



Multimodal Machine Learning for Early Alzheimer's Disease Prediction: Integrating Biomarkers, Artificial Intelligence, and Precision Neurology

OVERVIEW: This narrative review examines how multimodal machine learning, which combines blood biomarkers, neuroimaging, genetics and cognitive data, could support early and individualised prediction of Alzheimer's disease. The evidence indicates that the value of each data type shifts with disease stage and amyloid burden, and a framework built on biological complementarity, temporal adaptation, robustness to missing data and interpretability is proposed. The validation and generalisability gaps that must be closed before clinical use are also outlined.

AUTHORS: Öykü Nefes Özinanç Eyüboğlu (Eyüboğlu, Türkiye)

KEYWORDS: Alzheimer's disease; multimodal machine learning; blood biomarkers; explainable artificial intelligence; precision neurology

ABSTRACT: Early and accurate prediction of the onset and progression of Alzheimer's disease represents a major medical challenge, since years of pathological changes precede the first clinical sign of cognitive decline. So far, clinical prediction has relied on the analysis of individual biomarkers or, at best, data from a limited number of modalities. However, the complex and largely unknown biological mechanisms underlying the heterogeneity of the disease make it unlikely that a single biomarker or a single data modality will allow robust prediction of disease in individual subjects or of the rate of progression in already affected subjects. There is therefore a strong need for machine learning algorithms that integrate, within a unifying framework, the different classes of data generated in the context of precision medicine. Here, a framework of multimodal machine learning is presented and its potential for early prediction is discussed, with special emphasis on individualised risk assessment, explainable artificial intelligence and precision neurology. A series of limitations that may hinder implementation of the proposed

framework in the clinical setting is also presented. These limitations are largely due to the heterogeneity of the different classes of data, the lack of large multimodal patient cohorts, the lack of external validation of performance, and the lack of model interpretability. By using models that are both interpretable by humans and validated against clinical endpoints, it may be possible to move towards more individualised methods of early prediction of Alzheimer's disease.

INTRODUCTION:

Alzheimer's disease (AD) is a primary cause of dementia and poses a major challenge in medical science. At the moment, more than 55 million individuals around the world live with dementia, with Alzheimer's accounting for about 60-70% of that number. Since people's life expectancy is increasing, the global incidence of dementia is likely to increase tremendously in the next decades, creating an unprecedented medical, social, and economic burden for the healthcare systems around the world.¹



Nonetheless, despite years of thorough scientific research, there is still one major difficulty that hinders progress in this area: Alzheimer's is usually diagnosed when irreversible changes in the brain have already occurred, different from other types of neurological diseases that manifest themselves in a quick way. The molecular and cellular changes that take place in the organism occur long before cognitive impairment can be observed, letting the disease progress while the patients remain healthy.¹ By the time the cognitive impairment reaches a sufficient degree for diagnosis, it is often too late to maximize the therapeutic effect of the drugs.²

Understanding this prolonged preclinical phase has fundamentally changed the objectives in research about Alzheimer's disease. Rather than focusing specifically on diagnosing established dementia, contemporary research increasingly aims to identify individuals during the earliest stages of disease development, when therapeutic intervention may still preserve neuronal function and delay clinical progression.³ Consequently, early prediction has become one of the principal goals of modern neuroscience, precision medicine, and computational neurology.

The complexity inherent in the biology of Alzheimer's disease is one of the greatest obstacles to reaching the objective. Alzheimer's disease is the result of a set of

different pathogenic processes that are continuously interacting with each other.⁴ This means that the disease process begins at different stages, proceeds with different speeds, and varies from person to person due to genetic predisposition, environmental influences and lifestyle and cognitive issues. Because of this, no single biomarker could properly reflect the disease.

