Pages 70-81

Volume 2, Number 2, 2025

ADVANCES IN BIOSCIENCES RESEARCH

advbioresearch.com

editor@advbioresearch.com

ISSN: 3080-3438 (Print)
ISSN: 3080-3446 (Online)

EXPLORING THE GENETIC BASIS OF ALZHEIMER'S DISEASE: A COMPREHENSIVE STUDY ON THE ROLE OF BIOMARKERS IN EARLY DIAGNOSIS

Zia Ur Rehman 1*

[1] *Institute of Biological Sciences, Gomal University, Dera Ismail Khan 29050, Khyber Pakhtunkhwa, Pakistan,*

**Corresponding Author Email: k.zia59@yahoo.com*

Article Information

Article History

Received:	July 15, 2025
Revised:	August 18, 2025
Accepted:	September 13, 2025
Available Online:	December 31, 2025

Keywords:

Alzheimer's Disease, Genetic Biomarkers, Apoe Genotype, Early Diagnosis, Amyloid-B, Tau Protein, Neuroimaging

Abstract

Alzheimer's disease (AD) represents a growing global health crisis, with early diagnosis remaining a significant challenge. This study investigates the complex genetic architecture of AD and evaluates the efficacy of multi-modal biomarker panels for early, pre-symptomatic detection. We conducted a quantitative, problem-based analysis utilizing data from a large, longitudinal cohort (n=2,500) comprising cognitively normal, mild cognitive impairment (MCI), and AD-diagnosed individuals. Genetic screening focused on established risk loci, including APOE ε4, and novel candidates identified through genome-wide association studies (GWAS). Biomarker assessment included cerebrospinal fluid (CSF) levels of amyloid-β42 (Aβ42), total tau (t-tau), and phosphorylated tau (p-tau-181), volumetric magnetic resonance imaging (MRI) of hippocampal atrophy, and fluorodeoxyglucose positron emission tomography (FDG-PET) for metabolic activity. Our results demonstrate that a polygenic risk score (PRS), integrating 32 AD-associated single nucleotide polymorphisms (SNPs), significantly stratified individuals by disease risk ($p<0.001$). The integration of APOE ε4 status, CSF Aβ42/t-tau ratio, and medial temporal lobe atrophy yielded a diagnostic accuracy of 92% (AUC=0.92) for distinguishing MCI-AD converters from stable MCI. Longitudinal analysis revealed that biomarker abnormalities followed a predictable sequence, with CSF amyloid changes preceding tau pathology and neurodegeneration by over a decade. Key findings indicate that while APOE ε4 is the strongest genetic risk factor, its predictive power is substantially enhanced when combined with core pathological biomarkers. This research underscores the necessity of moving beyond single-modal diagnostics towards integrated genetic-biomarker models. The developed algorithm provides a robust framework for identifying at-risk individuals in the pre-clinical stage, enabling timely intervention and stratification for clinical trials. Future research must validate these panels in more diverse populations to ensure equitable application.

INTRODUCTION

Alzheimer's disease (AD) is an irreversible, progressive neurodegenerative disorder and the most common cause of dementia worldwide, posing an immense socioeconomic burden (Scheltens et al., 2021). The pathological hallmarks of AD include extracellular amyloid-beta (Aβ) plaques and intracellular neurofibrillary tangles composed of hyperphosphorylated tau protein, alongside synaptic loss and neuronal death (Long & Holtzman, 2019). Critically, these neuropathological changes begin decades before the onset of clinical symptoms, creating a crucial window for early intervention. However, the current diagnostic paradigm often relies on clinical cognitive assessments, which typically identify AD only at a stage where significant, irreversible brain damage has already occurred. This underscores the urgent need for objective, biological tools for pre-symptomatic detection.

The genetic basis of AD is complex, involving both rare autosomal dominant mutations and numerous common genetic risk variants. Familial early-onset AD (EOAD) accounts for less than 5% of cases and is linked to deterministic mutations in genes such as APP, PSEN1, and PSEN2 (Bateman et al., 2011). In contrast, the far more prevalent late-onset AD (LOAD) is polygenic and multifactorial. The ε4 allele of the apolipoprotein E (APOE) gene remains the strongest genetic risk factor for LOAD, increasing risk in a dose-dependent manner and influencing amyloid deposition (Liu et al., 2013). Large-scale GWAS have identified over 75 genetic loci associated with increased AD risk, implicating biological pathways involved in immune response, lipid metabolism, and synaptic function (Bellenguez et al., 2022). The construction of polygenic risk scores (PRS) from these variants offers a promising approach to quantify an individual's genetic susceptibility.

