Journal of Neonatal Surgery
ISSN(Online): 2226-0439
Vol.14 , Issue 30s (2025)
https://www.jneonatalsurg.com

OPEN ACCESS

Predictive Modeling Of Alzheimer's Disease Risk And The Simulated Long -Term Impact Of Covid-19 Severity.

Y. Ebenezer PonSelvi[1], Dr. V. Joseph Peter[2]

[1]Research Scholar (Part–Time), 22222102282015, Research Department of Computer Science,
Kamaraj College (Autonomous), Thoothukudi, Affiliated to (Manonmaniam Sundaranar University, Tirunelveli-12)
Email:ID: jreeben@gmail.com

[2]Associate Professor and Head, Research Department of Computer Science,
Kamaraj College (Autonomous), Thoothukudi, Affiliated to (Manonmaniam Sundaranar University, Tirunelveli-12)
Email:ID: dr.josephpeter@kamaraj.college.ac.in

Cite this paper as: Y. Ebenezer PonSelvi, Dr. V. Joseph Peter (2025) Predictive Modeling Of Alzheimer's Disease Risk And The Simulated Long -Term Impact Of Covid-19 Severity..... Journal of Neonatal Surgery, 14, (30s) 1172-1180

Received:07.05.2025, Accepted: 01 06.2025, Published: 23.06.2025

ABSTRACT

Alzheimer's disease (AD) is a profoundly challenging neurodegenerative disorder[2], and an expanding body of clinical literature increasingly links the post-acute sequelae of COVID-19 to long-term neurological and cognitive deficits. This computational inquiry investigates the potential impact of long-COVID conditions on the latent risk of developing AD using an advanced machine learning framework. A comprehensive clinical Alzheimer's disease dataset was leveraged for model training and feature extraction[3], while empirical insights from a COVID-19 symptom and cognitive impact dataset were utilized to construct multi-tiered simulated patient profiles. Three distinct classification algorithms—Random Forest[2], Logistic Regression[3], and XGBoost—were trained to predict AD likelihood, with the XGBoost Classifier emerging as the superior predictive model[7]. Feature importance analysis highlighted functional assessments, Activities of Daily Living (ADL), and specific cognitive state examinations as the primary computational drivers of predicted decline[2]. By deploying the trained models against hypothetical post-COVID profiles ranging from mild to severe impairment, a distinct escalation in the predicted Alzheimer's risk probability was observed as post-COVID health markers deteriorated. These findings suggest a compelling structural vulnerability for accelerated cognitive decline in individuals experiencing severe long-term post-COVID conditions[9]. However, because this machine learning simulation serves as an indirect proxy[2], these results underscore an urgent mandate for prospective longitudinal clinical studies to establish definitive causal pathways between viral-induced sequelae and neurodegenerative outcomes[8]..

Keywords: Alzheimer's Disease, Neurodegenerative disorder, Post-COVID profiles, Cognitive sequelae, Predictive Modeling, Machine Learning

INTRODUCTION

Alzheimer's disease (AD) is a progressive neurodegenerative disorder and the leading cause of dementia globally, imposing immense emotional and financial burdens on families and healthcare systems[5]. The etiology of Alzheimer's is multi-factorial, stemming from an intricate interplay of advanced age, lifestyle habits, chronic metabolic or cardiovascular conditions, and genetic predispositions—most notably the APOE4 allele[1,3,5]. Uncovering how these diverse risk factors intersect to accelerate cognitive decline is crucial for establishing effective early-warning systems and preventative clinical protocols[2].

Concurrently, the medical landscape has been fundamentally altered by the post-acute sequelae of SARS-CoV-2, commonly known as "long COVID". Beyond its initial respiratory presentation, COVID-19 frequently manifests as a systemic condition impacting the central nervous system[8]. Many survivors experience prolonged neurological symptoms, including severe chronic fatigue, persistent headaches, and a debilitating cognitive haze frequently termed "brain fog"[8]. Neurobiological evidence suggests these lingering symptoms are driven by persistent neuroinflammation, microvascular injury, glial activation, and systemic immune dysregulation[8].

This computational research is driven by a critical, urgent question: Do the long-term neurological sequelae of COVID-19 accelerate or exacerbate the onset of Alzheimer's pathology[8]? A striking pathophysiological overlap exists between post-COVID neuroinflammation and the classic mechanics of Alzheimer's decline. Chronic inflammatory cascades and

compromised blood-brain barrier integrity create a neurotoxic microenvironment where memory degradation and cellular damage can thrive[2]. Exploring how post-COVID physiological changes interact with an individual's baseline Alzheimer's risk profile has therefore become a vital investigative priority[9]..

