assessed the diagnostic performance of plasma p-tau biomarkers for AD across both cognitively impaired and unimpaired individuals.[71] While both reviews applied rigorous methodologies, including the GRADE approach, the scope and design of the present study differ in several key respects. First, our review was restricted to literature published in the past 5 years and focused exclusively on cognitively impaired populations, aligning closely with the intended use population of the accompanying CPG. Second, this review did not include tau PET imaging as a reference standard. Third, while Therriault et al. examined additional p-tau epitopes such as p-tau205 and p-tau212, these were not included in our review due to limited data. Finally, Therriault et al. presented results by pooling data across different assays measuring the same analyte, while we intentionally did not pursue this approach. This is because the main objective of this review was to support the development of

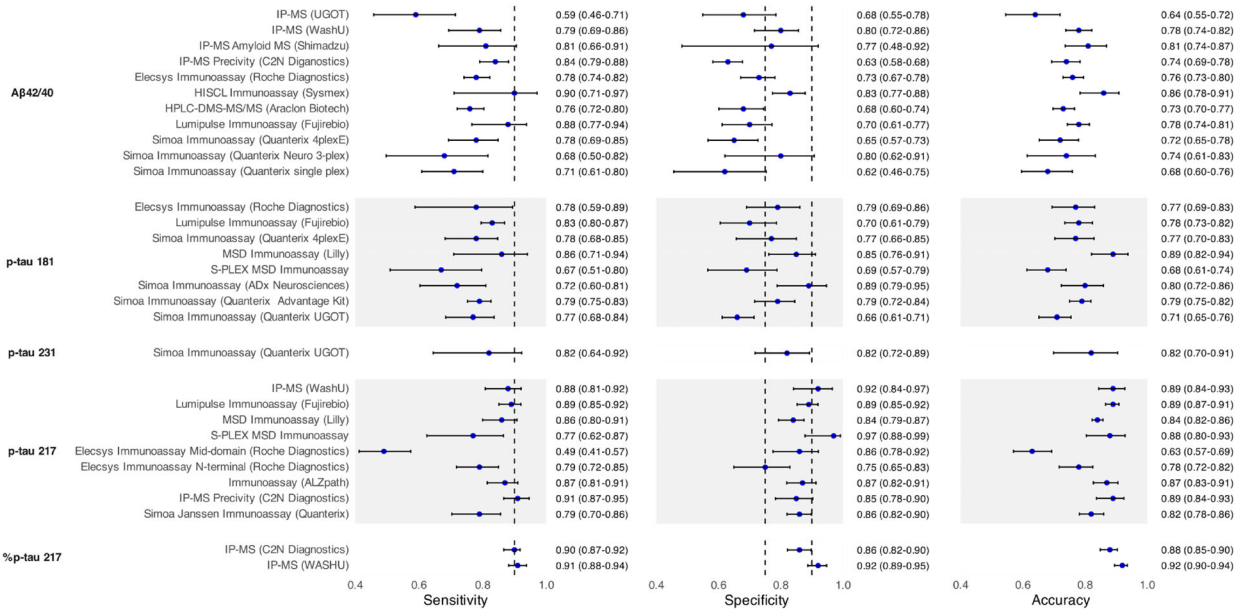


FIGURE 3 Forest plots of pooled sensitivity, specificity, and accuracy for evaluated tests. Aβ, amyloid beta; HPLC-DMS-MS/MS, high-performance liquid chromatography differential mobility spectrometry tandem mass spectrometry; IP-MS, immunoprecipitation mass spectrometry; MSD, Meso Scale Diagnostics; p-tau, phosphorylated tau.

associated CPG for which clinicians are the primary target audience.[13] From a clinical perspective, decisions are made at the level of which specific test to use, not solely at the analyte level. Because assay performance can vary substantially, pooling data across assays could be potentially misleading researchers to assume that any assay measuring a "best-performing" analyte would be sufficient. Therefore, to avoid such misinterpretation, the panel made a deliberate decision to present the results based on individual assay/test performance rather than pooling data across analytes.

When interpreting the results of this systematic review several nuances should be considered. For example, most included studies used data from single-batch plasma analysis, indicating that samples were analyzed in the same analytical run. This does not mimic a blood test in clinical practice, in which samples are often analyzed on a daily, weekly, or bi-weekly basis as patients enter the clinic and start their investigation. In this context, the coefficient of variation becomes critically important, but it has not been evaluated in this meta-analysis. High assay precision is necessary to ensure that small differences in biomarker concentrations are not due to analytical variability.[72] Especially when applying fixed diagnostic cutoffs (which is the case in clinical practice), even modest analytical imprecision can impact clinical interpretation.