In the last two decades, advances in research of biomarkers have turned Alzheimer's disease from a purely clinical syndrome to a biological pathology. Blood biomarkers such as phosphorylated tau (pTau181, pTau217, and pTau231), the amyloid- β 42/40 ratio, GFAP, and NfL can detect pathological processes in the body via less invasive techniques.⁵ Structural magnetic resonance imaging, PET, CSF biomarkers, genetic profiling and standardized neuropsychological tests provide additional information about the stage and course of disease.⁶ Nonetheless, no biomarker can capture the full extent of the biological complexity of Alzheimer's disease. Though structural MRI scans show changes associated with neurodegeneration, genetic variants can again shed light on the

¹ Jack, C. R., Jr., et al. "NIA-AA Research Framework: Toward a Biological Definition of Alzheimer's Disease." *Alzheimer's & Dementia* 14, no. 4 (2018): 535–562. <https://doi.org/10.1016/j.jalz.2018.02.018>.
Dubois, B., et al. "Clinical Diagnosis of Alzheimer's Disease: Recommendations of the International Working Group." *The Lancet Neurology* 20, no. 6 (2021): 484–496.

² Cummings, J., et al. "Alzheimer's Disease Drug Development Pipeline: 2022." *Alzheimer's & Dementia: Translational Research & Clinical Interventions* 8, no. 1 (2022): e12295. <https://doi.org/10.1002/trc2.12295>.

³ Jack et al., "NIA-AA Research Framework." Hansson, O. "Biomarkers for Neurodegenerative Diseases." *Nature Medicine* 27, no. 6 (2021): 954–963. <https://doi.org/10.1038/s41591-021-01382-x>.

⁴ Long, J. M., and D. M. Holtzman. "Alzheimer Disease: An Update on Pathobiology and Treatment Strategies." *Cell* 179, no. 2 (2019): 312–339. <https://doi.org/10.1016/j.cell.2019.09.001>. Knopman, D. S., et al. "Alzheimer Disease." *Nature Reviews Disease Primers* 7 (2021): 33. <https://doi.org/10.1038/s41572-021-00269-y>.

⁵ Palmqvist, S., et al. "Blood Biomarkers to Detect Alzheimer Disease in Primary Care and Secondary Care." *JAMA* 332, no. 15 (2024): 1245–1257. <https://doi.org/10.1001/jama.2024.13855>. Ashton, N. J., et al. "Diagnostic Accuracy of a Plasma Phosphorylated Tau 217 Immunoassay for Alzheimer Disease Pathology." *JAMA Neurology* 81, no. 3 (2024): 255–263. <https://doi.org/10.1001/jamaneurol.2023.5319>.

⁶ Jack, C. R., Jr., et al. "Revised Criteria for Diagnosis and Staging of Alzheimer's Disease: Alzheimer's Association Workgroup." *Alzheimer's & Dementia* 20, no. 8 (2024): 5143–5169. <https://doi.org/10.1002/alz.13859>. Ossenkoppele, R., et al. "Tau-PET to Predict Conversion from Mild Cognitive Impairment to Dementia; a Head-to-Head Comparison against Amyloid-PET and MRI." *Alzheimer's & Dementia* 19 (2023): e078754. <https://doi.org/10.1002/alz.078754>.



genetic predisposition of a patient without providing an insight into the ongoing disease process. Blood biomarkers provide information about specific pathways or lesions, whereas cognitive tests usually detect impairment only after the loss of numerous brain cells.⁷ Despite the fact that each biomarker is useful on its own, none of them can mirror the entire biological picture of Alzheimer's syndrome.

This is why in recent years multimodal machine learning has become well established. Multimodal machine learning integrates separate types of biological and clinical data, neuroimaging, blood markers, genetic information, demographic data, and cognitive tests, unlike traditional methods which analyze each of the biomarkers in isolation. Regardless of the fact that improving predictive accuracy alone may not imply clinical readiness, it is true that reported results differ greatly from different studies. That is due to the fact that studies differ substantially in their patient population, choice of biomarkers, machine learning solutions utilized, validation techniques, definitions of outcomes, and the composition of the datasets used. Numerous models have been created based on very homogeneous research samples of patients, such as ADNI, while only a small number of them underwent proper independent validation in different populations in terms of geography and ethnic diversity. For this reason, it is difficult to compare various models. The feasibility of applying some proposed models is unknown.⁸

⁷ Hansson, "Biomarkers for Neurodegenerative Diseases.". Karikari, T. K., et al. "Blood Phospho-Tau in Alzheimer's Disease: Diagnostic Performance and Clinical Utility." *Nature Reviews Neurology* 18, no. 7 (2022): 400–418. <https://doi.org/10.1038/s41582-022-00665-2>.