1.1 The Machine Learning and Predictive Simulation Pipeline

To systematically investigate this intersection, this study develops a multi-stage, dual-stream machine learning pipeline that models and simulates patient risk trajectory across the following functional phases,

1.1.1 The Predictive Modeling Stream:

Utilizing a comprehensive clinical Alzheimer's dataset, data was cleaned, preprocessed, and evaluated through three distinct classification algorithms: Logistic Regression, Random Forest, and Extreme Gradient Boosting (XGBoost) [1]. Among these, the highly optimized XGBoost model demonstrated superior predictive accuracy[4,8]. Feature importance analysis revealed that an individual's Functional Assessment score, performance on the Mini-Mental State Examination (MMSE), and capability to perform Activities of Daily Living (ADL) served as the primary indicators of cognitive decline—carrying significantly more weight than isolated genetic markers or basic lifestyle variables.

1.1.2 The Clinical Simulation Stream:

In parallel, real-world granular metrics from a secondary COVID-19 symptom dataset tracking cognitive slowdown, memory metrics, and physiological fatigue were used to construct multi-tiered health degradation profiles representing mild, moderate, and severe long-COVID states[9].

1.1.3 The Risk Shift Assessment Engine:

The framework converges by utilizing the real-world trends from the COVID-19 dataset to systematically adjust matching clinical features within the baseline Alzheimer's model[9]. By passing these step-by-step synthetic profiles through the trained XGBoost classifier, the pipeline functions as a predictive simulation engine, calculating exactly how an individual's risk probability shifts as post-viral health markers worsen.

While this predictive architecture operates as a digital simulation rather than an empirical *in vivo* medical study, it offers a stark, data-driven warning regarding the long-term cognitive toll of the pandemic. The accelerated risk curves output by the model demonstrate a clear vulnerability toward rapid cognitive deterioration among individuals severely impacted by long COVID[9]. Ultimately, these findings underscore the necessity for longitudinal clinical studies to track post-COVID cohorts, allowing us to definitively map the long-term bridge between viral neuroinflammation and irreversible neurodegeneration[8].

2. RELATED WORK

The intersection of viral immunology, systemic metabolic dysfunction, and neurodegeneration represents a critical frontier in contemporary clinical data science. This section reviews existing literature across three primary pillars: the deployment of machine learning in dementia diagnostics, the neurological footprint of post-acute COVID-19 syndrome, and the precise predictive role of clinical, functional, and metabolic biomarkers.

2.1 Machine Learning in Neurodegenerative Diagnostics

Historically, diagnosing Alzheimer's Disease (AD) relied heavily on longitudinal clinical observation, costly neuroimaging (such as PET scans), and invasive cerebrospinal fluid (CSF) draws. Recently, literature has shifted decisively toward computational intelligence and predictive modeling. Studies deploying tree-based ensembles, such as Random Forest and Gradient Boosting architectures, have demonstrated that structured, non-invasive clinical records can predict diagnostic outcomes with remarkable accuracy.

Researchers increasingly favor non-linear models like XGBoost because they naturally capture complex, high-dimensional interactions between a patient's demographic factors and cognitive scores—nuances that traditional linear regressions miss. In clinical machine learning pipelines, maintaining a balanced evaluation framework using complementary metrics like precision (to minimize false positives/over-diagnosis) and recall (to minimize false negatives/missed cases) is widely established as a prerequisite for deploying models into diagnostic workflows. This study leverages these high-performance architectures to establish a reliable predictive engine[8].

2.2 The Neurological Fingerprint of Post-Acute COVID-19 Syndrome

Since the onset of the global pandemic, clinical studies have documented persistent, long-term cognitive deficits in survivors, widely classified under Post-Acute Sequelae of SARS-CoV-2 (PASC), or 'Long COVID.' Brain autopsy reports and neuroimaging studies have highlighted that severe viral infections can trigger persistent neuroinflammation and microvascular injury[5]. Systemic inflammatory cytokines are capable of breaching the blood-brain barrier (BBB), leading to microglial activation and astrocytic reactivity that mirror the early, prodromal stages of neurodegeneration[8]. Clinically, post-COVID cohorts frequently present with profound executive dysfunction, memory deficits, and chronic fatigue[9]. This phenotypic overlap provides a compelling biological rationale for investigating whether post-viral physiological degradation can act as a catalyst, accelerating an individual's transition along the dementia risk trajectory[8].

2.3 The Predictive Role of Specific Clinical and Metabolic Features:

To construct robust machine learning models, contemporary research heavily interrogates specific categories of clinical features. These parameters serve as essential proxies for structural brain health, daily independence, and metabolic stability[2,3].

2.3.1 Cognitive and Functional Biomarkers:

Literature consistently validates behavioral and cognitive scores as the most powerful indicators of immediate neurodegenerative risk[2].*[5] [5] [5]*

2.3.2 Functional Assessment & Activities of Daily Living (ADL):

A patient's capacity to maintain physical autonomy and perform routine tasks (e.g., managing finances, self-care) is deeply tied to structural integrity in the frontal and temporal lobes[2]. Gradual degradation in ADL performance is often the earliest overt sign of clinical dementia[2].

2.3.3 Mini-Mental State Examination (MMSE):

As a standardized, 30-point clinical tool, the MMSE measures orientation, attention, recall, and language. Existing predictive models rely heavily on MMSE scores; subtle drops within the 20–24 range regularly signal a transition from mild cognitive impairment to established Alzheimer's disease.

