BMC Neurology

SYSTEMATIC REVIEW Open Access

Blood-based biomarkers for early diagnosis of Alzheimer's disease: a current systematic review of diagnostic accuracy studies

Jaiyeoba-Ojigho Jennifer Efe[1], Ese Anibor[1], Priyanka Pandey[2], Oladipo Gabriel Sunday[3], David Chinaecherem Innocent[4,5*], Precious Ebube Anyakorah[4] and Okolie Ikechukwu Emmanuel[6]

Abstract

Background Alzheimer disease (AD) is the most common cause of dementia in the world and its prevalence is increasing and it has important public health consequences. Early diagnosis remains challenging due to reliance on costly and invasive methods such as positron emission tomography (PET) and cerebrospinal fluid (CSF) analysis. Blood-based biomarkers have emerged as a promising, less invasive alternatives to identify AD pathology, especially in the early and preclinical stages. This systematic review aim to evaluate the diagnostic accuracy of blood-based biomarkers for the early detection of Alzheimer's disease.

Methods An extensive literature search was performed in PubMed/MEDLINE, EMBASE, Scopus, Web of Science, and Cochrane Library on studies published since 2010 up to March 2026. The search strategies involved the use of Medical Subject Headings (MeSH), and free-text words associated with Alzheimer disease, blood-based biomarkers, and diagnostic accuracy. Peer-reviewed diagnostic accuracy studies focused on adult populations were included in line with PICO framework. The selection of studies was based on PRISMA guidelines and were critically appraised using QUADAS-2. Because of heterogeneity across the included studies, synthesis of findings was carried out using a narrative approach.

Results Six studies were included in this review. Diagnostic performance of phosphorylated tau biomarkers (p-tau181 and p-tau217) was consistently high (AUC up to 0.93), and glial fibrillary acidic protein (GFAP) also showed strong performance (AUC up to 0.87). Amyloid-β ratios (Aβ42/Aβ40) showed moderate to high accuracy, especially in preclinical detection, but had varying performance across assays (AUC 0.69–0.94). Neurofilament light chain (NfL) demonstrated moderate diagnostic value (AUC of up to 0.79) and was more predictive of progression than an early diagnosis. Combinations of biomarkers, especially those that included genetic variables like APOE genotype, were consistently more effective than individual biomarkers (AUC up to 0.92).

Conclusion Biomarkers in blood, especially p-tau and GFAP, have a strong potential for early detection of Alzheimer disease, with their combination yielding better results. Although promising, assay variation and lack of standardisation

*Correspondence:
David Chinaecherem Innocent
davidinnocent@centicini.com

Full list of author information is available at the end of the article

hinder clinical translation. To facilitate their integration into standard diagnostic practice further large-scale validation and harmonisation effort is required.

Keywords Alzheimer's disease, Blood-based biomarkers, Diagnostic accuracy, Early diagnosis, Phosphorylated tau (p-tau), Amyloid-beta (Aβ42/Aβ40), Neurofilament light chain (NfL), Glial fibrillary acidic protein (GFAP), Systematic review, QUADAS-2

Introduction

Alzheimer Disease (AD) is a neurodegenerative progressive disease and the most common cause of dementia and is characterised by a deterioration in memory, cognitive ability, and capacity to carry out everyday tasks [1]. It is pathologically characterised by the presence of extracellular amyloid-β plaques and intracellular neurofibrillary tangles consisting of hyperphosphorylated tau protein, resulting in the dysfunction of the synapses and neuronal loss [2]. AD is one of the significant health issues in the world, especially among older generations. World Health Organization [3] states that the number of individuals living with dementia worldwide is more than 55 million with Alzheimer disease comprising about 60–70% of these cases. It is estimated that the global burden is likely to increase to 139 million by 2050, and most of it is due to demographic ageing but with low- and middle-income countries showing the greatest increases [3]. This increasing prevalence highlights the necessity to develop better approaches to early detection and treatment.

Clinical assessment of the patient (patient history, cognitive testing, and functional evaluation) has long been used as a tool to diagnose Alzheimer disease and in many cases, neuroimaging and cerebrospinal fluid (CSF) biomarkers have been used to supplement the diagnosis [4]. Developed diagnostic models, including those put forward by the National Institute on Aging-Alzheimer's Association (NIA-AA), include biomarkers of amyloid deposition, tau pathology, and neurodegeneration [2, 4]. Nevertheless, the existing methods of diagnosis have a number of limitations. The neuroimaging methods like positron emission tomography (PET) are costly and not widely accessible especially in resource constrained environments. On the same note, CSF biomarkers analysis necessitates lumbar puncture which is invasive but might not be acceptable by some patients and is not feasible in large scale screening [5]. As a result, there are cases where the diagnosis is usually made during relatively late phases of the disease, which limits the effectiveness of possible treatment methods.

To respond to these challenges, blood-based biomarkers have become a promising, less invasive alternative for the early detection of AD. Blood-based biomarkers are measurable biological molecules present in plasma or serum that reflect pathological processes associated with AD, including amyloid pathology (Aβ42/Aβ40), tau pathology (p-tau181 and p-tau217), neurodegeneration (NfL), and astrocytic activation (GFAP [6, 7]. The development of ultra-sensitive techniques of assays, including single molecule array (Simoa) and mass spectrometry, has greatly facilitated the accuracy and reliability of these biomarkers in blood detection [8]. Notably, these markers have demonstrated high levels of correlations with known CSF and imaging biomarkers, and capacity to predict disease onset at preclinical stages [9]. In this way, blood-based biomarkers have significant potential in large-scale screening, early diagnosis, and disease progression monitoring both in the clinical and community environment [10].

Although these progresses have been made, a lot of variations exists in the reported accuracy of various blood-based biomarkers in various studies, such as variations in sensitivity, specificity, assay procedures, and population characteristics [8, 10, 11]. While several reviews have explored the potential of these biomarkers, there remains a lack of comprehensive synthesis specifically focused on diagnostic accuracy across diverse settings and study designs [5, 12]. Furthermore, the methodological inconsistencies in quality and reporting criteria in different diagnostic accuracy studies restrict the application of the results to clinical practice. Thus, this systematic review aims to critically review and synthesise the existing evidence on the diagnostic accuracy of blood-based biomarkers to identify Alzheimer disease at an early stage. By identifying the most scalable biomarkers and exposing gaps in the current body of knowledge, this review seeks to provide insights into clinical decision-making, inform future studies, and support the creation of scalable diagnostic tools, especially in resource-limited settings.

Methods

This review follows the Preferred Reporting Items of Systematic Review and Meta-Analysis (PRISMA) guideline [13]. This review was registered on PROSPERO as http://www.crd.york.ac.uk/PROSPERO/view/CRD420261345590.