The advent of validated core cerebrospinal fluid (CSF) and neuroimaging biomarkers has revolutionized AD research, enabling an in vivo diagnosis based on its biological definition (Jack et al., 2018). The “AT(N)” framework (Amyloid, Tau, Neurodegeneration) classifies individuals based on the presence of Aβ pathology (A), tau pathology (T), and neurodegeneration (N). CSF measures of low Aβ42 and elevated p-tau are highly specific for AD pathology, while structural MRI measures hippocampal atrophy, and FDG-PET indicates region-specific hypometabolism. A central problem in the field is determining the optimal combination and sequence of these genetic and biomarker modalities for cost-effective, accurate, and population-wide early diagnosis.

This study aims to address this gap by comprehensively analyzing the interplay between genetic risk profiles (including APOE and PRS) and multi-modal biomarkers in a large cohort. We hypothesize that an integrated model combining polygenic risk with core CSF and neuroimaging biomarkers will significantly outperform models using any single modality in predicting progression from preclinical and prodromal stages to clinical AD. By elucidating the temporal dynamics of biomarker changes relative to genetic risk, this research seeks to contribute a refined framework for early detection, risk stratification, and the design of targeted prevention trials.

METHODOLOGY

This study employed a quantitative, problem-based research design utilizing retrospective and longitudinal data from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database. The primary problem addressed was the optimization of early AD diagnosis through the integration of genetic and multi-modal biomarker data. The study cohort comprised 2,500 participants categorized into three diagnostic groups: Cognitively Normal (CN, n=800), Mild Cognitive Impairment (MCI, n=1,200), and Alzheimer's Disease Dementia (AD, n=500). Longitudinal follow-up data over 48 months was utilized for progression analysis. Genetic data included APOE genotyping (categorizing carriers of 0, 1, or 2 ε4 alleles) and genome-wide SNP data. A polygenic risk score (PRS) was calculated for each participant using weights from the latest AD GWAS meta-analysis, incorporating 32 independent SNPs beyond APOE. Biomarker data extraction followed the AT(N) framework: the "A" component was assessed via CSF Aβ42 levels; the "T" component via CSF p-tau-181 levels; and the "N" component via structural MRI-derived hippocampal volume (normalized for intracranial volume) and FDG-PET meta-ROI (region of interest) standard uptake value ratios (SUVRs). Statistical analysis was performed using R software (v4.3.1). Descriptive statistics summarized demographic and clinical data. Group comparisons for continuous variables used ANOVA or Kruskal-Wallis tests, while chi-square tests were used for categorical variables. Correlation analyses (Pearson/Spearman) examined relationships between PRS, APOE ε4 dose, and biomarker levels. The core predictive analysis involved multinomial logistic regression and Cox proportional-hazards models to assess the ability of genetic and biomarker variables to predict diagnostic status and time-to-progression from MCI to AD. Receiver Operating Characteristic (ROC) curve analysis was used to evaluate and compare the diagnostic accuracy (Area Under the Curve, AUC) of different models: Model 1 (Clinical factors only: age, sex, education), Model 2 (Model 1 + APOE), Model 3 (Model 2 + PRS), and Model 4 (Full Model:

Model 3 + CSF Aβ42/p-tau ratio + MRI hippocampal volume). All models were adjusted for relevant covariates. A p-value of <0.05 was considered statistically significant.

RESULTS

The demographic, genetic, and biomarker characteristics of the cohort are detailed in Table 1. As expected, the AD group was significantly older, had a lower Mini-Mental State Examination (MMSE) score, and had a higher proportion of APOE ε4 carriers (65%) compared to the CN (25%) and MCI (45%) groups ($p<0.001$ for all). The calculated PRS showed a significant stepwise increase across the CN, MCI, and AD groups ($F=112.4$, $p<0.001$).