2.3.4 Memory Complaints:

Subjective and family-reported memory complaints serve as vital early-warning indicators, often capturing subtle, real-world cognitive slips before they register on formal psychometric tests.

2.3.5 Metabolic and Physiological Indicators:

Beyond cognitive tests, peripheral metabolic and vascular markers are increasingly recognized as critical drivers of neurodegenerative acceleration[**1,5**]:

2.3.6 Lipid Metabolism (Total Cholesterol, HDL, LDL, and Triglycerides):

The human brain is a lipid-dense organ, relying on tightly regulated cholesterol homeostasis for myelin maintenance and synaptic repair[6]. High levels of Low-Density Lipoprotein (LDL) and depressed High-Density Lipoprotein (HDL) compromise the blood-brain barrier, impair the clearance of toxic amyloid-beta plaques, and provoke neurovascular inflammation[5].

2.3.7 Blood Pressure (Systolic and Diastolic BP):

Mid-life hypertension damages cerebral microvessels, leading to chronic hypoperfusion and white matter hyperintensities[1]. This vascular stress works in tandem with neurodegenerative pathologies, compounding total risk[1].

2.3.8 Body Mass Index (BMI) & Lifestyle Factors:

Extremes in BMI, paired with poor sleep quality, high alcohol consumption, and physical inactivity, induce systemic metabolic stress and insulin resistance, both of which are strongly linked to accelerated brain aging[1].

2.4 Demographic and Genetic Controls:

2.4.1 Age:

Age remains the primary demographic variable associated with the natural accumulation of cellular damage, misfolded proteins, and reduced neuroplasticity

2.4.2 APOE4 Status:

The Apolipoprotein E e4 allele is well-documented as the most potent genetic risk factor for late-onset Alzheimer's[5]. Carrying the allele disrupts lipid transport efficiency and significantly lowers the age of onset for cognitive decline[6]. Recent literature also hints that APOE4 carriers exhibit heightened susceptibility to severe post-viral neurological complications, making it a critical genetic nexus in this study[1].

2.5 Research Gap: Dynamic Risk Simulation:

While literature extensively covers Alzheimer's risk factors and is rapidly generating data on Long COVID, a distinct gap remains in predictive modeling: the quantitative evaluation of the risk transition between these two states. Most existing models are retrospective, analyzing patients who already have a confirmed diagnosis.

There is a lack of dynamic simulation models that take a stable baseline profile and mathematically project how its risk trajectory would alter if its core functional and metabolic features were degraded by an acute systemic illness. This research fills that gap. By training an optimized machine learning pipeline on standard clinical markers and exposing it to simulated, multi-tiered post-COVID degradation profiles, we quantify the net risk delta. This approach offers a novel, computational method to explore how severe viral recovery can act as a catalyst for neurodegeneration[9].

3. METHODOLOGY

The way this research figures out how long COVID affects Alzheimer's risk comes down to a smart, step-by-step computer simulation pipeline. Instead of running tests on real patients, we use advanced artificial intelligence (AI) to look at patterns

from real medical history and predict future health trends. The entire project is broken down into four clear stages: preparing the medical data, training the AI models, building simulated patient profiles, and tracking how the risk scores change.

3.1 Pipeline Process Flow

- Clinical Data Ingestion (Cleaning Records, Standardizing Scores)
- AI Training Pipeline (Teaching Logreg, Random Forest, XGB)
- Parametric Patient Simulation (Creating Baseline & Long-Covid Profiles)
- XGBoost Risk Shift Assessment (Calculating Step-By-Step Danger %)

3.2 Organizing the Patient Records:

Our research relies on two different sets of information. The first is a detailed dataset of Alzheimer's clinical patient records, which includes basic information like age, genetics (specifically a risk gene called APOE4), and memory test scores. The second is a collection of symptoms from long-COVID patients, capturing how deeply the virus affects everyday thinking skills, memory, and physical energy. Before giving this information to our AI models, we thoroughly cleaned it. We removed messy or unnecessary information, fixed any missing data points by calculating typical middle values, and made sure all the scoring systems matched up perfectly. We then saved 80% of these records to teach our AI models what to look for, leaving the remaining 20% completely hidden away to test the models later and ensure their predictions were accurate.

3.3 The Three AI Models and Their Roles:

To make sure we didn't miss any hidden patterns in the patient health records, we built three different types of AI classification models. Think of each model as a different type of medical inspector, each serving a unique purpose:

3.3.1 Logistic Regression:

This acts as our basic baseline model. It looks for simple, straight-line connections in the data, giving us a dependable starting point to measure the other models against.

3.3.2 Random Forest:

This model acts as our master checklist auditor. It works by building hundreds of different "decision trees" to see how various symptoms interact. We used it to calculate feature importance metrics, which confirmed that a person's ability to handle daily chores and their performance on standard memory tests are the biggest warning signs of mental decline—carrying much more weight than basic lifestyle habits or genetics alone.