Search strategy

The search strategy was elaborate, systematic, and designed to locate the appropriate research on the diagnostic accuracy of blood-based biomarkers of early Alzheimer disease. The systematic search of electronic databases: PubMed/MEDLINE, EMBASE, Scopus, and

Web of Science and Cochrane Library was conducted for studies carried out between January 2010 and March 2026. A search strategy that was used comprised of a combination of controlled vocabulary (Medical Subject Headings [MeSH]) and free-text words to maximise sensitivity and specificity. Key MeSH terms included *"Alzheimer Disease"*, *"Biomarkers"*, *"Blood"*, *"Plasma"*, and *"Sensitivity and Specificity"*. These were complemented by synonymous keywords such as "Alzheimer*", "blood-based biomarker*", "plasma tau", "amyloid beta", "diagnostic accuracy", "early diagnosis", and "screening". Boolean operators (AND, OR, NOT) were applied to structure the search, for example: ("Alzheimer Disease" OR "Alzheimer*") AND ("blood biomarker*" OR "plasma biomarker*" OR "serum biomarker*") AND ("diagnostic accuracy" OR "sensitivity and specificity" OR "ROC"). Where necessary truncation and wildcard symbols were applied to represent terminology variations. Included studies and relevant reviews were also searched manually to find other eligible studies.

Inclusion and exclusion criteria

This review adhered to PICO (Population, Index test, Comparator, Outcome) framework to select the inclusion and exclusion criteria, which would allow a systematic and clear selection process [14]. Table 1 shows the inclusion and exclusion criteria.

Study selection

The study selection was done by a comprehensive multistage screening procedure in line with (PRISMA) [13]. All the records retrieved were exported to Zotero reference management software, where duplicates were detected and eliminated [15]. The rest of the records were subjected to preliminary title and abstract screening by two independent reviewers so as to eliminate obviously irrelevant studies based on the eligibility criteria. Full texts of potentially eligible studies were then retrieved and assessed independently by the same reviewers for final inclusion. Any discrepancy in any step was sorted out either by discussion or involvement of a third reviewer. The reasons for exclusion during the full-text stage were recorded to promote transparency. The study selection process was presented in a PRISMA flow diagram [13], that contained the number of records identified, screened, excluded, and included in the final synthesis.

Data extraction

A standardised and piloted data extraction form was used to extract data to ensure consistency and completeness. The characteristics of the included studies (author, year, country, study design), study participant (sample size, age, clinical status), type of blood-based biomarker measured, assay procedures (e.g., Simoa, ELISA, mass spectrometry), reference standards employed, and the diagnostic accuracy scores (sensitivity, specificity, AUC, cut-off values) were extracted. The data extraction was done by two reviewers and consensus was made in case of disagreement. The data extracted were then subjected to narrative synthesis to compare the results across studies.

Quality assessment

Methodological quality and risk of bias of included studies were measured with the use of the QUADAS-2 (Quality Assessment of Diagnostic Accuracy Studies-2), which is a validated and highly recommended tool of quality assessment of studies about diagnostic accuracy [16]. QUADAS-2 evaluates studies based on four important areas namely patient selection, index test, reference standard and flow and timing with the first three areas also evaluated based on applicability issues [16]. The reason behind choosing this tool is that it is very specific to the diagnostic accuracy research and can be used to determine possible sources of bias that can influence validity and generalisability of results. The quality assessment was done by two independent reviewers and where there was disagreement, a third reviewer was consulted. The studies that were deemed to be at risk of bias in various aspects, especially in the critical areas like the wrong selection of

Table 1 Inclusion and Exclusion Criteria

PICO Element	Inclusion Criteria	Exclusion Criteria
Population (P)	Adults (≥ 18 years) with Alzheimer's disease or at risk (preclinical/MCI)	Children, non-human/animal studies, other dementia types without AD-specific analysis
Index Test (I)	Blood-based biomarkers (plasma/serum Aβ, p-tau, NfL, GFAP, etc.)	Non-blood biomarkers (e.g., CSF-only, imaging-only studies)
Comparator (C)/Reference Test	CSF biomarkers, PET imaging, or clinical diagnostic criteria	Studies without a reference standard
Outcome (O)	Diagnostic accuracy metrics (sensitivity, specificity, AUC, ROC)	Studies not reporting diagnostic accuracy outcomes
Study Design	Diagnostic accuracy studies (cross-sectional, cohort, case-control)	Reviews, qualitative research, editorials, case reports, conference abstracts
Timeframe	2010–2026	Studies published before 2010
Language	English	Non-English publications
Publication Type	Peer-reviewed journal articles	Grey literature without sufficient data

patients or absence of valid reference standard were classified to be of low quality. These researches did not make it to the final synthesis in order to increase the reliability and strength of the findings of the review. This methodology is reasonable because inclusion of methodologically weak studies can generate bias and eventually invalidate the conclusions reached on the diagnostic value of blood-based biomarkers.

Data synthesis

To summarise and interpret the findings of the studies included, a narrative synthesis approach was used [17]. The rationale behind the selection of this approach is the perceived heterogeneity of studies in the type of biomarkers (e.g., Aβ, p-tau isoforms, NfL, GFAP), assay techniques, study populations, diagnostic thresholds, and reported outcomes. This inconsistency can restrict the practicability and suitability of statistical pooling by meta-analysis. An organised transparent synthesis of findings was carried out using narrative synthesis without losing the significant contextual variations among researches [18]. The synthesis proceeded in a number of phases: to begin with, studies were categorised based on the nature of the biomarker and diagnostic use (e.g. early detection, distinguishing AD among other forms of dementias); second, important traits and findings were tabulated and compared across studies; third, trends, consistency, and inconsistency in measures of diagnostic accuracy (e.g. sensitivity, specificity, AUC) were identified and critically discussed. Quality of methodology, variability of assays and differences in the populations were also taken into consideration to explain the heterogeneity in results. This narrative approach made it possible to gain a detailed and in-depth insight into the evidence base, outline the most promising biomarkers, and find gaps and areas where more research is needed.

Result

Overview of search process

As shown in Fig. 1, a search conducted across PubMed/MEDLINE, EMBASE, Scopus, Web of Science, and Cochrane Library identified 1,128 records. Following removal of duplicates and title/abstract screening, 214 studies underwent full-text assessment. Only six studies met the predefined eligibility criteria and were included in the final synthesis. Although the number of included studies was relatively small, this reflects the highly focused nature of the review rather than limitations in the search process. Specifically, studies were required to report quantitative diagnostic accuracy outcomes (e.g., sensitivity, specificity, AUC, ROC measures) and include appropriate reference standards. Many excluded studies investigated biomarker associations, disease progression, assay development, or non-diagnostic outcomes and therefore did not satisfy the review objectives.