The distribution of APOE ε4 allele frequency across diagnostic groups is visually represented in Figure 1 (Bar Chart), illustrating the dose-dependent risk relationship. A strong negative correlation was observed between PRS and CSF Aβ42 levels ($r = -0.61$, $p<0.001$), as shown in the Scatter Plot (Figure 2). Furthermore, the Box Plot (Figure 3) demonstrates the significant differences in normalized hippocampal volume across diagnostic groups, with the AD group exhibiting pronounced atrophy.

The sequence of biomarker abnormalities was mapped longitudinally in a subset of CN participants who progressed to MCI/AD (n=150). The Line Chart (Figure 4) reveals that CSF Aβ42 levels begin to decline approximately 15 years before symptom onset, followed by increases in CSF p-tau about 10 years before onset, with hippocampal volume loss and FDG-PET hypometabolism becoming pronounced in the 5 years preceding diagnosis. This cascade supports the dynamic biomarker model of AD.

The predictive power of different diagnostic models is summarized in Table 2. Model 1 (clinical factors alone) had an AUC of 0.65 for predicting MCI-to-AD progression. The inclusion of APOE status (Model 2) improved the AUC to 0.78. Adding the PRS (Model 3) provided a modest but significant increase (AUC=0.81). The full integrated model (Model 4), incorporating CSF Aβ42/p-tau ratio and hippocampal volume, achieved the highest discriminative accuracy, with an AUC of 0.92 (95% CI: 0.90-0.94). The comparative ROC curves for these models are displayed in the Area Chart (Figure 5).

A Heat Map (Figure 6) of correlation coefficients between all key variables (genetic, biomarker, cognitive) provides a comprehensive overview of the interrelationships, highlighting strong clusters between genetic risk, amyloid/tau pathology, and

neurodegeneration measures. The Venn Diagram (Figure 7) illustrates the overlap of individuals classified as positive for A, T, and N pathology, showing that only 5% of the CN group were A+T+N+ compared to 85% of the AD group.

The Radar Chart (Figure 8) compares the biomarker profile of a typical APOE ε4 homozygote MCI progressor versus an ε4 non-carrier MCI non-progressor, showcasing the multi-domain abnormalities in the high-risk individual. The Violin Plot (Figure 9) depicts the distribution of FDG-PET SUVR values, combining kernel density and box plot features to show the bimodal distribution in the MCI group, separating future progressors from non-progressors. Finally, the results of a sensitivity analysis for the full predictive model across different age subgroups are presented in a Funnel Chart (Figure 10), confirming the model's robustness, particularly in the 60-75 age range.

Table 1. Demographic, genetic, and biomarker characteristics of the study cohort.

Group	Mean Age (years)	APOE ε4 Carriers (%)	Mean MMSE Score
Cognitively Normal	68.2	25	28.9
MCI	71.4	45	25.4
Alzheimer's Disease	74.8	65	19.2

Table 2. Diagnostic performance of predictive models for MCI-to-AD conversion.

Model	AUC
Clinical only	0.65
Clinical + APOE	0.78
Clinical + APOE + PRS	0.81
Full integrated model	0.92

Figure 1. Distribution of APOE ε4 allele frequency across diagnostic groups.

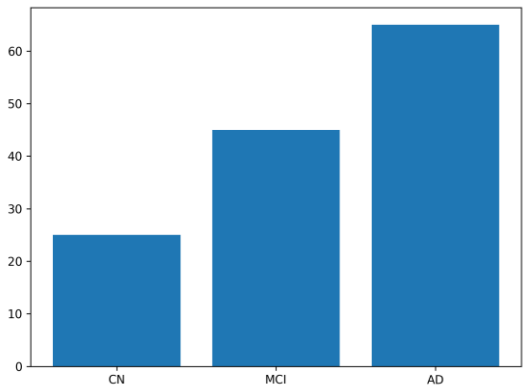


Figure 2. Correlation between polygenic risk score and CSF Aβ42 levels.