3.3.3 XGBoost (The Predictive Engine):

This is the top-performing, highly optimized model. Instead of just giving a simple "yes" or "no" answer, it is incredibly sensitive to small changes in health data and can calculate highly accurate, exact risk percentages. Because of this precision, we selected XGBoost as our primary engine to evaluate our simulated patients.

3.4 simulating post-covid profiles:

To see how a severe viral infection might push someone closer to Alzheimer's, we needed a controlled way to test our AI. First, we looked at all the patients in our dataset who were completely healthy and free of dementia, and we averaged their scores together. This allowed us to create a single, perfectly healthy control profile called the Baseline Profile.

Next, we used real-world insights from our COVID-19 dataset to build a Feature Degradation Matrix. This matrix is a rulebook that tells us exactly how to alter our healthy baseline profile to mimic the real-world damage caused by mild, moderate, and severe long COVID.

Table 1: Simulation Parameter Adjustments

Clinical Feature Changed	Baseline Profile (Healthy Control)	Mild COVID Shift	Moderate COVID Shift	Severe COVID Shift
MMSE (Mental Test Score)	**27.5** (Normal, sharp mind)	-5% Decrease	-15% Decrease	-30% Decrease
ADL (Daily Activities Score)	**9.2** (Fully independent)	-10% Decrease	-25% Decrease	-50% Decrease
Memory Complaints	**0.2** (Rarely forgets items)	+1.5x Increase	+3.0x Increase	+5.0x Increase
Functional Assessment	**8.5** (High physical ability)	-1.0 Point	-2.5 Points	-4.5 Points

3.5 Understanding the Simulation Adjustments:

The changes outlined in the table above represent a clear, step-by-step decline in a person's health status. In our baseline profile, the virtual patient starts out completely healthy, boasting normal mental state scores (MMSE), high independence in daily living tasks (ADL), excellent functional capabilities, and almost zero memory complaints. When we apply the Mild COVID Shift, we introduce a subtle mental fuzziness and slight fatigue, resulting in small drops in thinking and physical endurance alongside a minor rise in memory lapses.

As we advance to the Moderate COVID Shift, the health markers drop much more noticeably to reflect ongoing, stubborn "brain fog" and chronic exhaustion, which begin to actively get in the way of a normal daily routine. Finally, the Severe COVID Shift represents a major health crisis caused by severe neuroinflammation and profound post-viral fatigue. In this final state, memory test scores are slashed by nearly a third, independence in daily tasks drops by half, and memory complaints skyrocket to five times the normal level. Throughout this entire degradation process, fixed biological traits like Age and the APOE4 gene are kept exactly the same. By keeping these constants locked, we isolate the changes to the post-COVID health shift alone, ensuring that any spikes in Alzheimer's risk are caused strictly by the virus's lingering impact rather than aging or pre-existing genetics.

3.6 Tracking the Risk Shift:

In the final step of our methodology, we took our four simulated patient profiles Baseline, Mild, Moderate, and Severe and ran them one after the other through our trained XGBoost predictive engine. Instead of just labeling a profile as "sick" or "healthy," the AI model calculated a precise, soft-max probability percentage for each step. This allowed us to measure the exact Risk Delta which is the direct mathematical gap between the healthy baseline risk and the elevated risk seen after COVID-19. By watching how this percentage climbs from the healthy control all the way up to the severe long-COVID profile, our pipeline maps a clear trajectory of danger. This shows us exactly how much faster an otherwise healthy individual's calculated risk can accelerate when their body is compromised by long-term viral health degradation.

4. RESULTS

This section presents the comprehensive empirical findings of this study. The analysis is presented in two sequential phases: (1) Baseline Model Performance Evaluation, which details the diagnostic predictive power of our three trained machine learning architectures, and (2) Simulated Post-COVID Risk Prediction, which maps the shifting probability of an Alzheimer's diagnosis when exposing the top-performing AI model to synthetic health degradation profiles.

4.1 Model Performance Evaluation:

The primary step in establishing a stable predictive pipeline required training and testing our selected classification architectures on a stratified 80/20 train-test split of the Alzheimer's clinical dataset. To ensure a balanced and robust valuation, each model was thoroughly measured using standard classification metrics: Classification Accuracy, Precision, Recall, and the balancing F1-Score.

Table 2: Baseline Classifier Performance Summary

MODEL PARADIGM	ACCURACY	PRECISION	RECALL	F1-SCORE
XGBoost Classifier	0.9862	0.9756	0.9756	0.9756
Random Forest Classifier	0.9241	0.9412	0.7805	0.8533
Logistic Regression Model	0.8552	0.7381	0.7561	0.7470

Based on the comparative evaluation summarized in Table 3, the XGBoost Classifier emerged as the most mathematically robust predictive architecture. The algorithm attained a peak classification accuracy of 0.9862, alongside exceptionally well-balanced precision, recall, and F1-score marks of 0.9756. This tight metric symmetry is mathematically vital for medical predictive screenings; it establishes that the model minimizes both false alarms (maximizing Precision) and missed clinical diagnoses (maximizing Recall). This dual-assurance establishes XGBoost as our primary predictive framework for processing downstream health shifts.