Characteristics of included studies

A total of six studies were included in this review. The studies were conducted in mostly Asian countries, US, Europe, Australia. Although only six studies were included, this reflects the rigorous and focused eligibility criteria employed rather than limitations of the search strategy. The review specifically targeted diagnostic accuracy studies reporting quantitative outcomes such as sensitivity, specificity, AUC, and ROC measures with appropriate reference standards. Numerous studies identified during screening evaluated biomarker associations, disease progression, assay development, or non-diagnostic outcomes and therefore did not align with the predefined review objectives and methodological requirements.

Table 2. details the characteristics of included studies.

Critical appraisal of the included studies

Table 3 shows the quality assessment of the included studies using the QUADAS-2 tool. Each study was evaluated across the four key domains: Patient Selection, Index Test, Reference Standard, and Flow & Timing, alongside Applicability Concerns. Judgements were categorised as Low Risk (L), High Risk (H), or Unclear Risk (U) based on reported methodology. The quality assessment of included studies showed an overall moderate to high methodological quality, although some variability exists. Most studies demonstrated low risk of bias in the index test and reference standard domains due to the use of validated biomarker assays and established diagnostic criteria such as amyloid PET or clinical frameworks. However, some studies showed unclear risk in patient selection and flow and timing domains, largely due to insufficient reporting of recruitment methods and timing between tests, which may introduce bias. Applicability concerns were generally low, suggesting good alignment with the review objectives.

Findings

Diagnostic accuracy of individual blood-based biomarkers

The included studies show that individual blood-based biomarkers have different degrees of diagnostics accuracy of AD, with some biomarkers more effective than others. Plasma amyloid-β (Aβ42/Aβ40 ratio) demonstrated moderate-to-high diagnostic accuracy. E.g. Udeh-Momoh et al. [19]. have reported a range of AUC values of 0.693–0.753 based on assay platform with sensitivity and specificity varying between methods with moderate discriminatory ability found in cognitively intact individuals. Conversely, Kubota et al. [23]. also found higher accuracy in Aβ42/40 (AUC up to 0.937), especially in

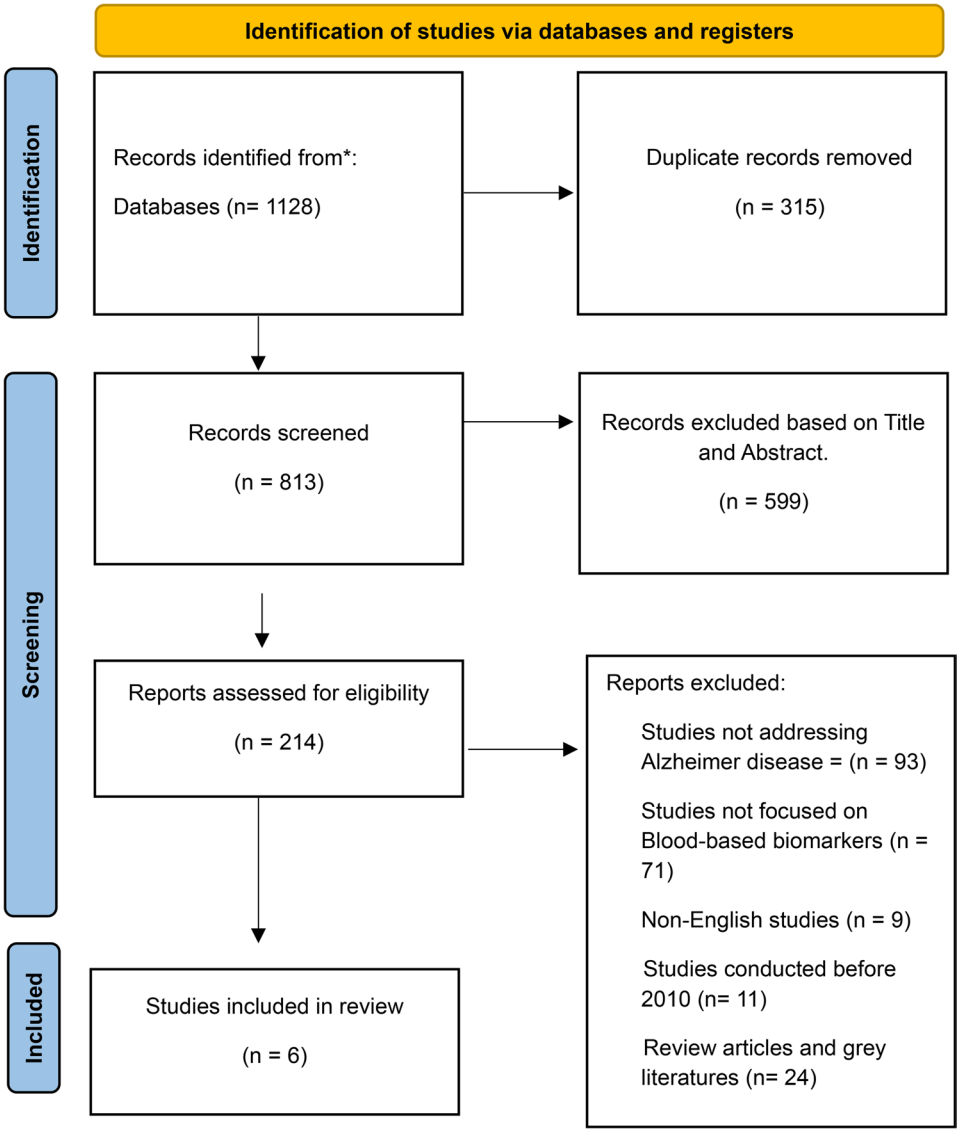

Fig. 1 PRISMA Flow Diagram [13], Demonstrating the Study Selection Process

predicting amyloid PET positivity, indicating better performance in well-characterised cohorts.

Phosphorylated tau biomarkers, especially p-tau181 and p-tau217, showed a high level of diagnostic performance persistently. Gao et al. [20]. found an AUC of 0.839 for detecting amyloid positivity in plasma p-tau, with even better results in plasma p-tau217 (AUC 0.926), which means that it is a strong method across populations. On the same note, NfL performed well in distinguishing between AD and cognitively normal individuals, with Li et al. [21]. reporting an AUC of 0.791, although its performance in early-stage detection was weaker.

Glial fibrillary acidic protein (GFAP) also came out as a potential useful marker, where Gao et al. [20]. reported an AUC of 0.871, which is better than a number of individual biomarkers. Nevertheless, there is inconsistency between studies with multi-protein panels, including those designed by Doecke et al. [8]. reporting higher diagnostic accuracy (AUC up to 0.93) and implying that single biomarkers are not as effective on their own. Overall, although single biomarkers, including p-tau and GFAP, show strong diagnostic abilities, their effectiveness is affected by disease stage, assay procedure, and population. Tau-related biomarkers appear to demonstrate more consistent diagnostic performance across studies.