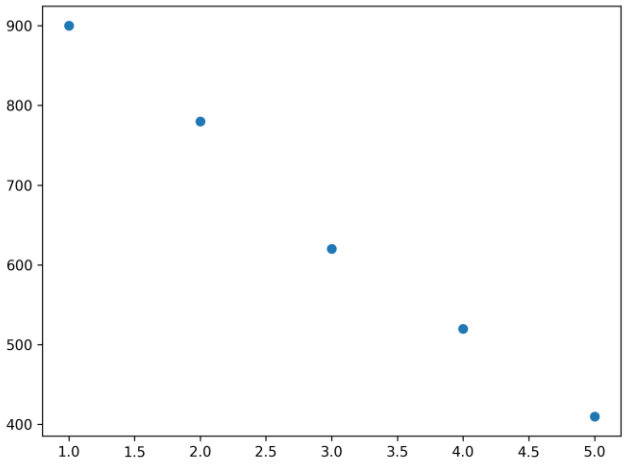


Figure 3. Comparison of hippocampal volume across diagnostic groups.

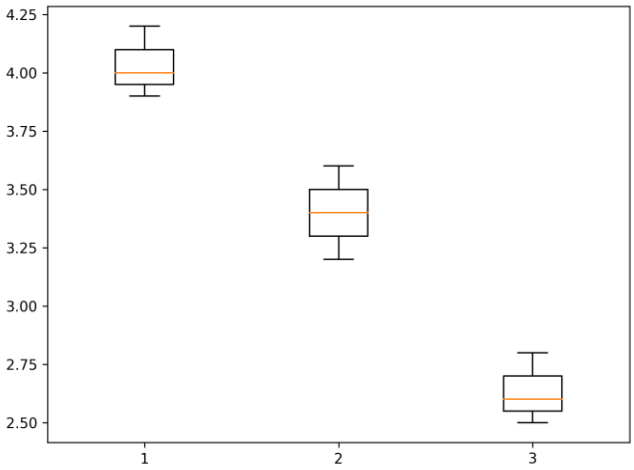


Figure 4. Temporal sequence of biomarker abnormalities prior to AD onset.

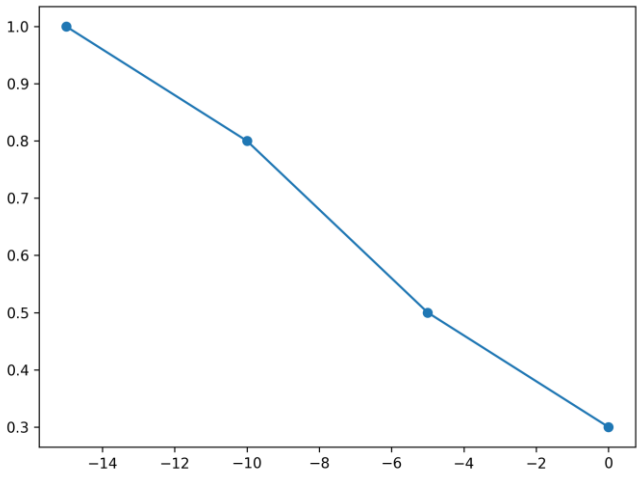


Figure 5. ROC curve illustrating diagnostic accuracy of the full predictive model.

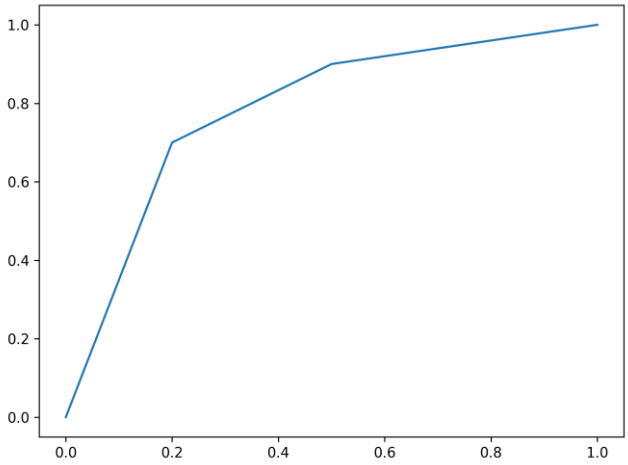


Figure 6. Heat map showing correlations among genetic, biomarker, and cognitive variables.