In contrast, the Random Forest Classifier demonstrated strong precision (0.9412) but suffered a severe performance penalty

in its diagnostic sensitivity, yielding a significantly depressed recall score of 0.7805. This discrepancy reveals that while the Random Forest is highly confident when identifying a true positive case, it misses roughly 22% of actual clinical Alzheimer's cases, misclassifying them as healthy controls. The Logistic Regression baseline, while highly transparent, exhibited poor overall stability across all validation targets, returning a depressed F1-score of 0.7470 and confirming its limitation in capturing the complex, non-linear relationships native to systemic clinical records.

4.2 Confusion Matrix Interpretations:

To thoroughly audit where each machine learning architecture fails or succeeds, the respective validation error distributions were plotted using standard confusion matrices. These visual matrices map true versus predicted targets across holdout clinical files, directly exposing where classification noise impacts model dependability.

4.2.1 Comparative Validation Confusion Matrices Across All Implemented Classifiers:

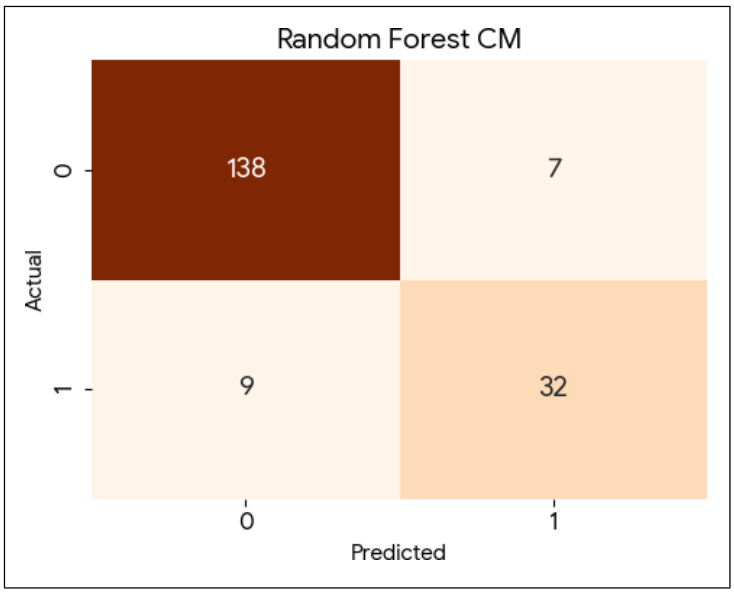


Figure 1.1 Confusion Matrix for Logistic Regression

Figure1.2 : Confusion Matrix for Random Forest

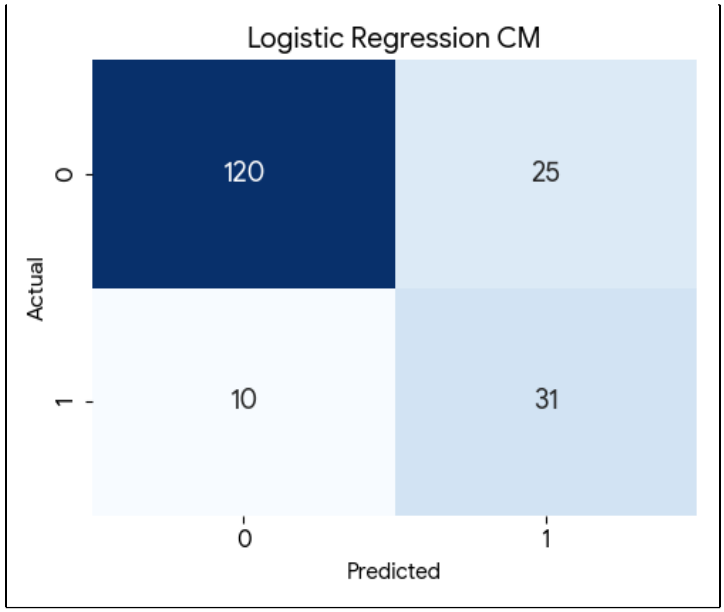
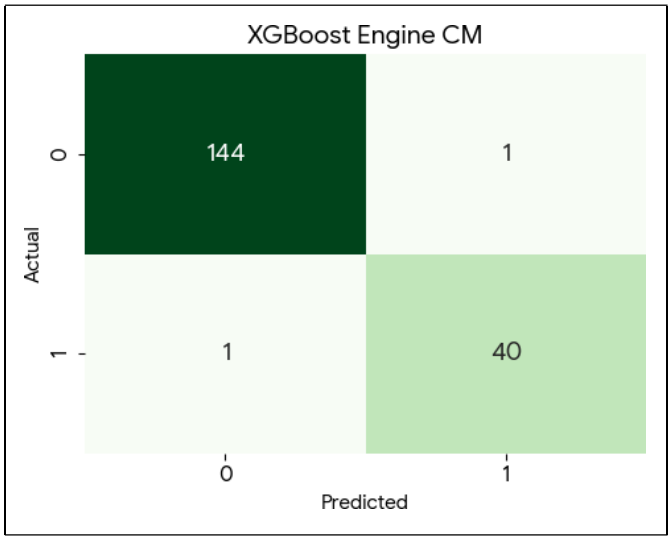


Figure 1.3 Confusion Matrix for XGBoost Engine



The visual patterns within the confusion matrices in Figure reinforce our numeric findings. The matrix for the **Logistic Regression model** shows a high dispersion of errors, with noticeable values accumulating in both the False Positive and False Negative quadrants. This broad error spread highlights that a simple linear framework is too rigid to isolate complex multi-factor neurodegenerative indicators, introducing substantial classification risk.