Furthermore, it is crucial to consider the fold change between biomarker concentrations in amyloid-positive versus amyloid-negative individuals, and this has not been accounted for in this meta-analysis. A large fold change provides a greater buffer against minor fluctuations in assay performance or instrument calibration. For example, plasma p-tau217 typically shows a 400% to 600% increase in amyloid-positive versus amyloid-negative individuals, making it more robust to small pre-analytical or analytical changes over time and less sensitive to pre-analytical factor variation overall (i.e., temperature, centrifugation, and storage).[73] That is not the case for plasma Aβ42/Aβ40, for which usually a small decrease of ≈ 10% is seen in AD pathology positive versus negative individuals.[59,74,75] Applying a predefined cutoff and implementing it in clinical practice might result in varying results over time. This means that some assays, especially Aβ42/Aβ40 assays, may meet our analytical performance criteria when data come from a single batch plasma analysis, yet remain unsuitable for routine clinical use.

Some studies have proposed using a two-cutoff approach for BBMs, whereby an upper cutoff indicates positive results and a lower cutoff indicates negative results, with intermediate values classified as indeterminate.[29] This strategy can increase both the positive predictive value and negative predictive value, assuming that those in between the cutoffs are considered intermediate without a clear result and are excluded. However, because the majority of studies included in this systematic review reported diagnostic accuracy using a single cutoff, meta-analysis of the two-cutoff approach was not feasible. Given many plasma tests lack sufficient accuracy at a single threshold, the field is shifting toward a two-cutoff approach.[29] As more evidence is published using this approach, future updates of this review will aim to consider the two-cutoff approach to assess test accuracy for determining clinical recommendations.

It is also important to note that 84 studies were excluded from this systematic review due to combining cognitively impaired and unimpaired individuals or lacking sufficient diagnostic accuracy parameters (e.g., TP, TN, FP, and FN; required for calculating various pooled estimates), limiting their usefulness for quantitative analysis and interpretation for the target population. These exclusions reveal critical gaps in the literature and emphasize the need for future research to clearly stratify populations by cognitive status and to adopt standardized, transparent reporting practices. Researchers and industry partners should provide comprehensive diagnostic data, including detailed 2 ×

Box 2. Key methodological features for future primary studies to explicitly report

1. Report frequencies of true positive, true negative, false positive, and false negative according to cognitive status.
2. Provide a detailed description of participant enrollment (e.g., whether consecutive or random enrollment was used).
3. Clearly specify the inclusion and exclusion criteria for participants.
4. Indicate whether analyses were conducted in a blinded manner.
5. Report any thresholds used (if applicable) and/or describe the method used to determine the threshold.
6. Explicitly report numerical threshold alongside sensitivity, specificity, or diagnostic accuracy even when based on continuous predictors.
7. Prioritize use of the same reference standard test to classify all participants.
8. Use a previously validated reference standard test (e.g., amyloid positron emission tomography or cerebrospinal fluid amyloid beta [Aβ]42/Aβ40) to classify participants.
9. Report any missing data and how missing or replaced data were handled (if applicable).
10. Disclose the time interval between the reference and index tests, aiming for the shortest possible interval.

2 tables, cutoff values, methods, and methodological design features to support robust evidence synthesis and quality assessment (Box 2).

Given the variability in assay methods, performance characteristics, and diagnostic thresholds across studies, clinicians are advised to interpret test results in the context of specific assay characteristics and integrate the results with a comprehensive clinical assessment.[9,74] Laboratory directors should be aware that performance estimates are specific to one test and may not be generalizable across platforms, pre-analytical sample handling and storage before analysis is key,[73] and careful local validation and ongoing quality monitoring are essential for implementation.

Finally, in the absence of strong evidence to guide test selection or sequencing, this review and the related guideline serve as an important resource to support more consistent and thoughtful application of BBMs in clinical practice, while minimizing unnecessary or invasive procedures.

4.1 | Strengths and limitations

A key strength of this systematic review is the involvement of a multidisciplinary review team, including clinical and subject-matter experts, as well as guideline methodologists with expertise in the GRADE framework. Their input guided the development of this work, such as determining eligibility criteria, data extraction, and interpretation of findings. Other key strengths of our methodology include extracting data from ROC curves when data were not directly reported in the original studies, detailed sensitivity analyses performed to assess robustness, and the assessment of how positive or negative test results alter the probability of having AD across a range of pre-test probabilities.