⁸ Varoquaux, G., and V. Cheplygina. "Machine Learning for Medical Imaging: Methodological Failures and Recommendations for the Future." *npj Digital Medicine* 5 (2022): 48. <https://doi.org/10.1038/s41746-022-00592-y>. Topol, E. J. "High-Performance Medicine: The Convergence of Human and Artificial Intelligence." *Nature Medicine* 25, no. 1 (2019): 44–56.

This review focuses on assessing the state of multimodal machine learning in the field of early prediction of Alzheimer's disease. It is not a simple review of the literature, but an evaluation of the biological reasons for multimodality, a comparison of different AI approaches, an assessment of the methodological restrictions limiting their implementation, and a consideration of the further development of interpretable and clinically efficient computing systems. The aim is to analyze the current advancements in biomarker creation, application of machine learning, and precision neurology.⁹

Biological basis of Alzheimer's disease

Alzheimer's disease is a progressive neurodegenerative disorder characterized by the gradual deterioration of cognitive function, memory, and behavior. Although it is most commonly recognized by its clinical manifestations, the disease begins many years before symptoms become apparent through a complex sequence of molecular and cellular events. Rather than resulting from a single pathological mechanism, Alzheimer's disease develops through the interaction of numerous biological processes that collectively disrupt neuronal function and ultimately lead to widespread neurodegeneration. This biological complexity explains why early diagnosis remains difficult and why no individual biomarker can fully characterize disease progression.

For many years, the amyloid cascade hypothesis has served as the dominant framework for understanding Alzheimer's disease. According to this model, abnormal processing of amyloid precursor protein (APP) produces amyloid- β ($A\beta$) peptides that aggregate into soluble oligomers and extracellular plaques within the brain. Although amyloid plaques represent one of the pathological hallmarks of Alzheimer's disease, growing evidence suggests that

<https://doi.org/10.1038/s41591-018-0300-7>.

⁹ Varoquaux and Cheplygina, "Methodological Failures.". Topol, "High-Performance Medicine.".



soluble A β oligomers may exert greater neurotoxicity than the plaques themselves by disrupting synaptic communication, impairing neuronal signaling, and triggering inflammatory responses. Importantly, amyloid accumulation often begins one to two decades before cognitive symptoms become clinically detectable, making it one of the earliest measurable biological changes associated with the disease.

Despite its central role, amyloid pathology alone does not adequately explain the clinical progression of Alzheimer's disease. Numerous studies have demonstrated that some individuals accumulate substantial amyloid deposits while remaining cognitively unimpaired, whereas others experience significant cognitive decline despite relatively modest plaque burden. These observations suggest that amyloid deposition represents only one component of a broader pathological network rather than the sole driver of neurodegeneration. Consequently, recent research has increasingly focused on understanding how amyloid interacts with other biological mechanisms throughout disease progression.

Among these mechanisms, tau pathology demonstrates a considerably stronger relationship with neuronal dysfunction and cognitive decline. Tau is a microtubule-associated protein responsible for maintaining neuronal structure and facilitating intracellular transport. Under pathological conditions, tau undergoes abnormal hyperphosphorylation, causing it to detach from microtubules and aggregate into intracellular neurofibrillary tangles. The accumulation of these tangles disrupts axonal transport, compromises neuronal integrity, and contributes directly to neuronal death. Unlike amyloid deposition, which may remain relatively stable over time, the regional spread of tau pathology closely parallels the progression of cognitive impairment, making phosphorylated tau one of the most promising biomarkers for disease staging and prognosis.

Neuroinflammation has emerged as another critical contributor to Alzheimer's disease pathogenesis. Microglia and astrocytes, which normally maintain neuronal homeostasis and protect the brain from injury, become chronically activated in response to persistent amyloid and tau pathology. While acute inflammatory responses may initially serve protective functions by clearing toxic proteins and cellular debris, prolonged activation promotes the sustained release of inflammatory cytokines, oxidative molecules, and complement proteins that amplify neuronal injury. This persistent inflammatory environment not only accelerates neurodegeneration but also influences disease progression through complex interactions with amyloid and tau pathology.