Comparative performance of biomarker combinations and panels

Across the included studies, the integration of multiple biomarkers always enhanced greater diagnostic accuracy than those of single markers. Doecke et al. [8]. have shown that a 18-biomarker panel has an AUC of 0.93,

Table 2 Characteristics of Included Studies

First Author and Year	Study aim	Setting and country	Sample size and participant characteristics	Study design and methodology	Index test	Reference test/comparator	Diagnostic accuracy scores	Main findings
Doecke et al. [8]	To identify plasma biomarkers for the diagnosis of Alzheimer's disease.	Australian Imaging, Biomarker and Lifestyle (AIBL) cohort, with validation in ADNI; Australia and USA.	Discovery: 754 healthy controls and 207 AD participants in AIBL. Validation: 58 healthy controls and 112 AD participants in ADNI. Participants were older adults in a community-based prospective cohort.	Baseline plasma screening of 151 multiplexed analytes, followed by biomarker selection and validation in an independent cohort. Models were adjusted for age, sex, and APOE genotype.	A blood-based protein biomarker panel, including cortisol, pancreatic polypeptide, IGFBP2, B2M, VCAM1, carcinoembryonic antigen, homocysteine, zinc, and others.	Clinical diagnosis of AD versus cognitively healthy control status.	AIBL: sensitivity/specificity 85%/85%, AUC 93%. ADNI validation: sensitivity/specificity 80%/80%, AUC 85%. Cut-off values were not reported in the extracted text.	The study identified an 18-biomarker plasma signature for AD diagnosis; a shorter 8-biomarker panel retained much of the performance and was considered more practical for future clinical use.
Udeh-Momoh et al. [19]	To validate plasma Aβ42/Aβ40 measured by six assays for predicting amyloid positivity in cognitively unimpaired adults and to assess discrimination of CU from AD; also to examine cost-effectiveness.	ADNI; USA.	115 ADNI participants with plasma Aβ42/Aβ40 measured across six platforms; 113 had valid Aβ-PET close to plasma sampling. In the CU subgroup, 29 were Aβ-negative and 19 were Aβ-positive; the broader sample also included 55 MCI and 12 AD cases. Mean age was 78.7 years, 43% were female, 40% were APOE ε4 carriers.	Cross-sectional analysis of plasma and brain amyloid data, with comparative economic analysis. Plasma Aβ42/Aβ40 was measured by three immunoassays and three mass-spectrometry-based assays.	Plasma Aβ42/Aβ40 measured on WashU, Shimadzu, Gothenburg, ADx, Quanterix, and Roche platforms.	Aβ-PET using [18F]florbetapir within three months of plasma sampling; clinical grouping by CU, MCI, and AD.	CU Aβ+ vs Aβ−: WashU AUC 0.753, cutoff 0.125, sensitivity 0.684, specificity 0.828; Roche AUC 0.737, cutoff 0.168, sensitivity 0.842, specificity 0.586; Shimadzu AUC 0.695, cutoff 0.040, sensitivity 0.737, specificity 0.586; Quanterix AUC 0.693, cutoff 0.038, sensitivity 0.684, specificity 0.724. AD vs CU: WashU AUC 0.734.	WashU was the best-performing assay for identifying amyloid-positive cognitively unimpaired individuals. The authors concluded that plasma Aβ42/Aβ40 could be used to prescreen candidates for PET and reduce screening costs in prevention trials.
Gao et al. [20]	To evaluate blood biomarkers for AD in the Chinese Han population, including their relationship with brain atrophy and their ability to detect AV-45 PET status.	Multicenter study in China: CANDI at the First Affiliated Hospital of USTC plus five additional centers in North, West, South, and East China.	817 Chinese Han individuals from 6 centers. The CANDI cohort included 477 participants: 104 cognitively normal, 110 MCI, 208 AD, and 55 non-AD dementia. The multicenter cohort also included CSVD, CAA, and AD+CAA groups.	Multicenter cross-sectional and longitudinal study. Plasma and serum biomarkers were measured and related to PET, MRI, cognition, and APOE genotype. Analyses included ANCOVA, Spearman correlations, GLM, LOWESS/quadratic models, and ROC analysis.	Plasma Aβ42/Aβ40, plasma pTau181/pTau, plasma tTau, serum NFL, serum GFAP, and APOE genotype; the strongest model combined APOE-e4, plasma pTau, and serum GFAP.	AV-45 PET for amyloid status; MRI for brain atrophy; clinical diagnoses of CN, MCI, AD, non-AD dementia, CSVD, and CAA.	In the CANDI cohort, pTau AUC 0.839 and GFAP AUC 0.871 for amyloid status; combined demographic model AUC 0.735; NFL AUC 0.645; the logistic regression model combining APOE-e4, plasma pTau, and serum GFAP had AUC 0.915. Cut-off values were not reported in the extracted text.	Plasma pTau and serum GFAP were the most informative markers, and their combination with APOE genotype improved amyloid-status discrimination. Baseline GFAP also tracked cognitive decline and brain atrophy over time.

Table 2 (continued)