One limitation of this work is that several studies have been published since our latest literature search conducted in November 2024, as they fall outside the predefined search window for this study. These studies will be considered for inclusion in future updates to maintain alignment with the fast-paced BBM evidence base. Another limitation is that most of the findings from the included studies come from specialized settings with specific participant groups, which limits their generalizability to what is typically observed in routine clinical practice among individuals with cognitive impairment. In addition, most of the available evidence to date has been generated in predominantly White populations, with limited data from other ethnic and racial groups. Future research will need to address these gaps to ensure broader applicability. Furthermore, this work focused on individual biomarkers rather than combinations of multiple biomarkers, deferring evaluation of biomarker combinations to future phases of this work.

5 | CONCLUSION

The results of this systematic review show a wide range of diagnostic test accuracy across 31 different BBM tests and provide the foundational evidence for the accompanying CPG that provides recommendations for the use of BBMs in the diagnostic workup of suspected AD within specialized care.[13] As the field evolves, continued assessment of emerging assays and combinations of biomarkers will be essential to refine diagnostic accuracy and inform future updates to this systematic review and accompanying guideline(s).

ACKNOWLEDGMENTS

The authors would like to acknowledge the panel for the associated clinical practice guideline for their contributions to defining the scope and assistance with requests for additional data. The authors also acknowledge the contributions of Ilaria Pola and Joel Simren for assistance with indexing included studies based on available data. The authors would also like to acknowledge primary study authors for providing diagnostic test accuracy data for analyses. The authors also acknowledge Farid Foroutan for providing methodological guidance related to diagnostic test accuracy. The authors also acknowledge the support and contributions of Alzheimer's Association staff including Maria Carillo, Sheena Aurora, Heather Snyder, Srilatha Raghuram, Lycia Tramujas Vasconcellos Neumann, Ozama Ismail, Emily Meyers, Emily Larson, and Nancy Cullen. This systematic review was funded by the Alzheimer's Association, but it was not involved in formulating the clinical questions.

CONFLICT OF INTEREST STATEMENT

Dr. Sebastain Palmqvist receives grant funding for research related to AD from Avid and ki elements through the Alzheimer's Drug Discovery Foundation (paid to his institution) and in the past 36 months he has received consultancy/speaker fees from BioArtic, Biogen, Eisai, Eli Lilly, Novo Nordisk, and Roche. He is supported by the National Institute on Aging (#R01AG083740), the Alzheimer's Association (#SG-23-1061717), the GHR Foundation, and the Innovative Health Initiate Joint Undertaking ("AD Riddle" #101132933). Dr. Suzanne E. Schindler has received grant funding for research directly related to this topic from the NIA (paid to her institution), and has served in speaking/advisory roles for blood-based biomarker research for Eli Lilly, Novo Nordisk, Eisai, Medscape, and universities (paid directly to her), and unpaid speaking/advisory roles with Eisai, Danaher, Eli Lilly, and the World Health Organization in the past 36 months. Dr. Charlotte Teunissen has received grant funding for research related to AD biomarkers from European Commission, Innovative Medicines Initiatives 3TR, European Platform for Neurodegenerative Diseases, EU Joint Programme—Neurodegenerative Disease Research, European Partnership on Metrology/EU Horizon Europe Research and Innovation Programme, CANTATE project funded by the Alzheimer Drug Discovery Foundation, Alzheimer Association, Michael J. Fox Foundation, Health Holland, the Dutch Research Council (ZonMW), Alzheimer Drug Discovery Foundation, The Selfridges Group Foundation, Alzheimer Netherlands (paid to her institution); has research contracts with Acumen, ADx Neurosciences, AC-Immune, Alamar, Aribio, Axon Neurosciences, Beckman-Coulter, BioConnect, Bioorchestra, Brainstorm Therapeutics, Celgene, Cognition Therapeutics, EIP Pharma, Eisai, Eli Lilly, Fujirebio, Instant Nano Biosensors, Novo Nordisk, Olink, PeopleBio, Quanterix, Roche, Toyama, and Vivoryon; and has consultancy/speaker contracts for Aribio, Biogen, Beckman-Coulter, Cognition Therapeutics, Eisai, Eli Lilly, Merck, Novo Nordisk, Novartis, Olink, Roche, Sanofi, and Veravas (paid to her institution) in the past 36 months. Dr. Heather E. Whitson receives grant funding in the diagnostic AD space from the National Institute on Aging (NIA; paid to her institution) and has served in speaking/education/advisory roles from UpToDate Inc., DrivenData Inc., ClinSTAR, American Federation for Aging Research, and universities (unrelated to the topic) in the past 36 months. Dr. Henrik Zetterberg has served in speaker/advisory/consulting roles for Denali, Apellis, Siemens, Biogen, Roche, Neumora Therapeutics Inc., Novo Nordisk, Amylyx, Enigma, WebMD Health Corp, LabCorp, Oy Medics (paid to him) in the past 36 months. Ms. Yara Yakoub has received grant funding from Alzheimer's Society of Canada Fonds de recherche du Québec in the past 36 months. Ms. Luiza S. Machado has received grant funding from Anna-Lisa och Bror Björnsons Stiftelse paid to her institution in the past 36 months. Ms. Lindsey A. Kuchenbecker, Dr. Lara A. Kahale, Dr. Anna Hoffman, Ms. Noëlle Warmenhoven, Dr. Thomas Claessen, Dr. Bernardo Sousa-Pinto, Dr. José Contador, Dr. José Vieira, and Ms. Mary Beth McAteer report no indirect or direct financial or intellectual disclosures related to the topic. Dr. Rebecca Edelmeyer, Dr. Simin Mahinrad, Ms. Sarah Pahlke, and Ms. Malavika P. Tampi are full-time employees of the Alzheimer's Association. The Alzheimer's Association funds research in numerous areas. The terms and conditions of its research awards contain a standard provision that requires the awardee to share with the Alzheimer's Association revenue derived from the licensing of intellectual property resulting from the funded research. A research award made to Washington University in St. Louis led to the development of technology for the measurement of p-tau217. This technology was subsequently licensed to C_2N Diagnostics. In accordance with the terms and condition of its award, Washington University shares a portion of the licensing fees that the university receives for this technology with the Alzheimer's Association. The Alzheimer's Association is not a party to the licensing agreement between Washington University and C_2N Diagnostics, and did not participate in its negotiation. Author disclosures are available in the supporting information.

