Toward Preventive Alzheimer's Risk Screening with Cell-Subtype-Aware Brain-Blood Graph Neural Network

Claire Xu
The Harker School
claire.78.xu@gmail.com

Abstract

Early Alzheimer's disease (AD) pathology begins decades before symptoms emerge, yet over 75% of the at-risk population lacks access to non-invasive screening methods. Current diagnostic tools like PET imaging and cerebrospinal fluid sampling are costly, invasive, and poorly suited for large-scale, proactive brain health monitoring. This research introduces a cell-subtype-aware brain-blood gene modeling framework that reframes AD assessment from reactive diagnosis to preventive risk evaluation for sustained cognitive health. Using graph neural networks, blood RNA-seq profiles are anchored to sex-specific, single-cell brain transcriptomics across neuronal layers, enhancing biological fidelity and interpretability. Explicit control of APOE4 genotype, age, sex, and education preserves meaningful variation while suppressing systemic noise. Gene set enrichment analysis confirmed pathways in neurodegeneration, inflammation, oxidative phosphorylation, and sensory function, with brain-derived signals reproducibly detected in blood. Sex-stratified analyses revealed female-specific signatures linked to addiction and mood regulation, pathogen-driven immune responses, and nutrient-based neuroprotection. This research identifies a blood-based gene panel for AD risk: GFAP, TREM2, C1QC, C1QB, PLCG2, TXNIP, CD163, CAMK1D, DAPK1, CCND3, LRP10, and COQ10A. By coupling fine-grained brain biology with interpretable AI, this work enables equitable, population-scale early risk identification, supporting proactive interventions to maintain cognitive function and delay disease onset.

Introduction

Alzheimer's Disease (AD) affects one in nine individuals aged 65 and older in the United States, with women being disproportionately affected, twice as likely to develop AD as men ("2024 Alzheimer's Disease Facts and Figures" 2024). Yet, AD pathology begins decades before clinical symptoms appear, progressing silently through vulnerable brain regions (Beason-Held et al. 2013). A striking case documented onset in a 19-year-old male with no known family history (Alzra 2023), highlighting the need for early, proactive detection strategies. Current diagnostic tools, such as PET imaging (Gentry et al. 2024) and cerebrospinal fluid (CSF) biomarkers (Bouwman et al. 2022) (Adeyemi et al. 2025), are both invasive and costly, making them unsuitable for large-scale, pre-symptomatic screening.

Critically, over 75% of the U.S. population falls between the ages of 20 and the typical onset of early AD pathology. This biologically susceptible but routinely overlooked population has a great potential to benefit from equitable, non-invasive screening tools aimed at identifying early risk and guiding preventive brain health interventions.

Figure 1: AD progression timeline

To address this need, I developed a confounder-aware, biologically grounded modeling approach that aligns blood transcriptomic with brain single-cell gene expression to detect early, subtle signals of Alzheimer's risk. By anchoring blood data in sex-specific, cell-type-aware representations of brain biology and modeling confounding variables such as age, APOE genotype, and education, this approach improves reliability and interpretability of the early detection.

This approach also identifies candidate blood-based gene signals linked to smell (OR2V1 and OR8B2, associated with olfactory bulb–related loss of smell) and vision impairment (ZNF503, associated with retinal dysfunction), supplementing blood-based early markers. These findings support the proposal to integrate low-cost, accessible sensory assessments to enhance early screening access. By focusing on trustworthy signal extraction and inclusive design, this work contributes to the foundation for scalable, preventive brain health systems that can reach the 75% of the population often excluded from the traditional AD diagnostics.

Copyright © 2025, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

Existing Research and Limitations

Cerebrospinal Fluid and PET-Based Approaches

Cerebrospinal fluid (CSF) assays and PET imaging remain the clinical gold standards for AD pathology, detecting amyloid and tau biomarkers with high sensitivity. However, these methods are costly, invasive, restricted to specialized centers, and detect signals only after protein misfolding or plaque formations, limiting their utility for population-wide proactive and preventive screening (Chen et al. 2024).

Biological innovations in 2025 have demonstrated the diagnostic power of CSF and imaging-based approaches. Adeyemi et al. (2025) combined CSF Aβ42, p-tau, and hippocampal atrophy, achieving a disease correlation coefficient of 0.596. Similarly, Sriman et al. (2025) integrated CSF biomarkers with PET imaging, achieving 92% accuracy in an AI-driven early detection model. While these methods demonstrate high performance, they are fundamentally limited by invasiveness, high cost, and restricted access to imaging facilities, making them unsuitable for large-scale or routine population screening.

Sun et al. (2024) developed a multimodal AI model that predicts CSF Aβ levels, integrating MRI-derived brain atrophy, EEG features, and behavioral assessments, eliminating the need for invasive lumbar puncture while maintaining high diagnostic value. Hata et al. (2025) applied portable EEG with machine learning to detect CSF amyloid and tau biomarker status, achieving 74.1% accuracy (AUC = 0.80) for amyloid and 73.1% (AUC = 0.80) for phosphorylated tau, demonstrating a non-invasive, cost-effective screening solution. While these advances mark significant progress, both methods are limited by reliance on specialized equipment (MRI or EEG) and moderate diagnostic accuracy, restricting their adoption in large-scale, routine population screening efforts.