As these pathological processes continue, progressive synaptic dysfunction and neuronal loss become increasingly apparent. Synaptic degeneration is strongly associated with declining cognitive performance because synapses represent the fundamental communication network of the brain. Over time, neuronal death produces characteristic structural changes, including hippocampal atrophy and cortical thinning, which can be visualized using structural magnetic resonance imaging (MRI). These imaging findings often correspond with later stages of disease progression, illustrating that structural neurodegeneration generally reflects damage that has already accumulated over many years rather than the earliest pathological events.

Genetic susceptibility further contributes to the heterogeneity of Alzheimer's disease. Variants of the APOE gene, particularly the $\epsilon 4$ allele, remain the strongest genetic risk factor for late-onset Alzheimer's disease, whereas mutations in APP, PSEN1, and PSEN2 are associated with rare familial forms of the disease. Nevertheless, genetics alone cannot fully explain disease onset or progression. Age, cardiovascular health, metabolic disorders, sleep quality, education, physical activity, and



environmental exposures all interact with genetic predisposition to influence individual disease trajectories. Consequently, patients with similar genetic profiles may exhibit markedly different clinical outcomes.

Collectively, these findings illustrate that Alzheimer's disease is not the consequence of a single abnormal protein or isolated molecular pathway but rather the result of multiple interconnected biological mechanisms that evolve over time. Amyloid deposition, tau pathology, neuroinflammation, synaptic dysfunction, neurodegeneration, and genetic susceptibility each provide valuable yet incomplete insight into disease progression. This multidimensional nature forms the biological rationale for the development of multimodal diagnostic strategies. Because different biomarkers capture different aspects of Alzheimer's disease and become abnormal at distinct stages, integrating complementary sources of biological information offers a more comprehensive representation of disease progression than any single biomarker alone.

RESULTS

Multimodal machine learning for early prediction

The inherent biological heterogeneity of Alzheimer's disease suggests that the integration of complementary information sources may provide a more robust foundation for individualized risk assessment than reliance on any solitary biomarker. Multimodal machine learning (ML) architectures facilitate the unification of neuroimaging data, molecular biomarkers, genetic profiles, and cognitive or demographic variables within a single predictive framework. However, the efficacy of such multimodal integration is contingent not merely upon the quantity of variables, but upon whether each modality provides biologically and clinically distinct insights that enhance the overall model.¹⁰

Empirical investigations have validated this principle through diverse combinations of data modalities. Ritter et al. (2015) explored the transition from mild cognitive impairment (MCI) to clinical AD using multimodal datasets from the Alzheimer's Disease Neuroimaging Initiative (ADNI), encompassing neuropsychological metrics, MRI, PET, and cerebrospinal fluid (CSF) analytes. Their optimized model, utilizing support vector machines and expert-derived features, achieved 73% accuracy, identifying neuropsychological and imaging measures as primary drivers of predictive performance. Critically, this research addressed the challenge of missing biomarker data, establishing that multimodal frameworks remain viable even when specific measurements are inaccessible in clinical practice.¹¹

The synergy of combining distinct biological dimensions was further elucidated by Reas et al. (2023), who constructed a multimodal hazard score integrating chronological age, polygenic risk, volumetric brain atrophy, and memory performance. This integrated approach facilitated superior risk stratification compared to individual predictors; individuals in the highest quintile of the multimodal score exhibited a substantially elevated risk of conversion to dementia. These results underscore that genetic predisposition, structural neurodegeneration, and cognitive decline offer complementary perspectives on the trajectory of disease progression.¹²

Hansson, "Biomarkers for Neurodegenerative Diseases."

¹¹ Ritter, K., et al. "Multimodal Prediction of Conversion to Alzheimer's Disease Based on Incomplete Biomarkers." *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring* 1, no. 2 (2015): 206–215.
<https://doi.org/10.1016/j.dadm.2015.01.006>.

¹² Reas, E. T., et al. "Improved Multimodal Prediction of Progression from MCI to Alzheimer's Disease Combining Genetics with Quantitative Brain MRI and Cognitive Measures." *Alzheimer's & Dementia* 19, no. 11 (2023): 5151–5158.
<https://doi.org/10.1002/alz.13112>.

¹⁰ Jack et al., "NIA-AA Research Framework."



processes, genetic measures define baseline susceptibility, and cognitive tests document the functional manifestations of disease. However, evidence also indicates that no singular combination of modalities is universally superior; the most relevant data sources fluctuate based on disease stage, amyloid burden, and specific clinical endpoints.¹⁵ Consequently, effective frameworks should emphasize biological complementarity and clinical adaptability over simple data aggregation.