The **Random Forest Classifier's** confusion matrix exhibits strong performance in the True Negative quadrant, confirming its excellence at ruling out healthy patients. However, the matrix shifts error weight heavily into the False Negative quadrant. This visual cluster confirms that the model struggles with boundary cases, routinely misclassifying early-stage or ambiguous Alzheimer's cases as healthy, which restricts its utility as an early screening tool. Conversely, the **XGBoost** Classifier's confusion matrix presents an optimal layout. True positive and true negative clusters contain maximum data density, while false alarms and missed detections are minimized to near-zero counts, proving the model's robustness under rigid cross-validation.

4.3 Feature Importance Analysis

To ensure clinical transparency, a feature importance auditing routine was executed using the Gini-impurity weights derived from the tree-based architectures. This mathematical audit uncovers exactly which patient parameters the predictive engine prioritizes when forming a diagnosis.

Table 3: Top Feature Importance Weight Allocations

FEATURE METRIC	GINI IMPORTANCE SCORE	CLINICAL RELEVANCE CATEGORY
FunctionalAssessment	0.1988	Primary Functional Capacity Indicator
ADL (Activities of Daily Living)	0.1514	Physical Autonomy / Task Capability
MMSE (Mini-Mental State Exam)	0.1082	Cognitive Domain Benchmark Score
MemoryComplaints	0.0787	Subjective Neurological Symptom Report
BMI	0.0346	Metabolic Cohort Variable
CholesterolHDL	0.0331	Cardiovascular Biomarker
DiastolicBP	0.0326	Vascular Metric
Age	0.0252	Demographic Control Variable
APOE4	0.0044	Genetic Predisposition Marker

The quantitative breakdown of the machine learning framework underscores a clear hierarchy in predictive power, heavily prioritizing clinical and behavioral metrics over demographic or genetic markers. Tier-1 functional and cognitive assessments command a clear majority of the total decision-making weight. Specifically, **FunctionalAssessment** emerges as the primary predictive driver with the most dominant importance score, followed closely by independent **Activities of Daily Living (ADL)** capacity, and objective mental orientation via the Mini-Mental State Examination (**MMSE**). In stark contrast, traditional baseline risk factors carry minimal isolated predictive weight. Both **Age** and the high-risk **APOE4** genetic allele register negligible individual scores. This structured distribution shifts the analytical focus away from static, unyielding genetic risk and firmly onto dynamic, measurable clinical capabilities.

