Heinz 94-887: Alzheimer's Disease Progression Prediction using ML Models

Ziyi Chen **Jenny Ding** **Yanyi Qian** **Zhaoyu Qiao** **Jinchen Xie**
Heinz College of Information Systems and Public Policy
Carnegie Mellon University
Pittsburgh, PA, United States

{ziyichen,yanyid,yanyiq,qzhaoyu,jinchenx}@andrew.cmu.edu

Abstract

Alzheimer's Disease (AD) is a progressive disease that destroys memory and thinking skills of patients. In 2006, 26.6 million people were affected by AD globally, and the number is expected to quadruple by 2050. There is currently no cure, and existing medications can only slow down the disease progression. Existing literature has demonstrated Machine Learning models' ability to predict AD; however, how early prediction can be performed and how clinicians could use the prediction results remain unexplored. Our study revolves around building machine learning models to detect AD transition from mild cognitive loss (MCI), 6, 12, and 24 months in advance. Our work utilizes data from Alzheimer's Disease Neuroimaging Initiative (ADNI). Using random forest model, which showed the best performance, we were able to achieve 83% sensitivity rate in the near future and 96% in the distant future. Moreover, we developed an interactive tool for clinicians to visualize risks of transition in the near and distant future to better develop treatment plans.

1. Introduction

Affecting an estimated 5.5 million Americans, Alzheimer's Disease (AD) is a progressive disease and the most common cause of dementia. It has serious symptoms that interfere with daily life and worsen over time. AD is currently ranked as sixth sixth leading cause of death the United States (Institute on Aging (2019)). Considering its devastating effects and the fact that treatments can only slow down the progression but not revert the process, researchers have been seeking better solutions in diagnosing early AD amid the challenge that early forms of the disease has no determinant clinical manifestation.

Led by the increasing application of Machine Learning on disease diagnosis, researchers have made substantial progress in applying state-of-art models on Alzheimer's progression prediction. For example, many researches have focused on using EHR and claims data to predict Alzheimer's risk among population. The Korean National Health Insurance Service database is used by Park et al. (2020) to predict likelihood of AD among payees, requiring no new collection of data. Similarly, data gathered from the Indiana Network for Patient Care is used to predict AD risk (Jessica Kent (2020)), in which natural language processing is used to extract features from medical notes. Recent advances in the collection of Alzheimer's Disease data (Aisen et al. (2017)) have led to the creation of ADNI database, which aggregates study data and provides open access to researcher. Using ADNI data, Beltrán et al. (2020) surveyed ML models, including tree-based and SVM methods to predict conversion from mild cognitive impairement (MCI) to AD, and Gupta et al. (2019) used

0. **Our code can be accessed here: Github Repository_AlzheimersDisease**

truncated singular value decomposition technique to reduce feature dimensionality and help with feature selection. Deep learning methods have also been explored in progression prediction tasks. Lee et al. (2019) had used RNN to predict conversion from MCI to probable AD, using neuroimaging biomarkers and cognitive performance data contained in ADNI, achieved 81 percent accuracy. Nonetheless, the earliness of progression prediction has not been fully explored.

Although previous researches in AD modeling has shown high prediction accuracy, most of the studies focused on predicting stages instead of transition. Moreover, results are often not advanced enough for clinicians to change regimens or increase medication dosage to slow down progression. Furthermore, predicting if a patient will convert to AD two years in advance would have been very beneficial for family members to search for appropriate caregivers and nursing homes ahead of time. In this work, we aim to predict the risk of conversion from the prodromal stage MCI to AD in 6-, 12- and 24-month using patient data collected at baseline visit. Our prediction results could help clinicians let the patients prepare for potential conversion to dementia in the future. Not only is our work unique in the sense that we provide transition probabilities over the span of two years, we also developed a user-friendly user interface dedicated for physicians with high interpretability. Another of our goal is to provide a consolidated list of predictors that are highly effective in predicting transition status, thereby improving the cost efficiency of check ups for prodromal stage AD patients.

Section 2 of this report provides background on Alzheimer's Disease and existing models. Then, in Sections 3 and 4, we explain our models and implementation setup, followed by a discussion on model results in Section 5. Lastly, we will discuss our results, use cases and future directions in Section 6.