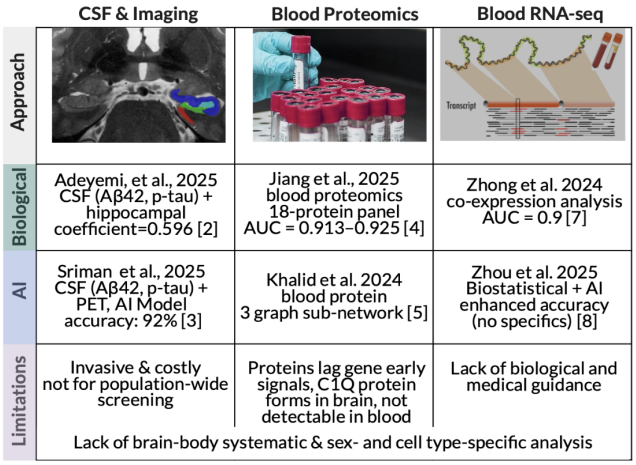

Figure 2: 2024-2025 innovations and limitations on CSF and non-invasive early AD detection

Blood Proteomics Approaches

Recent studies in 2024–2025 demonstrate that blood proteomics offers strong diagnostic performance for Alzheimer's detection. However, it faces critical limitations: many early protein markers, such as brain-synthesized C1Q, may not be detectable in peripheral blood during preclinical stages. Protein expression typically occurs downstream of gene activity, making it less sensitive to the earliest molecular changes. Additionally, reliance on high-dimensional proteomic assays presents challenges for routine clinical implementation.

Jiang et al. (2025) developed an 18-protein panel from blood proteomics, achieving an AUC between 0.913 and 0.925. The panel was derived using co-expression network analysis and captures distinct proteomic changes linked to AD progression. While outperforming some existing plasma ATN markers, the model depends on high-dimensional proteomic assays and complex feature selection, which may challenge implementation in routine clinical settings. Khalid et al. (2024) applied graph sub-network analysis to blood proteins for AD prediction. However, proteomic signals often lag transcriptomic changes.

Blood Transcriptomics Approaches

Recent research in 2024 explores blood RNA-seq as a promising direction for non-invasive Alzheimer's diagnosis, though limitations remain around biological grounding and interpretability. Zhong et al. (2024) performed transcriptomic profiling on whole-blood samples from Alzheimer's disease (AD), mild cognitive impairment (MCI), and control cohorts using bulk RNA-sequencing data from multiple GEO datasets. By applying machine learning models (e.g., XGBoost, Random Forest, SVM) to gene expression data, they achieved classification AUCs between 0.7 and 0.9 depending on the model and cohort size. The key limitation is the lack of brain-blood biological grounding, making RNA-based models prone to overfitting on blood-specific noise rather than early brain pathology. Zhou et al. (2025) incorporated biostatistical priors to enhance AI-driven detection, though 50 specific performance metrics remain unclear. These methods lack rigorous integration of biological knowledge such as cell-type, sex-specificity, or brain pathology.

Limitations of the Existing Research and Proposed Solutions

Existing approaches are limited by invasiveness, high cost, or delayed detection, often identifying signals only after protein misfolding or plaque formation. Blood-based methods, while non-invasive, frequently lack alignment with neuronal pathology, omit sex-specific analysis, and are not guided by underlying biological mechanisms.

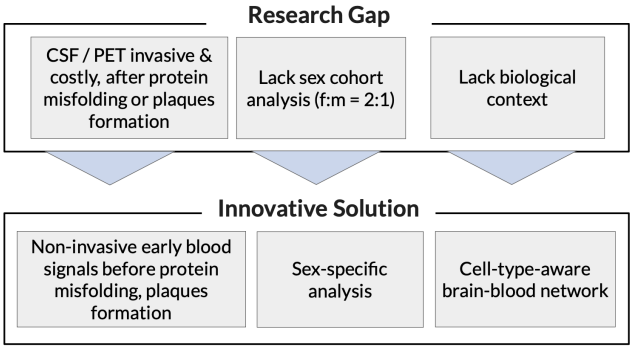

Figure 3: Identified research gap and solution design

Methods

In response to current research limitations, I developed a graph neural network that models brain- gene interactions using sex-specific, stratified single-cell transcriptomics to identify early Alzheimer's disease (AD) markers through interpretable, biologically grounded AI. This scalable, non-invasive approach enables more precise screening of individuals at risk for AD.