First Author and Year	Study aim	Setting and country	Sample size and participant characteristics	Study design and methodology	Index test	Reference test/comparator	Diagnostic accuracy scores	Main findings
Li et al. [21]	To assess the diagnostic and predictive value of plasma proteins in AD and aMCI, and to test whether baseline proteins predict future cognitive decline.	Shanghai elderly cohort, China. Baseline participants were recruited from the community and Shanghai Mental Health Center; a 2-year follow-up was conducted in normal older adults.	Cross-sectional: 508 normal controls, 82 aMCI, 103 AD. Follow-up: 506 normal older adults; 136 converted to aMCI and 370 remained normal. Mean age at baseline was about 69 years.	Cross-sectional and longitudinal study using Simoa plasma assays. Baseline biomarker levels were compared across groups; ROC curves were used for diagnosis; Cox regression was used for prediction of future cognitive decline.	Plasma NfL, T-tau, p-tau181, Aβ40, Aβ42, and Aβ42/40 measured by Simoa.	Clinical diagnosis of NC, aMCI, and AD; longitudinal conversion from NC to aMCI.	aMCI: Aβ40 AUC 0.558, T-tau 0.532, NfL 0.478, Aβ42 0.473, Aβ42/40 0.471. AD: NfL 0.791, p-tau181 0.693, Aβ40 0.653, T-tau 0.599, Aβ42 0.383, Aβ42/40 0.197. Cut-off values were not reported in the extracted text.	NfL was the best single discriminator for AD and also predicted future cognitive decline. The study concluded that NfL is useful for distinguishing cognitively normal from cognitively impaired individuals at the syndrome level.
Souchet et al. [22]	To identify peripheral blood biomarkers that can predict which cognitively impaired patients will develop AD dementia symptoms with high specificity.	Seven retrospective cohorts from the US, Europe, and Australia.	345 cognitively impaired individuals followed for up to 13 years: 193 MCI and 152 mild dementia at baseline; final diagnoses classified 249 as AD and 96 as non-AD brain disorders. Amyloid status was available for 82.9%, and 61.9% were amyloid-positive.	Retrospective multicenter clinical study. Biomarkers were first identified in an AAV-AD rat model and then validated in humans. The final model was trained on 70% of the data and blindly tested on the remaining 30%.	The B-HEALED model based on 19 blood biomarkers plus age, derived from a larger set of 81 blood biomarkers (45 proteins and 36 metabolites).	Final clinical diagnosis of AD versus non-AD brain disorders, with amyloid status from CSF or PET used in comparative analyses.	Internal validation at cut-off 0.76 gave specificity 93.0%, sensitivity 65.4%, AUROC 81.9%. In the combined amyloid-test + B-HEALED approach, specificity reached 100% and sensitivity 52.8%. External blind validation was consistent.	The model reduced false positives compared with amyloid testing alone and was especially useful for selecting a near-pure prodromal AD cohort for therapeutic trials.

Table 2 (continued)

First Author and Year	Study aim	Setting and country	Sample size and participant characteristics	Study design and methodology	Index test	Reference test/comparator	Diagnostic accuracy scores	Main findings
Kubota et al. [23]	To evaluate multiple plasma biomarkers for early diagnosis and staging of AD in a Japanese cohort and to assess their utility as alternatives to Aβ-PET.	Keio University Hospital and external organizations in Japan; recruitment between July 2018 and May 2024.	69 healthy controls, 13 preclinical AD, 38 AD-MCI, 44 AD-D, and 79 non-AD cognitively impaired participants. All participants were 40–85 years old and had at least 12 years of education.	Cross-sectional diagnostic accuracy study with comprehensive neuropsychological testing, Aβ PET, plasma biomarker measurement, and APOE genotyping. Comparisons were made across AD stages and against PET/Centiloid status.	Plasma Aβ42/40 by HISCL; plasma p-tau181, p-tau217, GFAP, and NfL by Simoa; ratios p-tau217/Aβ42 and p-tau217/Aβ40.	Aβ-PET visual reading and Centiloid quantification; clinical staging (HC, preclinical AD, AD-MCI, AD-D, non-AD CI).	Overall Aβ PET prediction: Aβ42/40 AUC 0.937, p-tau217 0.926, p-tau217/Aβ42 0.946. Cognitively normal: 0.968, 0.958, 0.979 respectively. Cognitively impaired: 0.919, 0.893, 0.923 respectively. Cutoff: Aβ42/40 cutoff 0.096; Centiloid threshold for Aβ42/40 separation 19.35. In the cognitively normal group, Aβ42/40 sensitivity was 0.884 and specificity 1.000.	Plasma Aβ42/40 and p-tau217 both showed strong diagnostic value, with the p-tau217/Aβ42 ratio performing best overall. Aβ42/40 appeared to detect Aβ accumulation earlier and then plateau after the preclinical stage.

which is higher than single-biomarker approaches and a smaller 8-biomarker panel still has a high percentage of accuracy, which means that a multi-marker design is both effective and feasible. Likewise, Gao et al. [20]. demonstrated that combining plasma p-tau, serum GFAP, and APOE genotype significantly enhanced diagnostic performance, with a larger AUC of 0.915 compared to lower values for individual biomarkers. This underscores the notion of combining genetic and biochemical markers so as to enhance predictive ability.

In a different study, Souchet et al. [22]. came up with the B-HEALED model, which used 19 blood biomarkers with age and achieved a result of an AUC of 0.819, with specificity reaching 93%. It is worth noting that when combined with amyloid testing, the specificity reached 100, which proves the validity of hybrid diagnostic methods. Moreover, Kubota et al. [23]. found that ratio of p-tau217 to Aβ42 achieved the highest diagnostic accuracy (AUC 0.946) outperforming each biomarker and the combination of the two simplest measures. This implies that ratio-based biomarkers can be more reflective of underlying disease pathology. Overall, it is clear that the evidence shows the superiority of the combined biomarker method over the use of single markers. Although single biomarkers are a useful source of information, combination models, especially models that take into account genetic and demographic elements, are much more accurate in their diagnosis and can be more appropriate for clinical practice.

Biomarkers in early and preclinical detection of Alzheimer's disease

The potential for blood-based biomarkers to detect the AD in its early stages or preclinical stages is the main focus of current research. Several included studies show good promise in detection of those with early pathological alterations prior to the development of clinical symptoms. As an illustration, Udeh-Momoh et al. [19]. reported that plasma Aβ42/Aβ40 could distinguish amyloid-positive from amyloid-negative cognitively unimpaired individuals, with moderate diagnostic accuracy (AUC up to 0.753), suggesting its potential as a screening tool in asymptomatic populations. Correspondingly, Kubota et al. [23]. found that Aβ42/40 and p-tau217 performed excellently in cognitively normal people with a value reaching AUC 0.968 and 0.958 respectively. It is noteworthy that tau biomarkers appeared to increase after detectable amyloid accumulation, suggesting that amyloid-related alterations may precede tau-related changes during disease progression. This finding may indicate that Aβ42/Aβ40 could have greater utility during the earliest pathological stages of AD.

Li et al. [21]. in contrast, found fairly low performance of the majority of biomarkers when used to detect mild

Table 3 Critical Appraisal Of The Included Studies

Author and Year	Patient Selection (Risk of Bias)	Index Test (Risk of Bias)	Reference Standard (Risk of Bias)	Flow & Timing (Risk of Bias)	Applicability Concerns (Patient Selection)	Applicability Concerns (Index Test)	Applicability Concerns (Reference Standard)	Overall Quality
Doecke et al. 2012 [8]	L	L	U	U	L	L	U	Moderate
Udeh-Momoh et al. 2022 [19]	U	L	L	U	L	L	L	Moderate–High
Gao et al. 2023 [20]	L	L	L	L	L	L	L	High
Li et al. 2024 [21]	L	L	U	U	L	L	U	Moderate
Souchet et al. 2024 [22]	U	L	L	U	U	L	L	Moderate
Kubota et al. 2025 [23]	L	L	L	L	L	L	L	High

cognitive impairment (AUCs generally below 0.60), indicating limitations in early stage discrimination. NfL was however found to be predictive of future cognitive decline, which means that it has a value as a prognostic and not diagnostic marker. Also, Souchet et al. [22]. established their biomarker model has a high specificity of identifying people at risk of developing AD dementia, even in cognitively impaired communities, suggesting its use in early intervention strategies. Overall, although certain biomarkers, like Aβ42/40 and p-tau, have a high probability of early detection, they respond differently in different disease stages. The evidence demonstrates the possibility of combining biomarkers or using them with other diagnostic measures in order to make early diagnosis reliable.