ORCID

Sarah Pahlke https://orcid.org/0000-0003-2444-5306

REFERENCES

1. Hansson O, Blennow K, Zetterberg H, Dage J. Blood biomarkers for Alzheimer's disease in clinical practice and trials. *Nat Aging*. 2023;3:506–519. doi:10.1038/s43587-023-00403-3
2. Beach TG, Monsell SE, Phillips LE, Kukull W. Accuracy of the clinical diagnosis of Alzheimer disease at National Institute on aging Alzheimer's disease centers, 2005–2010. *J Neuropathol Exp Neurol*. 2012;71:266–273. doi:10.1097/NEN.0b013e31824b211b
3. Dickerson BC, Atri A, Clevenger C, et al. The Alzheimer's association clinical practice guideline for the diagnostic evaluation, testing, counseling, and disclosure of suspected Alzheimer's disease and related disorders (DETeCD-ADRD): executive summary of recommendations for specialty care. *Alzheimers Dement*. 2025;21:e14337. doi:10.1002/alz.14337
4. VandeVrede L, Schindler SE. Clinical use of biomarkers in the era of Alzheimer's disease treatments. *Alzheimers Dement*. 2025;21:e14201. doi:10.1002/alz.14201
5. van Dyck CH, Swanson CJ, Aisen P, et al. Lecanemab in Early Alzheimer's Disease. *N Engl J Med*. 2023;388:9–21. doi:10.1056/NEJMoa2212948
6. Sims JR, Zimmer JA, Evans CD, et al. Donanemab in early symptomatic Alzheimer disease: the TRAILBLAZER-ALZ$_2$ randomized clinical trial. *JAMA*. 2023;330:512–527. doi:10.1001/jama.2023.13239
7. Schindler SE, Galasko D, Pereira AC, et al. Acceptable performance of blood biomarker tests of amyloid pathology—recommendations from the Global CEO initiative on Alzheimer's disease. *Nat Rev Neurol*. 2024;20:426–439. doi:10.1038/s41582-024-00977-5
8. Warmenhoven N, Salvadó G, Janelidze S, et al. A comprehensive head-to-head comparison of key plasma phosphorylated tau 217 biomarker tests. *Brain J Neurol*. 2025;148:416–431. doi:10.1093/brain/awae346
9. Schindler SE, Petersen KK, Saef B, et al. Head-to-head comparison of leading blood tests for Alzheimer's disease pathology. *Alzheimers Dement*. 2024;20:8074–8096. doi:10.1002/alz.14315
10. Hansson O, Edelmayer RM, Boxer AL, et al. The Alzheimer's association appropriate use recommendations for blood biomarkers in Alzheimer's disease. *Alzheimers Dement*. 2022;18:2669–2686. doi:10.1002/alz.12756
11. Jack CR, Andrews JS, Beach TG, et al. Revised criteria for diagnosis and staging of Alzheimer's disease: Alzheimer's Association Workgroup. *Alzheimers Dement*. 2024;20(8):5143–5169. doi:10.1002/alz.13859
12. Geneva: World Health Organization. *Preferred Product Characteristics of Blood-Based Biomarker Diagnostics for Alzheimer Disease*. World