Clinical feasibility remains a paramount concern for real-world implementation. While research cohorts often provide comprehensive multimodal datasets, routine clinical environments frequently involve incomplete records. Predictive models that maintain robustness despite missing measurements are significantly more practical for precision neurology. The work of Ritter et al. (2015) highlights that strategic handling of missing data can preserve the utility of multimodal systems. Future developments should prioritize identifying parsimonious combinations of accessible and complementary markers that can reliably deliver individualized risk assessments in diverse clinical settings.¹⁶

In summary, the current body of evidence identifies multimodal ML as a transformative approach for the early prediction of AD, yet it cautions that increasing data volume does not inherently guarantee improved accuracy. Instead, a selective strategy is warranted, one that aligns modality selection with biological relevance, clinical accessibility, and the specific stage of the disease. This principle underpins the framework proposed herein, which seeks to integrate complementary biomarkers and clinical data while addressing the critical needs for model

¹³ Mirabnahrzham, G., et al. "Machine Learning Based Multimodal Neuroimaging Genomics Dementia Score for Predicting Future Conversion to Alzheimer's Disease." *Journal of Alzheimer's Disease* 87, no. 3 (2022): 1345–1365. <https://doi.org/10.3233/JAD-220021>.

¹⁴ Gebre, R. K., et al. "Can Integration of Alzheimer's Plasma Biomarkers with MRI, Cardiovascular, Genetics, and Lifestyle Measures Improve Cognition Prediction?" *Brain Communications* 6, no. 5 (2024): fcae300. <https://doi.org/10.1093/braincomms/fcae300>.

¹⁵ Mirabnahrzham et al., "Neuroimaging Genomics Dementia Score." Gebre et al., "Plasma Biomarkers with MRI."

¹⁶ Ritter et al., "Incomplete Biomarkers." Topol, "High-Performance Medicine."



integration of complementary biological signals that a solitary modality cannot fully encapsulate.

Moreover, the framework must account for the temporal evolution of AD pathology. While genetic risk remains constant, molecular shifts often manifest decades before structural or cognitive decline, with the latter becoming increasingly relevant as the disease advances.²⁰ Longitudinal architectures should therefore refine individualized risk assessments as iterative biomarker, imaging, and cognitive data become available. This approach facilitates a shift from static classification toward the modeling of specific disease trajectories for each patient.

A critical requirement for clinical implementation is resilience to incomplete datasets. Since comprehensive multimodal profiles are frequently unavailable in routine practice, a clinically efficacious system must assess whether the existing data provide sufficient confidence for a reliable assessment rather than treating missing variables as a systemic failure. Research by Ritter et al. (2015) suggests that multimodal models can maintain predictive utility despite missing analytes. Such a system could further improve clinical workflows by identifying specific instances where an additional diagnostic test would substantively reduce predictive uncertainty.²¹

Finally, the framework must maintain a clear distinction between computational risk assessment and definitive clinical diagnosis. Model outputs should offer individualized risk estimates alongside transparent insights into the underlying features driving those results. While explainable AI techniques like SHAP or imaging-based attribution provide valuable context for model behavior, these interpretations should be viewed as evidence

¹⁷ Varoquaux and Cheplygina, "Methodological Failures."

¹⁸ Mirabnahrzazam et al., "Neuroimaging Genomics Dementia Score.". Gebre et al., "Plasma Biomarkers with MRI."

¹⁹ Hansson, "Biomarkers for Neurodegenerative Diseases.". Karikari et al., "Blood Phospho-Tau."

²⁰ Jack et al., "NIA-AA Research Framework.". Mirabnahrzazam et al., "Neuroimaging Genomics Dementia Score."

²¹ Ritter et al., "Incomplete Biomarkers."



of algorithmic logic rather than confirmed biological causality.²²

In summary, the proposed architecture is governed by four foundational principles: biological complementarity, temporal adaptation, robustness to missing data, and model interpretability. Rather than seeking performance gains through increased data volume alone, the objective is to isolate the most parsimonious combination of accessible and biologically meaningful markers to achieve reliable, individualized prediction in the early stages of AD.