Variability in diagnostic performance across assay methods and populations

Significant heterogeneity in diagnostic performance was found among the studies, which was largely driven by differences assay techniques, studies population and methods of analysis. For example, Udeh-Momoh et al. [19]. found a significant difference in AUC values for plasma Aβ42/Aβ40 across six assay platforms, ranging from 0.693 to 0.753, indicating the impact of analytical approaches on diagnosis. Likewise, Kubota et al. [23]. also achieved high diagnostic accuracy with Simoa and HISCL platforms meaning that more sensitive technologies can achieve better performance. Conversely, the earlier research like Doecke et al. [8]. used multiplex assays that, as effective as they are, might not be as precise as the newer methods.

Variation was also caused by population differences. Gao et al. [20]. performed a multicenter study in a Chinese cohort and reported good performance of p-tau and GFAP but Li et al. reported lower diagnostic accuracy, in a community-based elderly cohort, especially for early stages of disease. Such variations can show genetic differences, prevalence of the disease, and clinical phenotypes. Moreover, some studies differ in the value of cut-offs and the reference standards, which makes it difficult to compare them directly. There were studies that employed amyloid PET as the gold standard and some of them employed clinical diagnosis which may have influenced reported accuracy. All in all, the results have noted the necessity of standardisation of assay procedures, diagnostic thresholds and study designs to guarantee similar and comparable findings in various locations.

Influence of assay generation on diagnostic performance

The included studies demonstrated notable differences in diagnostic performance according to assay generation. Earlier studies, such as Doecke et al. [8], primarily relied on multiplex protein panels and conventional immunoassay techniques, which provided useful diagnostic information but were constrained by lower analytical sensitivity and greater variability in biomarker quantification. More recent investigations utilised ultrasensitive technologies including Single Molecule Array (Simoa), high-sensitivity chemiluminescent assays, and mass spectrometry-based platforms. These methods permit detection of biomarkers at substantially lower concentrations and generally demonstrated improved diagnostic performance. For example, Kubota et al. [23] and Li et al. [21] employed ultrasensitive technologies and reported improved performance of p-tau and Aβ-based biomarkers. Therefore, improvements in analytical technologies may partly explain differences in biomarker performance across studies and highlight the importance of assay standardisation.

Biomarker concentrations may also vary according to biological matrix selection, particularly between plasma and serum samples. Plasma and serum differ in protein composition and processing characteristics, which may influence measured concentrations of biomarkers such as GFAP and NfL. For example, coagulation processes occurring during serum preparation may alter protein concentrations and contribute to analytical variability. Such differences may influence diagnostic thresholds and complicate direct comparisons across studies. Therefore, future investigations should consider harmonisation of sample collection and processing procedures to improve comparability and reproducibility.

Clinical utility and implementation potential of blood-based biomarkers

The studies included in this review all indicate that blood-based biomarkers could be used in clinical practice, especially as a cost-effective and more convenient substitute to current diagnostic tools. Udeh-Momoh et al. [19]. showed that plasma Aβ42/Aβ40 could serve as a pre-screening instrument in reducing the need for expensive PET imaging, decreasing clinical trials and practice expense. Likewise, Souchet et al. [22]. demonstrated that their biomarker model was able to generate substantial false positive reduction compared to amyloid tests alone, increasing the efficacy of patient selection in clinical trials. This has significant consequences on the designing of specific treatments. Gao et al. [20]. also emphasised the possibility of using the combination of biomarkers and genetic data to increase diagnostic accuracy, which supports personalised medicine approaches. In the meantime, Kubota et al. [23]. opined that plasma biomarkers might prove to be effective alternatives to PET imaging, especially where advanced imaging is not readily accessible.

Despite these benefits, some challenges still exist, such as differences in assay methods, lack of standardized cut-offs, and absence of validation across many different populations. Also, as Li et al. [21]. state, there are biomarkers with poor performance at early stages of the disease, which could limit their application as a standalone diagnostic tool. In sum, blood-based biomarkers represent a good avenue to scalable, non-invasive and affordable diagnosis of Alzheimer disease. Nevertheless, additional validation and standardisation as well as integration into clinical processes are needed to realize wide implementation.

Discussion

This systematic review has brought together recent evidence on the diagnostic accuracy of blood-based biomarkers for early detection of AD, with the results indicated that phosphorylated tau (p-tau), amyloid-β ratios (Aβ42/Aβ40), NfL, and glial fibrillary acidic protein (GFAP) are promising but inconsistent in their performance. Across the included studies, p-tau isoforms, especially p-tau181 and p-tau217, demonstrated high-quality diagnostic accuracy across the studies, which is in line with large-scale cohort studies like that of Karikari et al. [7] which is a multicentre diagnostic study undertaken across European and North American cohorts that reported high-quality discrimination of AD to other neurodegenerative diseases (AUC > 0.90). Likewise, a prospective cohort study involving the BioFINDER study in Sweden by Palmqvist et al. [6], showed that plasma Aβ42/Aβ40 is closely related to amyloid PET status and therefore can be used as a screening biomarker. These results correspond with that of this current review, especially those studies which have reported high AUC values for both Aβ and tau markers, including Kubota et al. [23] and Gao et al. [20].

Nevertheless, some disparities were noted in the performance of some biomarkers, particularly Aβ42/Aβ40. Although others were found to have a good diagnostic accuracy (e.g., Kubota et al. [23]), others like Udeh-Momoh et al. [19] found themselves performing moderately with significant differences between assay platforms. Such contradiction reflects the findings of the longitudinal cohort study by Schindler et al. [9] in the United States, which emphasised that plasma Aβ ratios may be used to predict amyloid positivity but their reliability is extremely sensitive to analytical techniques and patient characteristics. This variability highlights the fact that there is an urgent gap in the existing evidence base, i.e. the absence of standardisation in assay methods and diagnostic limits.

The findings of this review also support the growing consensus that biomarker combinations outperform individual markers. These included studies, e.g. Gao et al. [20] and Souchet et al. [22] showed that the combination of biomarkers and genetic factors (e.g. APOE genotype) can substantially enhance diagnostic accuracy. This is supported by the finding of Hampel et al. [5], a critical review of biomarker development, which stated that multimodal solutions involving biochemical, genetic, and imaging data will be the most clinically valuable diagnostic instruments. Such combinations however increase accuracy but can also make things more complex and expensive which can limit scalability in low-resource environments.