Datasets Description and Quality Control

I used single-cell brain transcriptomics from Cellxgene (Cellxgene Data Portal, n.d.), including 282,930 excitatory neurons, 79,294 inhibitory neurons, 18,890 astrocytes, and 1,621 microglia. Peripheral blood RNA-seq data from Gene Expression Omnibus (GEO: GSE63060 and GSE63061) ("GEO Accession Viewer," n.d.) provided gene expression profiles from AD and control individuals.

Clean data drives AI accuracy. To ensure the reliability and biological validity of downstream analyses, all datasets underwent rigorous preprocessing. Low-quality cells and genes were filtered to remove noise, and log-transformed normalization was applied so that the range of gene expression values becomes more consistent and comparable across different genes, enabling the AI model to more reliably detect true biological changes.

Dimensionality reduction (PCA, UMAP) reduced feature complexity while preserving important signals, and Synthetic Minority Oversampling Technique (SMOTE) was applied to balance training classes (normal vs. early AD), helping prevent the model from overfitting to the majority class.

Confounder Control

Confounding variables, such as APOE4 status ("APOE E4: The Most Prevalent yet Understudied Risk Factor for Alzheimer's Disease - ScienceDirect," n.d.), age, sex (Frigerio et al. 2019), and education level (Xu et al. 2015), are known to significantly influence gene expression and Alzheimer's disease (AD) progression. For example, APOE4 increases AD risk and alters neuroinflammation patterns, while aging and education affect cognitive resilience and gene activity patterns. Failing to account for these variables can obscure early markers.

To address this, I stratified multi-dimensions, sex, APOE4 status, age, and education level, into the model architecture. This stratified modeling reduces dilution of AD-relevant signals caused by cellular heterogeneity and demographic variability. As a result, differential expression (DE) sensitivity is enhanced, and biologically meaningful variation is preserved. DE analysis revealed minimal signals in mixed-cell data and significantly more pronounced gene changes in cell-type-specific layers.

1. Cell-Subtype Stratification

In the Cellxgene dataset, neuronal cells are classified into excitatory (Ex01–Ex13) and inhibitory (In1–In7) subtypes, each corresponding to specific cortical layers and marker gene profiles. Excitatory subtypes align with cortical layers L2–L6 and express markers such as CUX2, RORB, FEZF2, and THEMIS, while inhibitory subtypes are defined by canonical interneuron markers including LHX6, PVALB, SST, and ADARB2. Please refer to Table 1 below. This layer–subtype mapping provides a reference framework for linking cell-type–specific gene expression to Alzheimer's disease–associated molecular signatures.

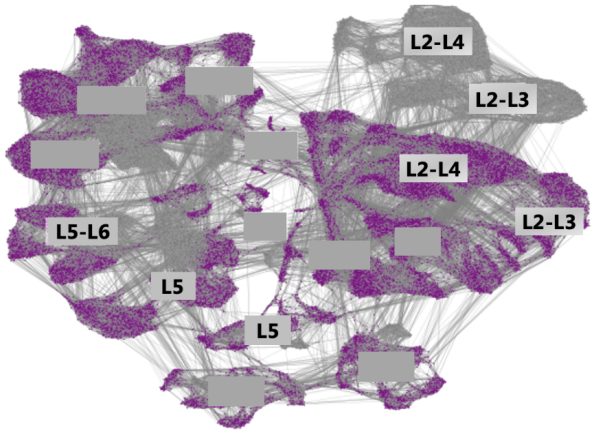

Figure 4: Graph of neuronal layers based on single-cell gene expression profiles. Labels denote distinct neuronal layers, with purple dots indicating Alzheimer's disease (AD) cells (darker shades represent higher disease-related signal) and grey indicating control cells. Clustering patterns show that the cell layers' identity exerts a stronger influence on gene expression profiles than AD status, emphasizing the need to account for biological context when detecting disease-associated signals.

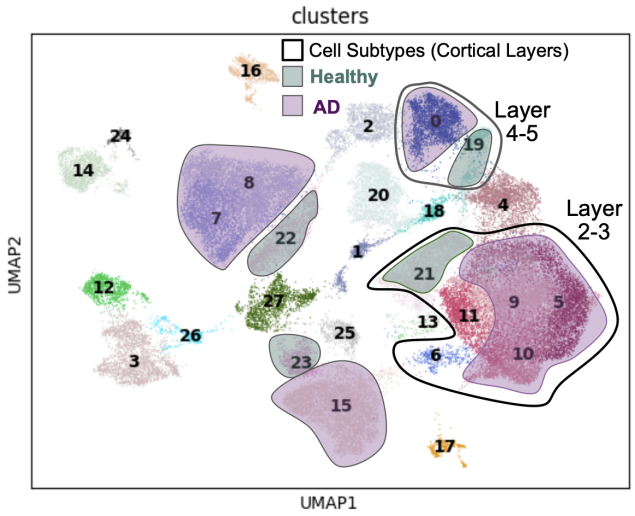

Figure 5: UMAP projection of neuron cells clustered by gene expression similarity. Each lassoed region corresponds to a distinct cell subtype, with numeric labels indicating cluster IDs. Purple shades indicate clusters dominated by Alzheimer's disease (AD) cells, and green shades indicate clusters dominated by control cells. Co-localization of cells from specific neuronal layers (e.g., L2–3) reflects strong cell-type–driven gene expression patterns, which can mask AD-associated signals.