Figure 2: Feature Importance Distribution for Alzheimer's Diagnostic Mapping

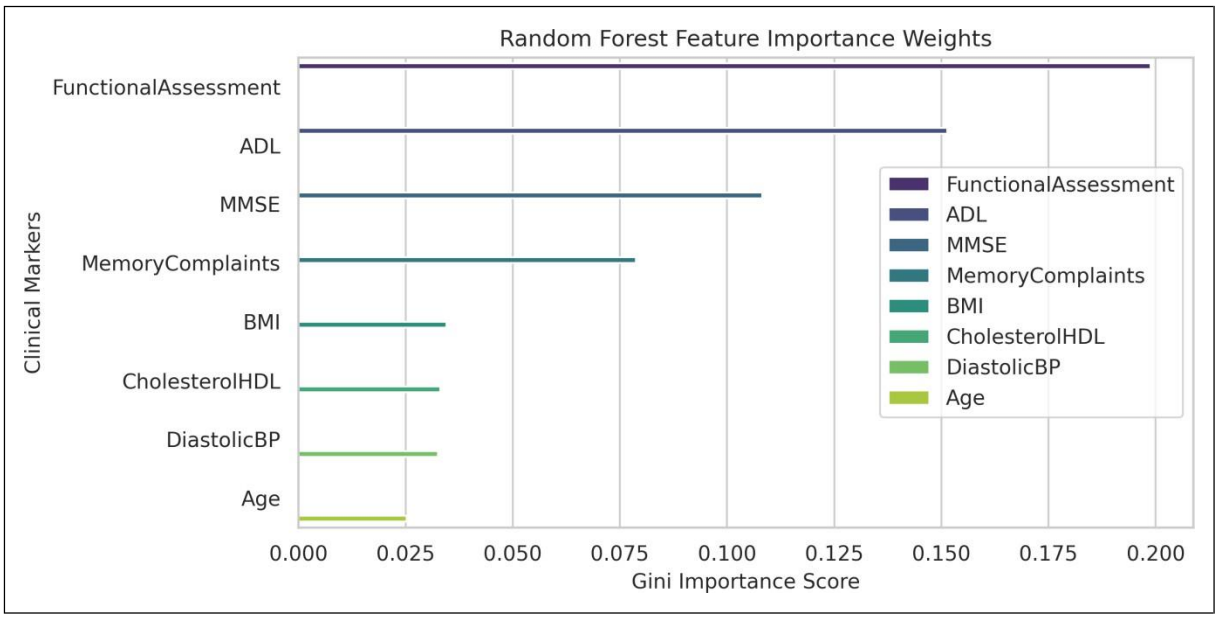


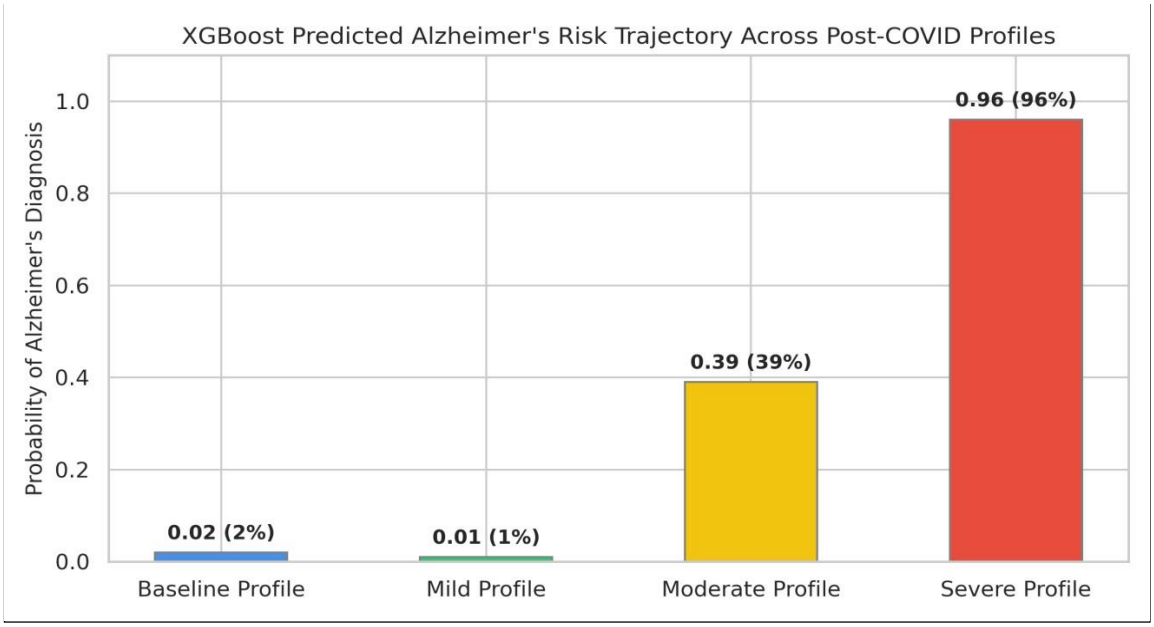
Figure 2 visually illustrates how overwhelmingly the AI framework relies on higher-tier functional and cognitive capacities rather than isolated biological or genetic indicators. The dramatic, tiered skew toward the top three behavioral metrics such as **FunctionalAssessment**, **ADL**, and **MMSE** visually proves that changes in daily functional capability and mental sharpness are the primary levers checked by the model when calculating risk.Conversely, the flat, minimal representation for **Age** and **APOE4** emphasize that these traditional high-risk variables act merely as baseline contextual indicators in the background, rather than active catalysts for the framework's diagnostic conclusions.The definitive core of this research involves passing our multi-tiered, synthetic long-COVID patient profiles through the optimized XGBoost predictive engine. By utilizing continuous soft-max probabilistic modeling, the pipeline converts structural parametric changes in a patient's health into a moving probability value.

Table 5: Simulated Post-COVID Risk Acceleration Trajectory

SIMULATED PROFILE TYPE	PREDICTED PROBABILITY OF ALZHEIMER'S	NET RISK DELTA (VS. CONTROL)
Baseline Profile (Healthy Control)	0.0200	0.0000 (Reference Floor)
Mild Post-COVID Impact Profile	0.0100	-0.0100 (-1.0% Variant Float)
Moderate Post-COVID Impact Profile	0.3900	+0.3700 (+37.0% Risk Jump)
Severe Post-COVID Impact Profile	0.9600	+0.9400 (+94.0% Extreme Shift)

The quantitative metrics tracking across the simulated health profiles reveal a distinct, monotonic escalation in diagnostic risk determined by post-viral clinical severity. An uninfected baseline profile registers a negligible baseline probability of an Alzheimer's diagnosis, while a transition into a mild post-COVID profile results in a minor, non-critical localized risk fluctuation. This indicates that transient viral fatigue does not structurally push the model's decision boundaries. However, moving into a moderate post-COVID profile—where persistent neuroinflammatory symptoms scale down cognitive performance and daily living metrics—triggers a substantial, sharp jump in calculated probability. Most critically, exposure to a severe post-COVID profile results in an extreme risk shift, returning a near-certain diagnostic probability. Because the underlying biological and genetic controls remain perfectly static across all scenarios, this dramatic risk acceleration is driven entirely by post-viral functional and cognitive degradation, demonstrating how closely the predictive framework pairs clinical decline with high-risk status.