The other important observation is associated with the practicality of biomarkers in early and preclinical diagnosis. Although While Aβ42/Aβ40 and p-tau performed well in identifying amyloid positivity in cognitively intact patients, as demonstrated in Kubota et al. [23], other biomarkers like NfL were reported to have low diagnostic ability at the early stage of the disease, but could be used to predict the progression of the disease, as reported by Li et al. [21]. This difference is justified by Mattsson et al. [24] longitudinal study carried out in the ADNI cohort, which discovered that plasma NfL is a marker of neurodegeneration, but not of disease-specific pathology. This implies that the various biomarkers can be used in a complementary fashion throughout the disease spectrum.

Although advances have been made, there are still a number of methodological limitations. Many included studies exhibited heterogeneity of study design, population characteristics and reference standards, making it difficult to make a direct comparison and restricting the generalisability. The same has been expressed in wider reviews, e.g., Hampel et al. [5], which emphasised that

harmonised protocols and large-scale validation studies are required. Moreover, the majority of the research has been carried out in high-income countries, which raises doubt on the generalisability to the low- and middle-income nations where the prevalence of dementia is very fast rising. An additional concern relates to population diversity and health equity. Most diagnostic thresholds reported in existing studies have been developed predominantly in Caucasian and selected Asian populations. Genetic differences, APOE genotype distributions, comorbidity patterns, and environmental exposures may influence biomarker concentrations and diagnostic performance across populations. Consequently, diagnostic cut-offs developed in one population may not necessarily demonstrate equivalent performance in other ethnic groups. This has important implications for clinical applicability and highlights the need for large multicentre studies involving ethnically diverse populations.

Altogether, this review supports the potential for blood-based biomarkers as convenient and non-invasive methods of AD diagnostics, it also supports the necessity of standardisation, validation, and context-specific implementation strategies.

These review findings indicate that blood-based biomarkers, especially p-tau and Aβ42/Aβ40, can greatly revolutionise clinical practice by allowing for early and less invasive detection of Alzheimer disease. As primary screening tests, they may help decrease the use of costly and invasive methods of screening, PET scanning, and cerebral fluid analysis, and can enhance the accessibility, particularly where resources are limited. Nevertheless, due to the inconsistency in the diagnostic performance and the absence of standardised thresholds, these biomarkers should currently be used as adjuncts rather than replacements for established diagnostic methods. Clinicians should also put into consideration the patient characteristics and disease stage when interpreting the results since various biomarkers may be more informative at various stages of the disease progression.

From a policy perspective, the adoption of blood-based biomarkers into the healthcare structure would benefit the mass dementia diagnosis and early intervention strategies. Policy-makers ought to focus on investing in research infrastructure, standardisation of the assay techniques and preparation of clinical guidelines to use biomarkers. Also, with the increasing burden of dementia in the low and middle-income countries, there is a necessity to guarantee accessibility of these new diagnostic instruments. The regulatory approval, cost-effectiveness analysis, and training of the workforce should be addressed by policies to help ensure the safe and effective deployment of biomarker-based diagnostics.

Future research must be dedicated to large, multicentre, longitudinal studies to confirm the diagnostic accuracy of the blood-based biomarkers in different populations and different clinical conditions. Standardisation of assay methods, cut-off value and diagnostic accuracy metrics reporting is also required to enhance comparability between studies. More research on the combinations of biomarkers and using them with other modalities, including digital biomarkers and neuroimaging, is necessary. Also, research should explore the utility of these biomarkers in predicting disease progression and response to emerging therapies, as well as their applicability in real-world clinical practice.

Limitations

This systematic review has several methodological limitations that should be considered when interpreting its finding. First, although a comprehensive search strategy was used, limited to studies published in English, there was a possibility of language bias and omission of pertinent studies carried out in non-English speaking countries. Second, peer-reviewed published articles were only included, which can also lead to publication bias since the studies that do not find any significant or negative result are less likely to be published. Third, there was a lot of heterogeneity amongst the included studies in terms of study design, characteristics of participants, type of biomarkers, methods of assays, and reference standards that limited the ability to conduct a meta-analysis and required the application of narrative synthesis that may have added subjectivity to interpretation [17]. Also, not all studies provided had clear reporting of their methodology especially in areas like flow and timing and patient selection which could compromise their findings even with the use of the QUADAS-2 tool [16]. In addition, the restriction to studies published prior to 2010, albeit by the necessity to keep up with recent biomarker technology, could have missed previous basis evidence. Lastly, inconsistency in diagnostic accuracy measures as reported across studies and absence of standardised cut-offs values across studies curtailed direct comparability and result synthesis. These methodological limitations emphasise that the findings should be carefully interpreted and that standardised reporting should be emphasised in future studies of diagnostic accuracy.

Conclusion

This systematic review evaluated the diagnostic accuracy of blood-based biomarkers for early detection of AD and demonstrated that biomarkers including p-tau isoforms, Aβ42/Aβ40 ratios, NfL, and GFAP show promising diagnostic performance. Among these, p-tau and GFAP consistently demonstrated favourable performance, whereas Aβ42/Aβ40 appeared particularly valuable during early and preclinical disease stages. Evidence also suggests that combinations of biomarkers and integration with genetic

variables such as APOE genotype may improve diagnostic accuracy beyond individual markers. However, substantial heterogeneity in assay techniques, diagnostic thresholds, and study populations continues to limit comparability and clinical translation. Although blood-based biomarkers have considerable potential to improve accessibility and reduce reliance on invasive diagnostic procedures, further validation across diverse populations and standardisation of analytical methods are required before widespread implementation into routine clinical practice.

Abbreviations

Aβ	Amyloid-beta
Aβ42/Aβ40	Amyloid-beta 42/40 ratio
AD	Alzheimer's Disease
ADNI	Alzheimer's Disease Neuroimaging Initiative
AIBL	Australian Imaging, Biomarker and Lifestyle Study
APOE	Apolipoprotein E
AUC	Area Under the Curve
CAA	Cerebral Amyloid Angiopathy
CANDI	Chinese Alzheimer's Disease Neuroimaging Initiative
CN	Cognitively Normal
CSF	Cerebrospinal Fluid
CSVD	Cerebral Small Vessel Disease
CU	Cognitively Unimpaired
ELISA	Enzyme-Linked Immunosorbent Assay
GFAP	Glial Fibrillary Acidic Protein
GLM	Generalised Linear Model
HISCL	High-Sensitivity Chemiluminescence Enzyme Immunoassay
LMICs	Low- and Middle-Income Countries
LOWESS	Locally Weighted Scatterplot Smoothing
MCI	Mild Cognitive Impairment
MRI	Magnetic Resonance Imaging
NfL	Neurofilament Light Chain
NIA-AA	National Institute on Aging–Alzheimer's Association
NTDs	Neglected Tropical Diseases
PET	Positron Emission Tomography
PICO	Population, Intervention (Index Test), Comparator, Outcome
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
QUADAS-2	Quality Assessment of Diagnostic Accuracy Studies-2
ROC	Receiver Operating Characteristic
Simoa	Single Molecule Array
t-tau	Total Tau

Acknowledgements
Not Applicable.