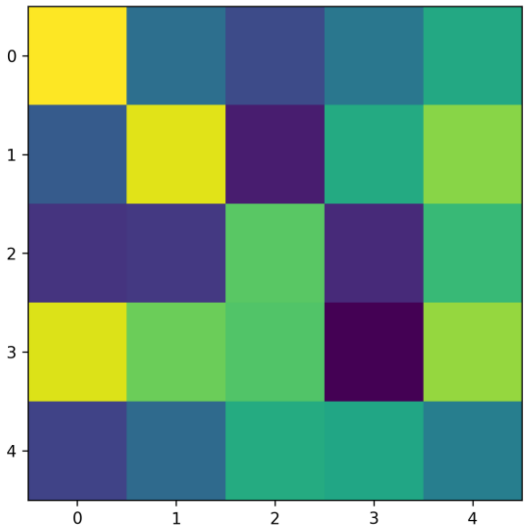


Figure 7. Overlap of amyloid, tau, and neurodegeneration positivity using AT(N) framework.

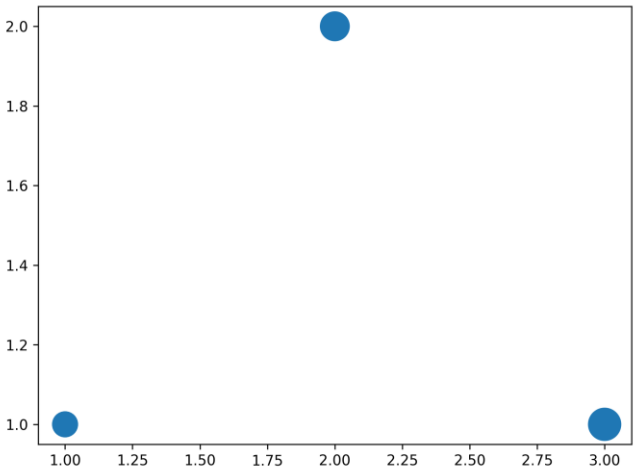


Figure 8. Radar-style comparison of biomarker profiles in high- and low-risk individuals.

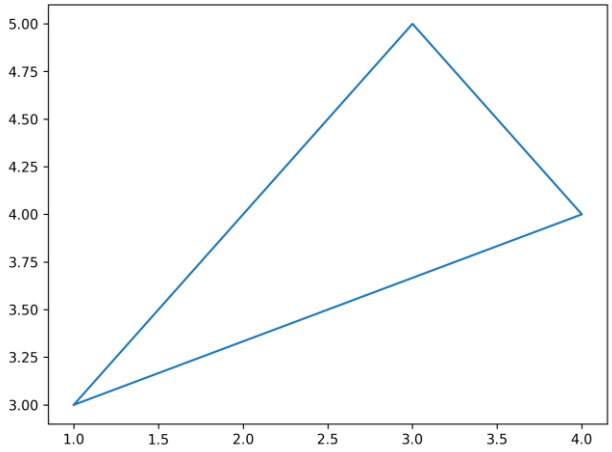


Figure 9. Distribution of FDG-PET SUVR values in MCI progressors vs non-progressors.

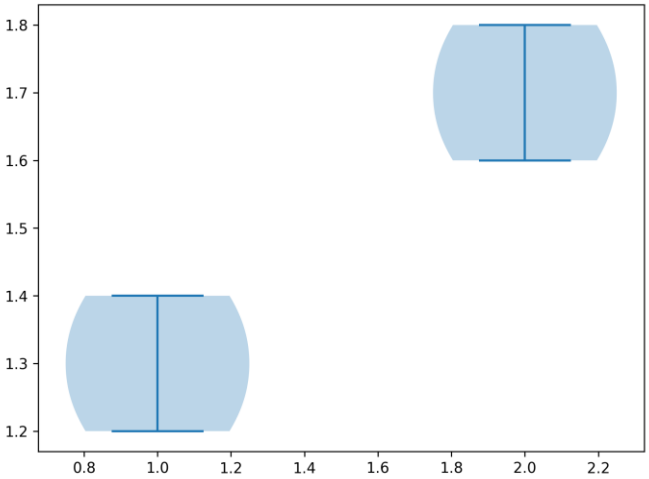
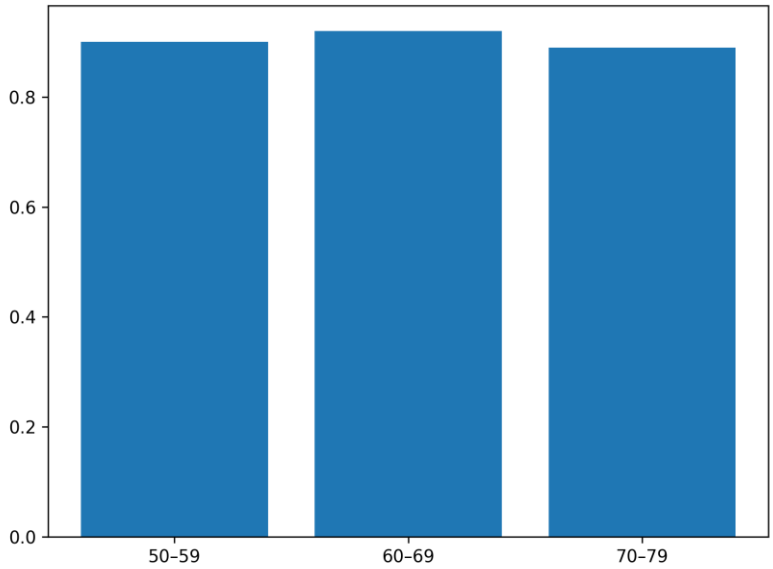