Neuron Excitatory Layers	Neuron Inhibitory Layers
Ex01_CUX2-LAMP5 (L2-L3) Ex02_CUX2-COL5A2 (L2-L4) Ex03_RORB-MME (L4-L5) Ex04_RORB-GABRG1 (L4-L5) Ex05_RORB-ADGRLA (L5) Ex06_RORB-PCP4-RPRM (L5) Ex07_RORB-PCP4 (L5) Ex08_FEZF2-PCP4-ROBO3 (L5) Ex09_FEZF2-ADRA1A (L5b) Ex10_THEMIS (L5-6) Ex11_THEMIS-NTNG2 (L6) Ex12_FEZF2-SYT6 (L6) Ex13_FEZF2-SEMA3D (L6b)	In1_LHX6-PVALB In2_LHX6-PVALB-Chandelier In3_LHX6-SST In4_LHX6-SST-NPY In5_LHX6-ADARB2-LAMP5 In6_ADARB2-LAMP5 In7_ADARB2-CALB2

Table 1: Neuronal Layers in the Human Brain

2. Sex-Specific Stratification

Given that males and females exhibit distinct hormone trajectories across the lifespan, I examined whether these differences are reflected at the molecular level. Using UMAP clustering of single-cell transcriptomic data from brain neurons, my analysis revealed clear sex-specific gene expression patterns, with male and female cells forming distinct molecular profiles.

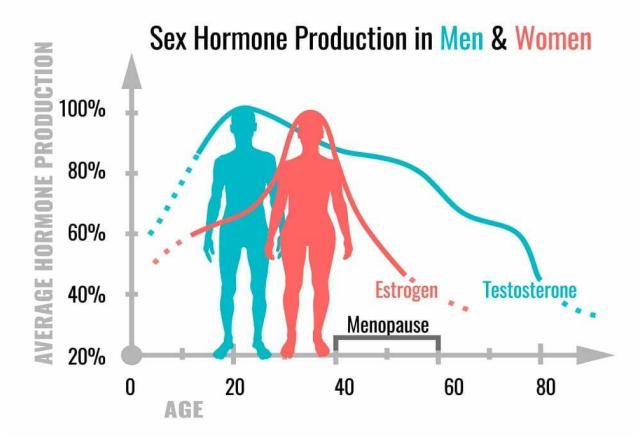

Figure 6: Sex hormone changes across lifespan. Men and women present distinct hormone change trajectory (LabMe, n.d.).

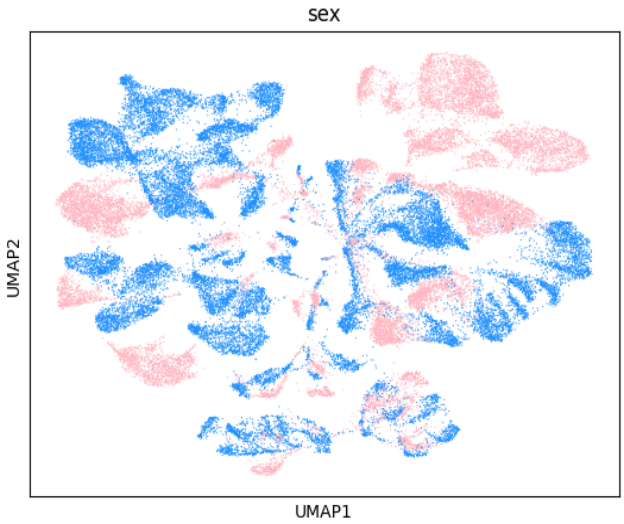

Figure 7: Distinct male (blue) & female (pink) molecular profiles in UMAP. Each dot represents one cell. Cells with similar gene expressions are co-located. The x-axis represents the 1st major variance, and the y-axis represents the 2nd major variance.

Brain-Blood Graph Neural Network Modeling

To enable early Alzheimer's risk screening that is biologically grounded and clinically interpretable, I developed a graph neural network (GNN) that models brain-blood transcriptomic profiles. The framework first anchors on the sex-specific, cell-type–aware representations of brain transcriptomics, since the Alzheimer's disease pathology is defined by changes in the brain. It then links these brain-based biomarkers to blood transcriptomic profiles to identify corresponding blood-based biomarkers. This way, the framework reduces the systemic noise and non-disease variation

expressed in the blood that circulates through the whole body. For example, although genes such as NCKAP1L, ATHL1, and FAM83F are differentially expressed in blood, their expression is barely detectable in the brain; therefore, they are excluded from this research.

The GNN network is explicitly designed to address population heterogeneity and biological confounding, two factors that can mask early disease signals in real-world datasets. Multiple GNN subnetworks are built, one for each sex (male, female) and cell type (neuronal cell subtype, microglia, astrocyte, blood) combination. Table 1 lists the excitatory and inhibitory neuronal cell subtypes.