Health Organization; 2024. (Accessed July 7, 2025) https://www.who.int/publications/i/item/9789240099067
13. Palmqvist S, Whitson HE, Allen LA, et al. Alzheimer's Association Clinical Practice Guideline on the use of blood-based biomarkers in the diagnostic workup of suspected Alzheimer's disease within specialized care settings. *Alzheimers Dement*. 2025;21:e70535. doi:10.1002/alz.70535
14. Guyatt GH, Oxman AD, Vist GE, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ*. 2008;336:924–926. doi:10.1136/bmj.39489.470347.AD
15. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*. 2021;372:n71. doi:10.1136/bmj.n71
16. Deeks JJ, Bossuyt PM, Leeflang MM, Takwoingi Y (eds.). *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy*. Cochrane; 2023. Version 2.0 (updated July 2023). Available from https://training.cochrane.org/handbook-diagnostic-test-accuracy/current
17. *Covidence Systematic Review Software*. Veritas Health Innovation. Covidence n.d. https://www.covidence.org/ accessed May 7, 2025)
18. Rohatgi A. WebPlotDigitizer (Version 4.6) [Computer software]. 2021. (Accessed May 6, 2025) https://automeris.io/
19. Whiting PF, Rutjes AWS, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med*. 2011;155:529–536. doi:10.7326/0003-4819-155-8-201110180-00009
20. Takwoingi Y, Guo B, Riley RD, Deeks JJ. Performance of methods for meta-analysis of diagnostic test accuracy with few studies or sparse data. *Stat Methods Med Res*. 2017;26:1896–1911. doi:10.1177/0962280215592269
21. Guyatt GH, Oxman AD, Kunz R, et al. What is "quality of evidence" and why is it important to clinicians? *BMJ*. 2008;336:995–998. doi:10.1136/bmj.39490.551019.BE
22. Schünemann HJ, Mustafa RA, Brozek J, et al. GRADE guidelines: 21 part 2. Test accuracy: inconsistency, imprecision, publication bias, and other domains for rating the certainty of evidence and presenting it in evidence profiles and summary of findings tables. *J Clin Epidemiol*. 2020;122:142–152. doi:10.1016/j.jclinepi.2019.12.021
23. Sousa-Pinto B, Bognanni A, Vieira RJ, et al. Disentangling the GRADE domains of inconsistency and imprecision 2025; Submitted.
24. Arranz J, Zhu N, Rubio-Guerra S, et al. Diagnostic performance of plasma pTau217, pTau181, $A\beta_{1\text{-}42}$ and $A\beta_{1\text{-}40}$ in the LUMIPULSE automated platform for the detection of Alzheimer disease. *Alzheimers Res Ther*. 2024;16:139. doi:10.1186/s13195-024-01513-9
25. Ashton NJ, Janelidze S, Mattsson-Carlgren N, et al. Differential roles of Aβ42/40, p-tau231 and p-tau217 for Alzheimer's trial selection and disease monitoring. *Nat Med*. 2022;28:2555–2562. doi:10.1038/s41591-022-02074-w
26. Barthélemy NR, Salvadó G, Schindler SE, et al. Highly accurate blood test for Alzheimer's disease is similar or superior to clinical cerebrospinal fluid tests. *Nat Med*. 2024;30:1085–1095. doi:10.1038/s41591-024-02869-z
27. Benedet AL, Brum WS, Hansson O, et al. The accuracy and robustness of plasma biomarker models for amyloid PET positivity. *Alzheimers Res Ther*. 2022;14:26. doi:10.1186/s13195-021-00942-0
28. Bermudez C, Graff-Radford J, Syrjanen JA, et al. Plasma biomarkers for prediction of Alzheimer's disease neuropathologic change. *Acta Neuropathol (Berl)*. 2023;146:13–29. doi:10.1007/s00401-023-02594-w
29. Brum WS, Cullen NC, Janelidze S, et al. A two-step workflow based on plasma p-tau217 to screen for amyloid β positivity with further confirmatory testing only in uncertain cases. *Nat Aging*. 2023;3:1079–1090. doi:10.1038/s43587-023-00471-5
30. Bucci M, Bluma M, Savitcheva I, et al. Profiling of plasma biomarkers in the context of memory assessment in a tertiary memory clinic. *Transl Psychiatry*. 2023;13:268. doi:10.1038/s41398-023-02558-4
31. Cecchetti G, Agosta F, Rugarli G, et al. Diagnostic accuracy of automated Lumipulse plasma p-tau217 in Alzheimer's disease: a real-world study. *J Neurol*. 2024;271:6739–6749. doi:10.1007/s00415-024-12631-7