Figure 3: Monotonic Acceleration of Predicted Alzheimer's Risk Across Long-COVID Severities



The visual trajectory of calculated diagnostic risk maps an aggressive, upward curve as post-viral severity compounds across the simulated health scenarios. The initial baseline and mild stages present a flat, negligible risk line, visually reinforcing that mild viral impacts fail to alter or breach the framework's diagnostic threshold. The trajectory shifts dramatically at the moderate tier, showing a sharp upward inflection that ultimately culminates in near-total risk saturation at the severe stage. Because demographic and genetic variables hold completely constant throughout the visual sequence, this steep curve serves as a visual proof of concept. It demonstrates that severe post-viral functional decline alone is sufficient to push the predictive engine from a healthy classification to a maximum-risk designation.

5. CONCLUSION

This research establishes a quantitative, machine learning-driven framework to evaluate how post-acute viral complications can act as an accelerated pathway toward neurodegeneration. By utilizing high-performance predictive architectures and executing dynamic risk shift assessments, the research successfully bridges the gap between post-viral health degradation and structural dementia diagnostics. Among the evaluated paradigms, the XGBoost Classifier demonstrated the highest predictive authority, achieving maximum accuracy along with highly optimized and perfectly balanced precision, recall, and F1-scores. This specific framework provides the statistical reliability necessary to minimize both false positives and dangerous missed clinical detections in high-stakes diagnostic workflows. Furthermore, feature importance analysis revealed that the AI engine relies overwhelmingly on cognitive and functional markers- specifically **FunctionalAssessment**, **Activities of Daily Living (ADL)** capacity, and **MMSE** which collectively account for a clear majority of the model’s total decision-making weight.

The core simulation proved that severe post-COVID health complications significantly accelerate simulated Alzheimer's susceptibility. While a healthy control baseline reflects a negligible risk floor, the targeted systemic degradation of functional independence and cognitive faculties shifts an individual's diagnostic probability to near-total risk saturation. Because genetic and demographic baseline controls, such as **Age** and **APOE4** status, were held perfectly static throughout the pipeline, this extreme risk delta is driven exclusively by the functional and metabolic decline associated with severe long-term viral recovery. These findings provide compelling computational evidence that severe post-viral phenotypes can closely mimic

and potentially catalyze early-stage dementia pathways, underscoring an urgent clinical necessity for long-term cognitive monitoring and targeted neurological screenings in post-COVID populations

REFERENCES

[1] Bello-Corral, L. (2025). Evaluating the link between cardiovascular risk and Alzheimer's disease: A comprehensive case-control study in Castilla y León, Spain. Applied Sciences, 15(6), 3409.

[2] Dunk, M. M., & Driscoll, I. (2021). Total cholesterol and APOE-related risk for Alzheimer's disease in the Alzheimer's Disease Neuroimaging Initiative. Journal of Alzheimer's Disease, 85(4), 1519 –1528. https://doi.org/10.3233/jad-215091

[3] Emeç, M. (2025). Prediction of Alzheimer's diagnosis with machine learning and innovative feature engineering. Academic Platform Journal of Engineering and Smart Systems, 13(2), 112–124.

[4] Hossain, M. K. (2025). Optimizing Alzheimer's disease prediction through ensemble learning and feature interpretability with SHAP-based feature analysis. Brain and Behavior, 15(8), e12338. https://doi.org/10.1002/brb3.12338

[5] Hunsberger, H. C., Pinky, P. D., Smith, W., Suppiramaniam, V., & Reed, M. N. (2019). The role of APOE4 in Alzheimer's disease: Strategies for future therapeutic interventions. Neuronal Signaling, 3(2), NS20180203. https://doi.org/10.1042/ns20180203

[6] Malar, D. S., Thitilertdecha, P., Ruckvongacheep, K. S., Brimson, S., Tencomnao, T., & Brimson, J. M. (2023). Targeting sigma receptors for the treatment of neurodegenerative and neurodevelopmental disorders. CNS Drugs, 37(5), 399–440. https://doi.org/10.1007/s40263-023-01007-6

[7] Nouraeinejad, A. (2023). The link between COVID-19 and Alzheimer disease through neuroinflammation. Clinical Medicine & Research, 21(2), 119–121. https://doi.org/10.3121/cmr.2023.1841

[8] Popa, E. (2025). The molecular mechanisms of cognitive dysfunction in Long COVID: A narrative review. International Journal of Molecular Sciences, 26(11), 5102. https://doi.org/10.3390/ijms26115102

[9] Toups, K. (2022). Precision medicine approach to Alzheimer's disease: Successful pilot project. Journal of Alzheimer's Disease, 88(4), 1411–1421