Before building the GNN network, an adjacency matrix is constructed to extract the gene-gene relationships from Cellxgene AnnData and convert it into the edge_index format required by PyTorch Geometric to represent the interaction network.

In the GNN network, each node represents one gene and each edge represents the co-expression between genes. There are three layers in the GNN network. Two graph convolutional layers are constructed with PyTorch Geometric to handle graph connectivity. Relu nonlinear activation function is used to learn complex relationships. These two layers update node features by aggregating information from neighboring genes connected in the adjacency matrix, a process known as message passing. After one round of message passing, each node's features contain information from its 1-hop neighbors. After two rounds, each node encodes information from its 2-hop neighbors. After GCN layers, each gene node's representation reflects both its own expression and the structure of its gene-gene interaction network. Then the output passes to the 3rd layer which is a fully connected dense layer that computes per-gene logits for the classification task: marker vs. non-marker. Softmax activation function is used for binary classification. After various experimentation, a learning rate 0.0005 is verified to achieve highest accuracy.

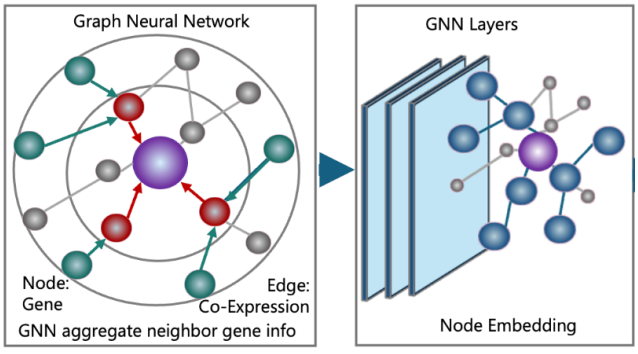

Figure 8: GNN architecture. The AI model was run on GPU A100 (NVIDIA-SMI 550.54.15, Driver Version: 550.54.15, CUDA Version: 12.4). Total RAM allocated was 89.6 gigabytes (partially consumed by the model).

The GNN model has been experimented on Chan Zuckerberg's Cellxgene AD single-cell transcriptomics datasets and Gene Expression Omnibus peripheral blood RNA-seq datasets (282,930 excitatory neurons, 79,294 inhibitory neurons, 18,890 astrocytes, and 1,621 microglia; GSE63060 and GSE63061 datasets). I randomly divided the dataset into 70% training, 15% validation, and 15% testing sets. The test set was strictly held out during training and hyperparameter tuning, such as selecting the learning rate. Given the large dataset size, this split provided sufficient representation for evaluation. To ensure reproducibility, I randomly repeated the data split three times and observed consistent results.
Genes are connected within the graph based on shared co-expression patterns across matched brain-blood subspaces. Through iterative message passing, the GNN amplifies features from genes with strong differential expression between early AD and control groups while reducing contributions from low-variance housekeeping genes. This ensures that disease-relevant signals are strengthened without being obscured by general background noise.

Results and Evaluation

Functional Enrichment of Early Marker Genes

The Gene Set Enrichment Analysis (GSEA) is derived from GNN outputs. Each GNN subnetwork corresponds to a specific analysis stratum defined by sex and cell subtype.

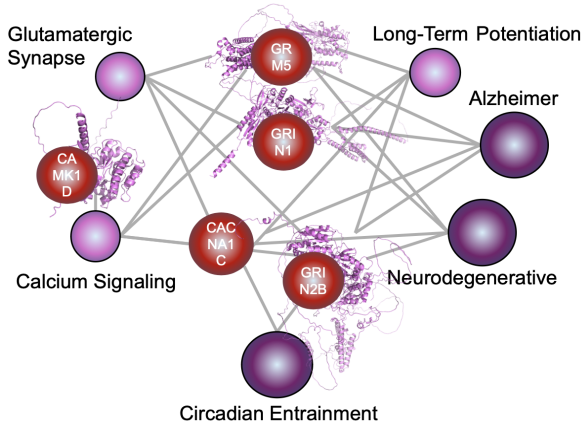

Figure 9: Selected AD risk marker genes from female brain and their associated KEGG pathways. Brain-derived signals serve as anchors for detecting corresponding patterns in blood, helping reduce systemic noise in peripheral data. The graph highlights biologically relevant pathways, including neurodegeneration, Alzheimer, calcium signaling, and circadian entrainment. Representative AD risk marker genes from female brain include (but are not limited to) GRM5, GRIN1, GRIN2B, CACNA1C, and CAMK1D.