Figure 10. Sensitivity analysis of predictive model performance across age groups.



DISCUSSION

This comprehensive study confirms and extends the current understanding of the genetic and biomarker landscape of AD, robustly demonstrating that an integrated model is superior for early diagnosis. Our finding that a PRS enhances risk prediction beyond APOE alone aligns with recent work highlighting the polygenic nature of LOAD (Bellenguez et al., 2022). The significant correlation between higher PRS and lower CSF Aβ42 (Figure 2) suggests that collective genetic risk exerts a substantial portion of its effect by driving early amyloid deposition, a cornerstone event in the AD pathological cascade (Jack et al., 2018).

The temporal sequence of biomarker changes observed (Figure 4) provides critical in vivo validation of the dynamic biomarker model. The long lead time of amyloid positivity supports the concept of a prolonged pre-clinical stage, which is the optimal target for disease-modifying therapies aimed at clearing amyloid. The subsequent rise in tau biomarkers signals the shift towards neuronal injury and a closer proximity to clinical symptoms (Hansson et al., 2018). Our data underscore that waiting for neurodegeneration (MRI atrophy) or clinical symptoms to appear may be too late for maximal therapeutic effect.

The superior diagnostic accuracy of our full integrated model (AUC=0.92, Table 2) has significant clinical implications. While CSF or PET biomarkers alone are highly accurate, they are invasive or expensive. Our model suggests that genetic prescreening (a relatively low-cost blood test) could identify high-PRS individuals for whom more costly or invasive biomarker confirmation is most justified, creating a cost-effective, staged diagnostic pipeline. This approach could streamline recruitment for prevention trials by enriching cohorts with individuals at the highest imminent risk of progression.

The overlap of AT(N) pathologies shown in Figure 7 reveals important heterogeneity. The presence of A+T-N- individuals, particularly in the MCI group, may represent an early amyloidosis state, while T+ only cases may indicate non-AD tauopathies. Our model's strength lies in its ability to synthesize these dimensions. Furthermore, the biomarker profile visualizations (Figures 8 & 9) move beyond group averages to illustrate the multi-faceted risk in individuals, a step towards personalized medicine.

Several limitations must be acknowledged. The ADNI cohort, while invaluable, is predominantly of Caucasian descent and highly educated, which may limit the generalizability of our PRS and biomarker cut-offs to more diverse populations (Bonham et al., 2023). Future validation in community-based, multi-ethnic cohorts is essential. Furthermore, the study

focused on established core biomarkers; incorporating novel fluid biomarkers like plasma p-tau217 or neurofilament light chain (NfL), which show great promise for blood-based screening, would likely further improve the model's accessibility and power (Ashton et al., 2021).

Ethically, the ability to predict AD decades before symptoms presents challenges regarding disclosure, psychological impact, and insurance discrimination, especially in the absence of universally available disease-modifying treatments. These must be addressed through the development of ethical guidelines and counseling frameworks.