Limitations

The present review is subject to several methodological constraints. Primarily, the significant heterogeneity across investigated cohorts, datasets, and predictive targets, coupled with varying machine-learning architectures and validation protocols, precludes direct performance comparisons. Furthermore, a substantial portion of the multimodal literature depends upon research-grade cohorts with highly standardized measurements, potentially limiting their generalizability to broader, real-world clinical populations.²³ Additionally, the rapid evolution of both blood-based biomarkers and artificial intelligence ensures that critical evidence may continue to surface beyond the scope of the literature synthesized herein.

Moreover, it is important to note that this work constitutes a focused narrative review rather than a comprehensive systematic review or meta-analysis. Consequently, quantitative pooling of model performance was not undertaken, and the proposed

framework should be viewed as a theoretical synthesis of contemporary evidence rather than a clinically validated predictive instrument.

Future directions

Subsequent research efforts should prioritize the development of adaptive and longitudinal multimodal architectures instead of merely expanding the volume of input variables. Given that the relative contribution of discrete modalities fluctuates according to disease stage and individual patient profiles, next-generation models must be capable of refining risk assessments as longitudinal data becomes available.²⁴ Clinical accessibility must also emerge as a foundational consideration. Future frameworks should maintain predictive utility even when certain modalities are missing and should provide guidance on whether further diagnostics would substantively reduce uncertainty. Such an approach would mitigate unnecessary testing while enhancing the practical utility of multimodal prediction in routine clinical settings.²⁵

Lastly, rigorous prospective validation across geographically and demographically diverse populations remains imperative. Models must undergo evaluation using clinically relevant outcomes, independent datasets, and interpretable outputs before integration into clinical decision-support systems. The objective, therefore, is not the pursuit of computational complexity for its own sake, but the realization of models that are robust, transparent, and clinically efficacious.

Conclusion

Alzheimer's disease arises from a convergence of interacting pathological processes that manifest across a prolonged continuum, complicating individualized prediction. The evidence synthesized here indicates that multimodal machine learning

²² Viswan Vimbi, V., et al. "Interpreting Artificial Intelligence Models: A Systematic Review on the Application of LIME and SHAP in Alzheimer's Disease Detection." *Brain Informatics* 11 (2024): 10. <https://doi.org/10.1186/s40708-024-00222-1>. Wang, F., et al. "Interpretable Machine Learning-Driven Biomarker Identification and Validation for Alzheimer's Disease." *Scientific Reports* 14 (2024): 30770. <https://doi.org/10.1038/s41598-024-80401-6>.

²³ Varoquaux and Cheplygina, "Methodological Failures."

²⁴ Mirabnahrzazam et al., "Neuroimaging Genomics Dementia Score." Gebre et al., "Plasma Biomarkers with MRI."

²⁵ Ritter et al., "Incomplete Biomarkers."



can navigate this complexity by unifying complementary insights from molecular biomarkers, neuroimaging, genetics, and cognitive metrics. However, these data also reveal that increasing data volume does not inherently yield clinically superior prediction; the utility of each modality is intrinsically tied to disease stage, amyloid burden, and specific clinical endpoints.²⁶

The analysis presented supports a paradigm shift from simple data aggregation toward adaptive, individualized multimodal prediction. Rather than subjecting every patient to a uniform diagnostic battery, future systems should identify the most parsimonious combination of accessible and biologically relevant data for the individual. Critical to this transition will be the integration of longitudinal updating, robustness to incomplete datasets, external validation, and model interpretability.

Ultimately, the impact of multimodal artificial intelligence in AD should be measured by its capacity to translate heterogeneous biological signals into reliable, individualized, and actionable clinical predictions. An adaptive and interpretable framework offers a pathway toward precision neurology by facilitating earlier risk identification and a more nuanced understanding of individual disease trajectories.