The early AD marker genes identified across all subnetworks are then aggregated and analyzed for KEGG pathway enrichment. Significant pathways (FDR < 0.05; size 10-500) are reported. The pathway-gene network figures 9, 10, and 11 visualize these significant pathways and their overlapping GNN-prioritized marker genes. This analysis confirms that the marker genes are functionally enriched in pathways related to neurodegeneration, inflammation, and oxidative phosphorylation. Early gene signals in the brain translate to the blood, ensuring biological relevance beyond the blood's systemic noise. The blood-based AD risk marker gene panel includes GFAP, TREM2, C1QC, C1QB, PLCG2, TXNIP, CD163, CAMK1D, DAPK1, CCND3, LRP10, and COQ10A.

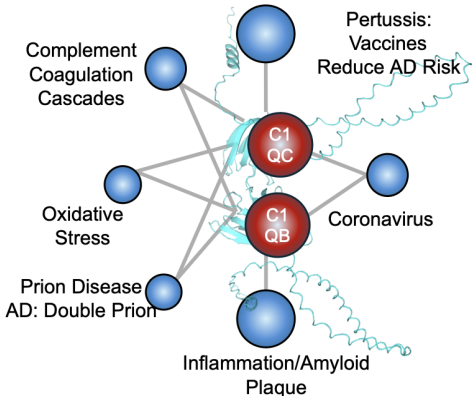

Figure 10: Selected AD risk marker genes from male brain and their associated KEGG pathways. Together with the female brain profile, these male-specific brain signals provide a complementary reference for cross-validating blood-based findings and controlling for sex-specific differences. The graph highlights biologically relevant pathways, including oxidative stress, inflammation/amyloid plaque, etc. Representative AD risk marker genes from male brain include (but are not limited to) C1QC, and C1QB.

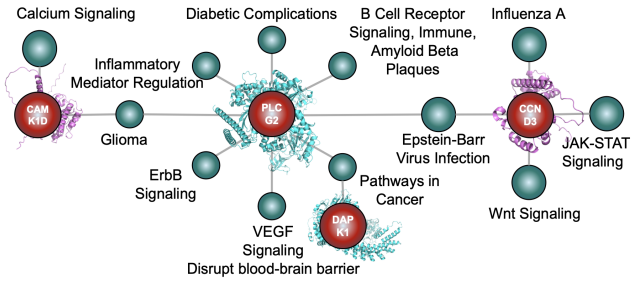

Figure 11: AD risk marker genes and KEGG pathways identified in blood samples, cross-referenced with both female and male brain–derived profiles. This alignment to brain-based anchors reduces systemic noise and enhances the specificity of blood biomarkers for Alzheimer's disease. Some of the known AD genes (BIN1, CR1, ABCA7, CLU, etc.) are not selected. These genes are well-known from genetic studies, but they didn't show strong or consistent RNA signals in the blood. My model filtered for genes with clear early-stage differences and reliable expression in blood data. Representative AD risk marker genes from blood include (but are not limited to) CAMK1D, PLCG2, DAPK1, and CCND3.

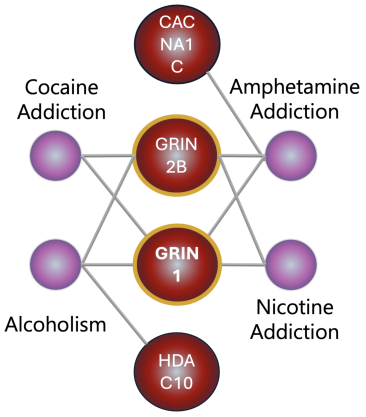

Figure 12: Female hormone influences the brain's reward system, increasing addiction.

Cognitive Risk and Longevity from Early Gene Signals

Early brain-blood gene signals reveal sex-specific pathways tied to cognitive decline, linking genetic markers to lifestyle and environmental factors for early risk reduction.

1. Addiction and mood in female

Sex-specific hormonal influences affect both neurological health and behavior. In females, hormone-driven changes in the thyroid and the brain's reward circuitry are linked to anxiety, weight changes, and addiction risk. Early marker genes illustrated in figure 10 map to pathways involved in cocaine, alcohol, amphetamine, and nicotine addiction.

2. Bacteria and Virus Fuel AD

Environmental pathogens can accelerate Alzheimer's progression through immune and inflammatory pathways. Gender-influenced infection routes, such as malaria, HPV, toxoplasmosis, and cytosolic DNA sensing, are associated with early marker genes. Preventive strategies include vaccination, avoiding vector exposure, safe food handling, and maintaining immune resilience.