32. Chatterjee P, Pedrini S, Doecke JD, et al. Plasma Aβ42/40 ratio, p-tau181, GFAP, and NfL across the Alzheimer's disease continuum: a cross-sectional and longitudinal study in the AIBL cohort. *Alzheimers Dement*. 2023;19:1117–1134. doi:10.1002/alz.12724
33. Cousins KAQ, Phillips JS, Das SR, et al. Pathologic and cognitive correlates of plasma biomarkers in neurodegenerative disease. *Alzheimers Dement*. 2024;20:3889–905. doi:10.1002/alz.13777
34. Dakterzada F, Cipriani R, López-Ortega R, et al. Assessment of the correlation and diagnostic accuracy between cerebrospinal fluid and plasma Alzheimer's disease biomarkers: a comparison of the Lumipulse and Simoa platforms. *Int J Mol Sci*. 2024;25:4594. doi:10.3390/ijms25094594
35. Devanarayan V, Doherty T, Charil A, et al. Plasma pTau217 predicts continuous brain amyloid levels in preclinical and early Alzheimer's disease. *Alzheimers Dement*. 2024;20:5617–5628. doi:10.1002/alz.14073
36. Dyer AH, Dolphin H, O'Connor A, et al. Performance of plasma p-tau217 for the detection of amyloid-β positivity in a memory clinic cohort using an electrochemiluminescence immunoassay. *Alzheimers Res Ther*. 2024;16:186. doi:10.1186/s13195-024-01555-z
37. Feizpour A, Doecke JD, Doré V, et al. Detection and staging of Alzheimer's disease by plasma p-tau217 on a high throughput immunoassay platform. *EBioMedicine*. 2024;109:105405. doi:10.1016/j.ebiom.2024.105405
38. Figdore DJ, Griswold M, Bornhorst JA, et al. Optimizing cutpoints for clinical interpretation of brain amyloid status using plasma p-tau217 immunoassays. *Alzheimers Dement*. 2024;20:6506–6516. doi:10.1002/alz.14140
39. Fowler CJ, Stoops E, Rainey-Smith SR, et al. Plasma p-tau181/$A\beta_{1\text{-}42}$ ratio predicts Aβ-PET status and correlates with CSF-p-tau181/$A\beta_{1\text{-}42}$ and future cognitive decline. *Alzheimers Dement (Amst)*. 2022;14:e12375. doi:10.1002/dad2.12375
40. Giacomucci G, Mazzeo S, Ingannato A, et al. Future perspective and clinical applicability of the combined use of plasma phosphorylated-tau181 and neurofilament light chain in subjective cognitive decline and mild cognitive impairment. *Sci Rep*. 2024;14:11307. doi:10.1038/s41598-024-61655-6
41. Gildengers A, Weinstein AM, Gujral S, et al. Where do plasma biomarkers fit in with current Alzheimer's disease detection? *Am J Geriatr Psychiatry*. 2025;33:428–437. doi:10.1016/j.jagp.2024.09.015
42. Giuffrè GM, Quaranta D, Vita MG, et al. Performance of fully-automated high-throughput plasma biomarker assays for Alzheimer's disease in amnestic mild cognitive impairment subjects. *J Prev Alzheimers Dis*. 2024;11:1073–1078. doi:10.14283/jpad.2024.58
43. Groot C, Cicognola C, Bali D, et al. Diagnostic and prognostic performance to detect Alzheimer's disease and clinical progression of a novel assay for plasma p-tau217. *Alzheimers Res Ther*. 2022;14:67. doi:10.1186/s13195-022-01005-8
44. Howe MD, Britton KJ, Joyce HE, et al. Clinical application of plasma p-tau217 to assess eligibility for amyloid-lowering immunotherapy in memory clinic patients with early Alzheimer's disease. *Alzheimers Res Ther*. 2024;16:154. doi:10.1186/s13195-024-01521-9
45. Hu Y, Kirmess KM, Meyer MR, et al. Assessment of a plasma amyloid probability score to estimate amyloid positron emission tomography findings among adults with cognitive impairment. *JAMA Netw Open*. 2022;5:e228392. doi:10.1001/jamanetworkopen.2022.8392
46. Janelidze S, Palmqvist S, Leuzy A, et al. Detecting amyloid positivity in early Alzheimer's disease using combinations of plasma Aβ42/Aβ40 and p-tau. *Alzheimers Dement*. 2022;18:283–293. doi:10.1002/alz.12395
47. Jang H, Kim JS, Lee HJ, et al. Performance of the plasma Aβ42/Aβ40 ratio, measured with a novel HPLC-MS/MS method, as a biomarker of