MATERIALS AND METHODS

Literature search

For the identification of literature pertinent to the topic under investigation, a focused review of academic literature was conducted to identify studies examining the use of machine learning techniques in the early detection and prediction of Alzheimer's disease. The search encompassed published articles concerning blood-based and cerebrospinal fluid protein biomarkers, neuroimaging techniques, genetic factors, and artificial intelligence techniques that

incorporate different modalities of data. The author searched databases including PubMed, Google Scholar, Web of Science, Scopus, Nature Portfolio, Cell Press, ScienceDirect, IEEE Xplore, SpringerLink, and the NCBI. The search was guided by terms related to Alzheimer's disease, machine learning, artificial intelligence, deep learning, blood biomarkers, plasma proteins, tau, amyloid- β , neurofilament light chain, glial fibrillary acidic protein, and other relevant keywords. Boolean operators and Medical Subject Headings (MeSH) were used where appropriate.

Study selection

Original research containing clinical or experimental evidence for multimodal approaches was prioritized, particularly studies that combined two or more biomarkers or data modalities using machine learning or deep learning techniques. Studies providing important biological evidence for AD, identifying key biomarkers, or introducing relevant technical approaches were also considered. Recent literature focusing on blood-based biomarkers, multimodal data integration, machine learning techniques, explainable AI, federated learning, and model validation was given particular attention.

Research conducted in universities, hospitals, governmental organizations, and interdisciplinary research institutes was prioritized. Relevant literature was identified from journals including Nature, Nature Medicine, Nature Reviews Neurology, Nature Aging, The Lancet Neurology, JAMA Neurology, Brain, Alzheimer's & Dementia, and Alzheimer's Research & Therapy, among others. Review articles were used primarily to provide background on the biological mechanisms of AD and major developments in the field, while the main findings discussed in this review were based mainly on original experimental and clinical studies. The cited references therefore combine key studies that contributed to the understanding of the biological mechanisms of Alzheimer's disease with more recent

²⁶ Mirabnabrazam et al., "Neuroimaging Genomics Dementia Score.". Gebre et al., "Plasma Biomarkers with MRI.".



research on multimodal machine learning, biomarker discovery, and clinical application. Whenever possible, preference was given to studies with large and representative samples, external evaluation, independent datasets, and transparent reporting of model performance.

Synthesis

Rather than describing each study separately, the existing literature was grouped into several key thematic areas concerning the biological mechanisms of Alzheimer's disease, current biomarkers, advances in multimodal machine learning, comparisons between prediction approaches, and methodological difficulties. This approach allowed connections between biological discoveries and computational developments to be examined while also identifying common findings and areas of disagreement. This review aimed to integrate information from molecular neuroscience, biomarker research, neuroimaging, genetics, and artificial intelligence in order to examine how multimodal machine learning could contribute to the early prediction of Alzheimer's disease.

ACKNOWLEDGMENTS

The author gratefully acknowledges the researchers whose work on Alzheimer's disease, biomarkers, neuroimaging, machine learning, and precision medicine provided the scientific foundation for this review. The author also thanks the educators and mentors who encouraged scientific curiosity, critical thinking, and independent research throughout the preparation of this manuscript.

Ethics statement

This manuscript is a literature review based entirely on previously published scientific studies. No human participants, animals, biological specimens, or identifiable personal data were directly involved in this work. Therefore, ethical approval and informed consent were not required.

Data availability statement

No new datasets, code, or experimental materials were generated or analyzed in this study. The information synthesized in this review was obtained from the published scientific literature cited in the reference list.

Conflicts of interest

The author declares that there are no financial, professional, or personal relationships that could have influenced the work presented in this manuscript. The author declares no competing interests.

REFERENCES

- Alzheimer's Disease International. World Alzheimer Report 2024: Global Changes in Attitudes to Dementia. Alzheimer's Disease International, 2024.
- Ashton, N. J., et al. "Diagnostic Accuracy of a Plasma Phosphorylated Tau 217 Immunoassay for Alzheimer Disease Pathology." *JAMA Neurology* 81, no. 3 (2024): 255–263. <https://doi.org/10.1001/jamaneurol.2023.5319>.
- Cummings, J., et al. "Alzheimer's Disease Drug Development Pipeline: 2022." *Alzheimer's & Dementia: Translational Research & Clinical Interventions* 8, no. 1 (2022): e12295. <https://doi.org/10.1002/trc2.12295>.
- Dubois, B., et al. "Clinical Diagnosis of Alzheimer's Disease: Recommendations of the International Working Group." *The Lancet Neurology* 20, no. 6 (2021): 484–496.
- Gebre, R. K., et al. "Can Integration of Alzheimer's Plasma Biomarkers with MRI, Cardiovascular, Genetics, and Lifestyle Measures Improve Cognition Prediction?" *Brain Communications* 6, no. 5 (2024): fcae300. <https://doi.org/10.1093/braincomms/fcae300>.
- Hansson, O. "Biomarkers for Neurodegenerative Diseases." *Nature Medicine* 27, no. 6 (2021): 954–963. <https://doi.org/10.1038/s41591-021-01382-x>.
- Jack, C. R., Jr., et al. "NIA-AA Research Framework: Toward a Biological Definition of Alzheimer's Disease." *Alzheimer's & Dementia* 14, no. 4 (2018): 535–562. <https://doi.org/10.1016/j.jalz.2018.02.018>.