Authors' contributions
J.O.J.E., E.A., and D.C.I. conceptualised and designed the study. J.O.J.E., E.A., and P.P. conducted the literature search and screening of studies.O.G.S. and P.E.A. performed data extraction and contributed to data synthesis.J.O.J.E. and D.C.I. led the writing of the original manuscript draft.O.I.E. and P.P. critically reviewed and revised the manuscript for important intellectual content.D.C.I. supervised the study and provided overall guidance throughout the research process. All authors read and approved the final manuscript.

Funding
No funds were received for this study.

Data availability
All data generated or analysed during this study are included in this published article and its supplementary information files.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]Department of Human Anatomy, Delta State University, Delta State, Nigeria
[2]Department of Anatomy, Era's Lucknow Medical College and Hospital, Lucknow, India
[3]Department of Human Anatomy, University of Port Harcourt, Port Harcourt, Rivers State, Nigeria
[4]Centicini Research Lab, Centicini Team LTD, Abuja, FCT, Nigeria
[5]Department of Epidemiology, School of Public Health, University of Port Harcourt, Port Harcourt, Rivers State, Nigeria
[6]Department of Nursing and Midwifery, University of Central Lancashire, Preston, United Kingdom

Received: 10 May 2026 / Accepted: 2 June 2026
Published online: 05 June 2026

References

1. Lane CA, Hardy J, Schott JM. Alzheimer's disease. Eur J Neurol. 2018;25(1):59–70.
2. Jack CR Jr, Bennett DA, Blennow K, Carrillo MC, Dunn B, Haeberlein SB, et al. NIA-AA research framework: toward a biological definition of Alzheimer's disease. Alzheimers Dement. 2018;14(4):535–62.
3. World Health Organization. Dementia. Geneva: World Health Organization; 2023.
4. McKhann GM, Knopman DS, Chertkow H, Hyman BT, Jack CR Jr, Kawas CH, et al. The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimers Dement. 2011;7(3):263–9.
5. Hampel H, O'Bryant SE, Molinuevo JL, Zetterberg H, Masters CL, Lista S, et al. Blood-based biomarkers for Alzheimer disease: mapping the road to the clinic. Nat Rev Neurol. 2018;14(11):639–52.
6. Palmqvist S, Janelidze S, Stomrud E, Zetterberg H, Karl J, Zink K, et al. Performance of fully automated plasma assays as screening tests for Alzheimer disease–related β-amyloid status. JAMA Neurol. 2019;76(9):1060–9.
7. Karikari TK, Bened, Ashton NJ, Lantero Rodriguez J, Snellman A, Suárez-Calvet M, et al. Diagnostic performance and prediction of clinical progression of plasma phospho-tau181 in the Alzheimer's Disease Neuroimaging Initiative. Mol Psychiatry. 2021;26(2):429–42.
8. Doecke JD, Laws SM, Faux NG, Wilson W, Burnham SC, Lam CP, et al. Blood-based protein biomarkers for diagnosis of Alzheimer disease. Arch Neurol. 2012;69(10):1318–25.
9. Schindler SE, Bollinger JG, Ovod V, Mawuenyega KG, Li Y, Gordon BA, et al. High-precision plasma β-amyloid 42/40 predicts current and future brain amyloidosis. Neurology. 2019;93(17):e1647–59.
10. Varesi A, Carrara A, Pires VG, Floris V, Pierella E, Savioli G, et al. Blood-based biomarkers for Alzheimer's disease diagnosis and progression: an overview. Cells. 2022;11(8):1367.
11. Tao QQ, Lin RR, Wu ZY. Early diagnosis of Alzheimer's disease: moving toward a blood-based biomarkers era. Clin Interv Aging. 2023;353–8.
12. Padala SP, Newhouse PA. Blood-based biomarkers in Alzheimer's disease: a mini-review. Metab Brain Dis. 2023;38(1):185–93.
13. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. Int J Surg. 2021;88:105906.
14. Luijendijk HJ. How to create PICO questions about diagnostic tests. BMJ Evid-Based Med. 2021;26(4):155–7.
15. Strothmann M. Zotero: A Guide for Librarians, Researchers and Educators. Ref User Serv Q. 2018;57(3):222–3.

16. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med. 2011;155(8):529–36.
17. Popay J, Roberts H, Sowden A, Petticrew M, Arai L, Rodgers M, et al. Guidance on the conduct of narrative synthesis in systematic reviews. Prod ESRC Methods Programme Version. 2006;1(1):b92.
18. Lisy K, Porritt K. Narrative synthesis: considerations and challenges. JBI Evid Implement. 2016;14(4):201.
19. Udeh-Momoh C, Zheng B, Sandebring-Matton A, Novak G, Kivipelto M, Jönsson L, et al. Blood derived amyloid biomarkers for Alzheimer's disease prevention. J Prev Alzheimers Dis. 2022;9(1):12–21.
20. Gao F, Dai L, Wang Q, Liu C, Deng K, Cheng Z, et al. Blood-based biomarkers for Alzheimer's disease: a multicenter-based cross-sectional and longitudinal study in China. Sci Bull. 2023;68(16):1800–8.
21. Li W, Sun L, Yue L, Xiao S. Diagnostic and predictive power of plasma proteins in Alzheimer's disease: a cross-sectional and longitudinal study in China. Sci Rep. 2024;14(1):17557.
22. Souchet B, Michaïl A, Heuillet M, Dupuy-Gayral A, Haudebourg E, Pech C, et al. Multiomics blood-based biomarkers predict Alzheimer's predementia with high specificity in a multicentric cohort study. J Prev Alzheimers Dis. 2024;11(3):567–81.
23. Kubota M, Bun S, Takahata K, Kurose S, Momota Y, Iwabuchi Y, et al. Plasma biomarkers for early detection of alzheimer's disease: a cross-sectional study in a Japanese cohort. Alzheimers Res Ther. 2025;17(1):131.
24. Mattsson N, Andreasson U, Zetterberg H, Blennow K. Association of plasma neurofilament light with neurodegeneration in patients with Alzheimer disease. JAMA Neurol. 2017;74(5):557–66.

Publisher's Note