amyloid PET status in a DPUK-KOREAN cohort. *Alzheimers Res Ther*. 2021;13:179. doi:10.1186/s13195-021-00911-7

48. Karikari TK, Pascoal TA, Ashton NJ, et al. Blood phosphorylated-tau181 as a biomarker for Alzheimer's disease: a diagnostic performance and prediction modelling study using data from four prospective cohorts. *Lancet Neurol*. 2020;19:422–433. doi:10.1016/S1474-4422(20)30071-5
49. Karikari TK, Benedet AL, Ashton NJ, et al. Diagnostic performance and prediction of clinical progression of plasma phospho-tau181 in the Alzheimer's disease neuroimaging initiative. *Mol Psychiatry*. 2021;26:429–442. doi:10.1038/s41380-020-00923-z
50. Kwon HS, Lee E, Kim H, et al. Predicting amyloid PET positivity using plasma p-tau181 and other blood-based biomarkers. *Alzheimers Dement (Amst)*. 2023;15:e12502. doi:10.1002/dad2.12502
51. Lehmann S, Schraen-Maschke S, Vidal J-S, et al. Plasma phosphorylated tau 181 predicts amyloid status and conversion to dementia stage dependent on renal function. *J Neurol Neurosurg Psychiatry*. 2023;94:411–419. doi:10.1136/jnnp-2022-330540
52. Mattsson-Carlgren N, Collij LE, Stomrud E, et al. Plasma biomarker strategy for selecting patients with Alzheimer disease for antiamyloid immunotherapies. *JAMA Neurol*. 2024;81:69. doi:10.1001/jamaneurol.2023.4596
53. Meyer MR, Kirmess KM, Eastwood S, et al. Clinical validation of the PrecivityAD2 blood test: a mass spectrometry-based test with algorithm combining %p-tau217 and Aβ42/40 ratio to identify presence of brain amyloid. *Alzheimers Dement*. 2024;20:3179–3192. doi:10.1002/alz.13764
54. Mielke MM, Frank RD, Dage JL, et al. Comparison of plasma phosphorylated tau species with amyloid and tau positron emission tomography, neurodegeneration, vascular pathology, and cognitive outcomes. *JAMA Neurol*. 2021;78:1108. doi:10.1001/jamaneurol.2021.2293
55. Moscoso A, Karikari TK, Grothe MJ, et al. CSF biomarkers and plasma p-tau181 as predictors of longitudinal tau accumulation: implications for clinical trial design. *Alzheimers Dement*. 2022;18:2614–2626. doi:10.1002/alz.12570
56. Ni M, Zhu Z-H, Gao F, et al. Plasma core Alzheimer's disease biomarkers predict amyloid deposition burden by positron emission tomography in Chinese individuals with cognitive decline. *ACS Chem Neurosci*. 2023;14:170–179. doi:10.1021/acschemneuro.2c00636
57. Niimi Y, Janelidze S, Sato K, et al. Combining plasma Aβ and p-tau217 improves detection of brain amyloid in non-demented elderly. *Alzheimers Res Ther*. 2024;16:115. doi:10.1186/s13195-024-01469-w
58. Palmqvist S, Janelidze S, Stomrud E, et al. Performance of fully automated plasma assays as screening tests for Alzheimer disease—related β-amyloid status. *JAMA Neurol*. 2019;76:1060. doi:10.1001/jamaneurol.2019.1632
59. Palmqvist S, Janelidze S, Quiroz YT, et al. Discriminative accuracy of plasma phospho-tau217 for Alzheimer disease vs other neurodegenerative disorders. *JAMA*. 2020;324:772. doi:10.1001/jama.2020.12134
60. Palmqvist S, Stomrud E, Cullen N, et al. An accurate fully automated panel of plasma biomarkers for Alzheimer's disease. *Alzheimers Dement*. 2023;19:1204–1215. doi:10.1002/alz.12751
61. Palmqvist S, Tideman P, Mattsson-Carlgren N, et al. Blood biomarkers to detect Alzheimer disease in primary care and secondary care. *JAMA*. 2024;332:1245. doi:10.1001/jama.2024.13855
62. Parvizi T, Wurm R, König T, et al. Real-world performance of plasma p-tau181 in a heterogeneous memory clinic cohort. *Ann Clin Transl Neurol*. 2024;11:1988–1998. doi:10.1002/acn3.52116