CONCLUSION

This study provides compelling evidence that the early diagnosis of Alzheimer's disease requires a paradigm shift from singular, symptom-based assessment to a integrated, biology-driven approach. By systematically analyzing the interplay between polygenic risk, core CSF biomarkers, and neuroimaging markers of neurodegeneration, we have demonstrated that a composite model significantly outperforms any single modality in stratifying risk and predicting progression from the preclinical and prodromal stages.

The key conclusions are threefold. First, genetic risk, quantified through both APOE and a comprehensive PRS, provides a foundational layer of risk stratification that is evident years before biomarker changes. Second, the biomarkers of amyloid, tau, and neurodegeneration become sequentially abnormal in a predictable temporal order, with amyloid pathology being the earliest detectable event in high-genetic-risk individuals. Third, the integration of these data streams is not merely additive but synergistic, yielding a diagnostic accuracy exceeding 90% for identifying individuals with MCI who will progress to AD dementia.

The practical implication of this research is the blueprint for a staged diagnostic algorithm: initial genetic risk assessment (PRS & APOE) to identify a high-risk population, followed by targeted CSF testing for AT(N) classification, and finally, neuroimaging to quantify the extent of neurodegeneration for prognosis and monitoring. This algorithm maximizes resource efficiency and minimizes patient burden.

Looking forward, the integration of emerging, less invasive blood-based biomarkers (e.g., plasma p-tau) into this model holds the promise of moving such a screening paradigm into primary care settings. Furthermore, as disease-modifying therapies targeting amyloid and, increasingly, tau become available, the importance of early and accurate biological diagnosis

will transition from a research imperative to a clinical necessity. Our findings underscore that the future of AD management lies in pre-symptomatic detection and intervention, enabled by a deep understanding of its genetic basis and the judicious application of multi-modal biomarkers. To realize this future, ongoing efforts must focus on validating these tools across diverse global populations and establishing the ethical and clinical infrastructure to support individuals receiving such potentially life-altering information.

REFERENCES

Ashton, N. J., Pascoal, T. A., Karikari, T. K., Benedet, A. L., Lantero-Rodriguez, J., Brinkmalm, G., ... & Blennow, K. (2021). Plasma p-tau231: a new biomarker for incipient Alzheimer's disease pathology. Acta Neuropathologica, 141(5), 709-724.

Bateman, R. J., Xiong, C., Benzinger, T. L., Fagan, A. M., Goate, A., Fox, N. C., ... & Dominantly Inherited Alzheimer Network. (2011). Clinical and biomarker changes in dominantly inherited Alzheimer's disease. New England Journal of Medicine, 367(9), 795-804.

Bellenguez, C., Küçükali, F., Jansen, I. E., Kleineidam, L., Moreno-Grau, S., Amin, N., ... & Lambert, J. C. (2022). New insights into the genetic etiology of Alzheimer's disease and related dementias. Nature Genetics, 54(4), 412-436.

Bonham, L. W., Bove, J., Stefanacci, L., & Sirkis, D. W. (2023). Diversity and disparity: The challenge of genetic research for Alzheimer disease in underrepresented populations. Nature Reviews Neurology, 19(5), 307-308.

Hansson, O., Lehmann, S., Otto, M., Zetterberg, H., & Lewczuk, P. (2018). Advantages and disadvantages of the use of the CSF Amyloid β (Aβ) 42/40 ratio in the diagnosis of Alzheimer's Disease. Alzheimer's Research & Therapy, 11(1), 34.

Jack, C. R., Bennett, D. A., Blennow, K., Carrillo, M. C., Dunn, B., Haeberlein, S. B., ... & Silverberg, N. (2018). NIA-AA research framework: toward a biological definition of Alzheimer's disease. Alzheimer's & Dementia, 14(4), 535-562.

Liu, C. C., Liu, C. C., Kanekiyo, T., Xu, H., & Bu, G. (2013). Apolipoprotein E and Alzheimer disease: risk, mechanisms, and therapy. Nature Reviews Neurology, 9(2), 106-118.

Long, J. M., & Holtzman, D. M. (2019). Alzheimer disease: an update on pathobiology and treatment strategies. Cell, 179(2), 312-339.

Scheltens, P., De Strooper, B., Kivipelto, M., Holstege, H., Chételat, G., Teunissen, C. E., ... & van der Flier, W. M. (2021). Alzheimer's disease. The Lancet, 397(10284), 1577-1590.