Jack, C. R., Jr., et al. "Revised Criteria for Diagnosis and Staging of Alzheimer's Disease: Alzheimer's Association Workgroup." *Alzheimer's & Dementia* 20, no. 8 (2024): 5143–5169. <https://doi.org/10.1002/alz.13859>.

Karikari, T. K., et al. "Blood Phospho-Tau in Alzheimer's Disease: Diagnostic Performance and Clinical Utility." *Nature Reviews Neurology* 18, no. 7 (2022): 400–418. <https://doi.org/10.1038/s41582-022-00665-2>.

Knopman, D. S., et al. "Alzheimer Disease." *Nature Reviews Disease Primers* 7 (2021): 33. <https://doi.org/10.1038/s41572-021-00269-y>.

Long, J. M., and D. M. Holtzman. "Alzheimer Disease: An Update on Pathobiology and Treatment Strategies." *Cell* 179, no. 2 (2019): 312–339. <https://doi.org/10.1016/j.cell.2019.09.001>.

Mirabnahrzham, G., et al. "Machine Learning Based Multimodal Neuroimaging Genomics Dementia Score for Predicting Future Conversion to Alzheimer's Disease." *Journal of Alzheimer's Disease* 87, no. 3 (2022): 1345–1365. <https://doi.org/10.3233/JAD-220021>.

Ossenkoppele, R., et al. "Tau-PET to Predict Conversion from Mild Cognitive Impairment to Dementia; a Head-to-Head Comparison against Amyloid-PET and MRI." *Alzheimer's & Dementia* 19 (2023): e078754. <https://doi.org/10.1002/alz.078754>.

Palmqvist, S., et al. "Blood Biomarkers to Detect Alzheimer Disease in Primary Care and Secondary Care." *JAMA* 332, no. 15 (2024): 1245–1257. <https://doi.org/10.1001/jama.2024.13855>.

Reas, E. T., et al. "Improved Multimodal Prediction of Progression from MCI to Alzheimer's Disease Combining Genetics with Quantitative Brain MRI and Cognitive Measures." *Alzheimer's & Dementia* 19, no. 11 (2023): 5151–5158. <https://doi.org/10.1002/alz.13112>.

Ritter, K., et al. "Multimodal Prediction of Conversion to Alzheimer's Disease Based on Incomplete Biomarkers." *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring* 1, no. 2 (2015): 206–215. <https://doi.org/10.1016/j.dadm.2015.01.006>.

Topol, E. J. "High-Performance Medicine: The Convergence of Human and Artificial Intelligence." *Nature Medicine* 25, no. 1 (2019):

44–56. <https://doi.org/10.1038/s41591-018-0300-7>.

Varoquaux, G., and V. Cheplygina. "Machine Learning for Medical Imaging: Methodological Failures and Recommendations for the Future." *npj Digital Medicine* 5 (2022): 48. <https://doi.org/10.1038/s41746-022-00592-y>.

Viswan Vimbi, V., et al. "Interpreting Artificial Intelligence Models: A Systematic Review on the Application of LIME and SHAP in Alzheimer's Disease Detection." *Brain Informatics* 11 (2024): 10. <https://doi.org/10.1186/s40708-024-00222-1>.

Wang, F., et al. "Interpretable Machine Learning-Driven Biomarker Identification and Validation for Alzheimer's Disease." *Scientific Reports* 14 (2024): 30770. <https://doi.org/10.1038/s41598-024-80401-6>.

World Health Organization. "Dementia." 2023. <https://www.who.int/news-room/fact-sheets/detail/dementia>.