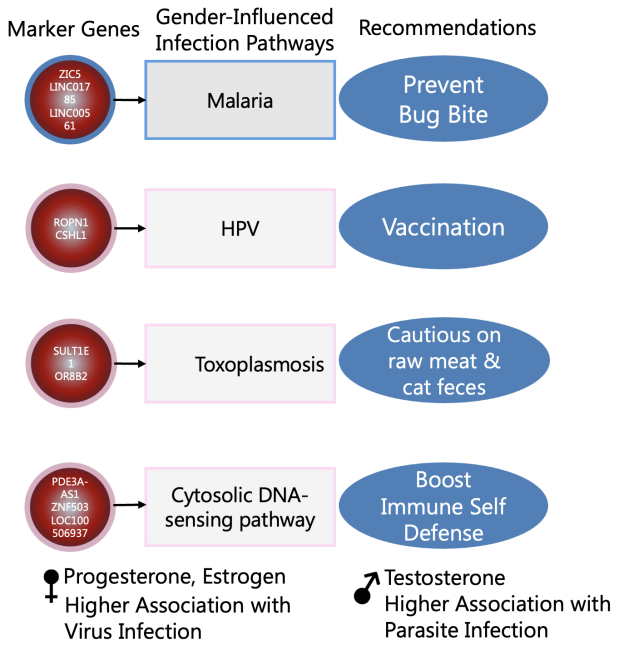

Figure 13: Environmental stressors

3. Neuron Protection from Toxic Effect

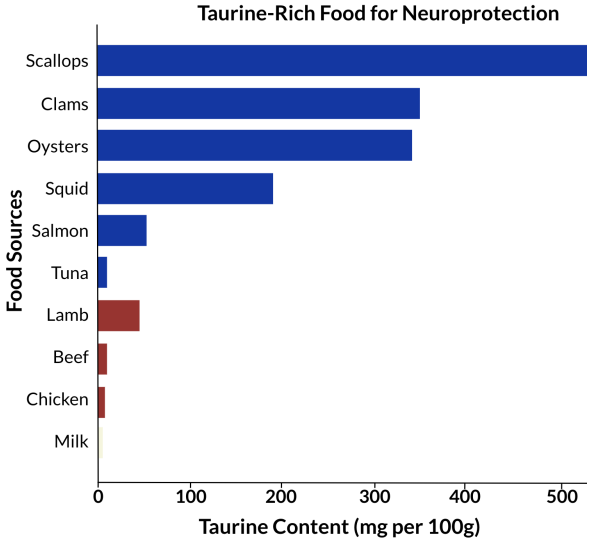

Figure 14: Taurine is a known therapeutic agent (Caine and Geracioti 2016)

Certain nutrients offer neuroprotective benefits against cellular toxicity. Taurine, a known therapeutic compound, is abundant in foods such as scallops, clams, oysters, squid, salmon, tuna, and lamb. These taurine-rich sources support neuronal health and mitigate neurodegenerative processes.

Applications and Future Work

Population-Scale Preventive Screening for Cognitive Health

This research transforms Alzheimer's detection from late-stage diagnosis to proactive risk assessment, enabling scalable identification of at-risk individuals years before symptoms emerge. By anchoring blood transcriptomic profiles to cell-subtype–resolved brain biology, it delivers biologically grounded, interpretable signals suitable for integration into routine health checkups, national screening programs, and community-based brain health initiatives. Coupled with low-cost cognitive, olfactory, or visual assessments, it offers an equitable, preventive pathway to preserve cognitive resilience—particularly for the 75% of the population in the preclinical age range who remain underserved by current detection methods.

Ethical and Implementation Considerations

Population-scale screening for neurodegenerative risk must be accompanied by safeguards to prevent unintended harms such as anxiety, stigma, or discrimination. Building trust requires transparency in how risk scores are generated, clear communication of uncertainty, and alignment with evidence-based preventive interventions.

Future Directions

Future work will focus on validating this cell-subtype-aware brain-blood modeling framework in diverse cohorts, integrating multimodal data sources, and exploring partnerships with public health systems for pilot implementations. Broader efforts will also investigate how systemic factors, such as peripheral inflammation or blood-brain barrier integrity, interact with brain gene expression in the earliest stages of Alzheimer's risk.

Acknowledgments

This work makes use of publicly available datasets from the Chan Zuckerberg Initiative's Cellxgene platform and the Gene Expression Omnibus (GEO). I gratefully acknowledge the individuals and their families who generously donated brain tissue and blood samples for scientific research, without whom this study would not have been possible.

References

"2024 Alzheimer's Disease Facts and Figures." 2024. *Alzheimer's & Dementia* 20 (5): 3708–821. https://doi.org/10.1002/alz.13809.

Adeyemi, Oluwatobi F., Penny Gowland, Richard Bowtell, Olivier Mougin, and Akram A. Hosseini. 2025. "Hippocampal Subfield Volume in Relation to Cerebrospinal Fluid Amyloid-ß in Early

Alzheimer's Disease: Diagnostic Utility of 7T MRI." *European Journal of Neurology* 32 (2): e70076. https://doi.org/10.1111/ene.70076.

Alzra. 2023. "Neurologists Report World's Youngest Alzheimer's Case." *Alzheimer's Research Association*, March 15. https://www.alzra.org/blog/neurologists-report-worlds-youngest-alzheimers-case/.

"APOE E4: The Most Prevalent yet Understudied Risk Factor for Alzheimer's Disease - ScienceDirect." n.d. Accessed August 9, 2025. https://www.sciencedirect.com/science/article/abs/pii/S1552526014024996.