63. Quaresima V, Pilotto A, Trasciatti C, et al. Plasma p-tau181 and amyloid markers in Alzheimer's disease: a comparison between Lumipulse and SIMOA. *Neurobiol Aging*. 2024;143:30–40. doi:10.1016/j.neurobiolaging.2024.08.007
64. Quispialaya KM, Therriault J, Aliaga A, et al. Plasma phosphorylated tau181 outperforms [^{18}F]fluorodeoxyglucose positron emission tomography in the identification of early Alzheimer disease. *Eur J Neurol*. 2024;31:e16255. doi:10.1111/ene.16255
65. Thijssen EH, La Joie R, Wolf A, et al. Diagnostic value of plasma phosphorylated-tau181 in Alzheimer's disease and frontotemporal lobar degeneration. *Nat Med*. 2020;26:387–397. doi:10.1038/s41591-020-0762-2
66. Thijssen EH, La Joie R, Strom A, et al. Plasma phosphorylated tau 217 and phosphorylated-tau181 as biomarkers in Alzheimer's disease and frontotemporal lobar degeneration: a retrospective diagnostic performance study. *Lancet Neurol*. 2021;20:739–752. doi:10.1016/S1474-4422(21)00214-3
67. Tosun D, Veitch D, Aisen P, et al. Detection of β-amyloid positivity in Alzheimer's disease neuroimaging initiative participants with demographics, cognition, MRI and plasma biomarkers. *Brain Commun*. 2021;3:fcab008. doi:10.1093/braincomms/fcab008
68. Yamashita K, Miura M, Watanabe S, et al. Fully automated and highly specific plasma β-amyloid immunoassays predict β-amyloid status defined by amyloid positron emission tomography with high accuracy. *Alzheimers Res Ther*. 2022;14:86. doi:10.1186/s13195-022-01029-0
69. Yang Z, Sreenivasan K, Toledano Strom EN, et al. Clinical and biological relevance of glial fibrillary acidic protein in Alzheimer's disease. *Alzheimers Res Ther*. 2023;15:190. doi:10.1186/s13195-023-01340-4
70. Janelidze S, Bali D, Ashton NJ, et al. Head-to-head comparison of 10 plasma phospho-tau assays in prodromal Alzheimer's disease. *Brain*. 2023;146:1592–1601. doi:10.1093/brain/awac333
71. Therriault J, Brum WS, Trudel L, et al. Blood phosphorylated tau for the diagnosis of Alzheimer's disease: a systematic review and meta-analysis. *Lancet Neurol*. 2025;24(9):740–752;In press.
72. Schöll M, Vrillon A, Ikeuchi T, et al. Cutting through the noise: a narrative review of Alzheimer's disease plasma biomarkers for routine clinical use. *J Prev Alzheimers Dis*. 2025;12:100056. doi:10.1016/j.tjpad.2024.100056
73. Verberk IMW, Misdorp EO, Koelewijn J, et al. Characterization of pre-analytical sample handling effects on a panel of Alzheimer's disease—related blood-based biomarkers: results from the standardization of Alzheimer's blood biomarkers (SABB) working group. *Alzheimers Dement*. 2022;18:1484–1497. doi:10.1002/alz.12510
74. Ashton NJ, Keshavan A, Brum WS, et al. The Alzheimer's Association Global Biomarker Standardization Consortium (GBSC) plasma phospho-tau Round Robin study. *Alzheimers Dement*. 2025;21:e14508. doi:10.1002/alz.14508
75. Barthélemy NR, Horie K, Sato C, Bateman RJ. Blood plasma phosphorylated-tau isoforms track CNS change in Alzheimer's disease. *J Exp Med*. 2020;217:e20200861. doi:10.1084/jem.20200861

SUPPORTING INFORMATION

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

How to cite this article: Pahlke S, Kahale LA, Mahinrad S, et al. Blood-based biomarkers for detecting Alzheimer's disease pathology in cognitively impaired individuals within specialized care settings: A systematic review and meta-analysis. *Alzheimer's Dement*. 2025;21:e70828. https://doi.org/10.1002/alz.70828