Beason-Held, Lori L., Joshua O. Goh, Yang An, et al. 2013. "Changes in Brain Function Occur Years before the Onset of Cognitive Impairment." *The Journal of Neuroscience* 33 (46): 18008–14. https://doi.org/10.1523/JNEUROSCI.1402-13.2013.

Bouwman, Femke H., Giovanni B. Frisoni, Sterling C. Johnson, et al. 2022. "Clinical Application of CSF Biomarkers for Alzheimer's Disease: From Rationale to Ratios." *Alzheimer's & Dementia (Amsterdam, Netherlands)* 14 (1): e12314. https://doi.org/10.1002/dad2.12314.

Caine, Jonathan J., and Thomas D. Geracioti. 2016. "Taurine, Energy Drinks, and Neuroendocrine Effects." Review. *Cleveland Clinic Journal of Medicine* 83 (12): 895–904. https://doi.org/10.3949/ccjm.83a.15050.

Cellxgene Data Portal. n.d. "Cellxgene Data Portal." Accessed January 6, 2025. https://cellxgene.cziscience.com/.

Chen, Yixin, Murad Al-Nusaif, Song Li, et al. 2024. "Progress on Early Diagnosing Alzheimer's Disease." *Frontiers of Medicine* 18 (3): 446–64. https://doi.org/10.1007/s11684-023-1047-1.

Frigerio, Carlo Sala, Leen Wolfs, Nicola Fattorelli, et al. 2019. "The Major Risk Factors for Alzheimer's Disease: Age, Sex, and Genes Modulate the Microglia Response to Aβ Plaques." *Cell Reports* 27 (4): 1293-1306.e6. https://doi.org/10.1016/j.celrep.2019.03.099.

Gentry, Charisse, Michael Malek-Ahmadi, Susan Bolas, and Jose Pena. 2024. "PET Imaging in Alzheimer Disease: Pathology and Research Insights for Technologists." Continuing Education. *Journal of Nuclear Medicine Technology*, ahead of print, November 12. https://doi.org/10.2967/jnmt.124.268916.

"GEO Accession Viewer." n.d. Accessed August 10, 2025. https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE63060.

Hata, Masahiro, Yuki Miyazaki, Kohji Mori, et al. 2025. "Screening of Aβ and Phosphorylated Tau Status in the Cerebrospinal Fluid through Machine Learning Analysis of Portable Electroencephalography Data." *Scientific Reports* 15 (1): 2067. https://doi.org/10.1038/s41598-025-86449-2.

Jiang, Yuanbing, Hyebin Uhm, Fanny C. F. Ip, et al. 2025. "Identification of High-Performance Blood-Based Biomarkers for Early Screening and Staging of Alzheimer's Disease by Large-Scale Plasma Proteomic Profiling in Mild Cognitive Impairments." *Alzheimer's & Dementia* 20 (S2): e086052. https://doi.org/10.1002/alz.086052.

Khalid, Maryam, Fadeel Sher Khan, John Broussard, and Arko Barman. 2024. "Graph-Based Biomarker Discovery and Interpretation for Alzheimer's Disease." arXiv:2411.18796. Preprint, arXiv, November 27. https://doi.org/10.48550/arXiv.2411.18796.

LabMe. n.d. "Latest Research on Number 1 Hormone Testosterone." Accessed April 18, 2025. https://labme.ai/blogs/health/knee-replacement.

"Precise Prediction of Cerebrospinal Fluid Amyloid Beta Protein for Early Alzheimer's Disease Detection Using Multimodal Data - Sun - 2024 - MedComm - Wiley Online Library." n.d. Accessed August 8, 2025. https://onlinelibrary.wiley.com/doi/full/10.1002/mco2.532.

Sriman, B., M. Vigneshkumar, K. S. Dhinesh Kumar, J. Praveenkumar, and K. Suganya. 2025. "Predictive Precision Harnessing AI for Early Alzheimer's Detection." In *Deep Generative Models for Integrative Analysis of Alzheimer's Biomarkers*. IGI Global Scientific Publishing. https://doi.org/10.4018/979-8-3693-6442-0.ch009.

Xu, Wei, Lan Tan, Hui-Fu Wang, et al. 2015. "Meta-Analysis of Modifiable Risk Factors for Alzheimer's Disease." Cognitive Neurology. *Journal of Neurology, Neurosurgery & Psychiatry* 86 (12): 1299–306. https://doi.org/10.1136/jnnp-2015-310548.

Zhong, Huan, Xiaopu Zhou, Hyebin Uhm, et al. 2024. "Using Blood Transcriptome Analysis for Alzheimer's Disease Diagnosis and Patient Stratification." *Alzheimer's & Dementia* 20 (4): 2469–84. https://doi.org/10.1002/alz.13691.

Zhou, Peng, Zixiu Li, Feifan Liu, et al. 2025. "BAMBI Integrates Biostatistical and Artificial Intelligence Methods to Improve RNA Biomarker Discovery." *Briefings in Bioinformatics* 26 (2): bbaf073. https://doi.org/10.1093/bib/bbaf073.