2. Background

AD is a neurodegenerative disease caused by misfolded proteins. The molecular basis lies in the accumulation of toxic senile amyloid beta plaques and neurofibrillary tangles caused by destablized microtubules, which resulted hyperphosphorylated tau as explained by Pospich and Raunser (2017). Brain imaging techniques can help researchers identify spread of abnormal amyloid and tau in patients' brains. Weller and Budson (2018) has characterized the three stages of AD. Patients who are in the pre-clinical phase might notice memory changes themselves; however, they are not detectable by doctors. In the MCI phase, cognitive changes become of concerns to individuals and family, with one more cognitive domains impaired significantly. In the last phase, cognitive impairment would become severe enough to interfere with everyday abilities. Currently, AD can be diagnosed by asking the patients and family members questions about the patients' overall ability to carry out daily activities, conducting cognitive tests, and performing brain scans and other standard medical tests to rule out other possible causes for symptoms. While genetic and molecular targeted therapies aimed at disease prevention remain to be top goals, current treatments of AD can only reduce symptoms and manage behavior problems. Morevoer, they may only be effective for a limited period of time. Caring for patients with AD is also costly, both financially and emotionally. There are often difficult decisions about coping with the patients' demand of intensive daily care, and whether the patients need to be placed in care facilities.

Designed for the detection and tracking of AD, the Alzheimer's Disease Neuroimaging Initiative (ADNI) study has been running for 17 years since 2004. The study recruits participants from 57 sites across the U.S. and Canada, 55 to 90 years of age, with or without memory impairment, and then followed them over time. The shared data contains a rich set of potential risk factors for AD by collecting five general types of data: clinical, genetic, MRI image, PET image and biospecimen. A diagnosis of CN, MCI or AD was given during each follow-up.

3. Methods

We applied logistic regression, decision tree, random forest and gradient boosting classifier to develop our classification task. Binary logistic regression transforms its output using the sigmoid function and return a probability value of two classes. Decision tree predicts the class by learning a series of decision rules. Random forest is an ensemble model of decision trees. Also, we used decision trees as weak learners in gradient boosting classifier. All these models do not require a linear relationship between the dependent and independent variables. In implementation, we utilized Tidymodels, Kuhn and Wickham (2020). To assess the generalization ability of our models, We utilized cross validation method to obtain more stable and accurate estimate by averaging the error across five folds.

Our code can be accessed here: **Github Repository_AlzheimersDisease**

4. Experimental Setup

Our pipeline involves five mains steps, data cleaning, exploratory analysis and cohort selection, data transformation, model building and lastly shiny implementation. The detailed breakdown of the steps are shown in a flowchart (Figure 1).

4.1 Data Extraction

From the ADNI databse, we were able to download 'ADNIMERGE.csv' which contained 2294 unique patients. Meaningless, non-baseline features as well as those had had over 40% missing values were deleted. Patients that only had baseline visits were removed as well as those that had missing diagnosis in all following visits. For patients that have the last one or few diagnosis missing, rows were removed from the end until the last non-missing diagnosis. Lastly, to fill in the missing diagnosis in middle visits, diagnosis labels are imputed if the previous and next non-missing labels are the same. Rows with different previous and next non-missing labels were removed. For missing baseline values, we used K-Nearest Neighbor to impute. For numerical features that also had non-numerical labels such as '>1700,' they were replaced with the numerical value, e.g. 1700. After all processing, we were left with 1159 patients and 41 features.

4.2 Cohort definition

The cohort are recruited participants of the ADNI study that spans from 2004 to 2019. Participants were from 57 sites across the U.S. and Canada. Only participants who were diagnosed as MCI at baseline and had diagnosis at 6-, 12-, and 24-month of follow-ups

(n=651) were retained in our studied cohort. The selected cohort had an average age of 72.7 years, with a range from 55 to 90 years, were mostly white (94.00%), married (78.49%) and males (59.14%). Regarding the APOE4 gene that heightens the risk of Alzheimer's, 321 (49.30%) don't have APOE4, 254 (39.02%) are carriers, and 76 (11.67%) have two copies.

For patients that have transitioned at 6-month, we removed them from the 12- and 24-month analysis, and similarly, for those that have transitioned at 12-month, they are removed from the 24-month analysis. Therefore, we had different cohorts at each time mark, and we ended up with n=651 for 6M (immediate), n=566 for 12M and n=471 for 24M early prediction. An illustration of cohort cohort breakdown is shown in Appendix A (Figure 2).

4.3 Feature Choices

There are four types of feature in the cleaned data: cognitive tests, biomarkers, imaging and demographics. Cognitive tests contain scores from Rey Auditory Verbal Learning Test (RAVLT), The Mini-Mental State Exam (MMSE), modified Preclinical Alzheimer Cognitive Composite (mPACC), Functional Activities Questionnaire Index(FAQ), Clinical Dementia Rating Scale (CDRS) and Alzheimer's Disease Assessment Scale (ADAS). Biomarkers include APOE4 allele and CSF levels of amyloid beta (ABETA), tau and phosphorylated tau (PTAU). Imaging data contained MRI volumetric data of different brain regions and various types of PET scan results.

We applied log transformation to features that have skewed distribution such as FAQ and PTAU (Figure 3). To scale all data into the same level, we applied min-max scaling. These conversions prevent models from being affected by the varying distributions and magnitudes of different features. Because we assumed that a natural ordering between categories may result in poor performance or unexpected results, we applied one-hot encoding to our categorical features, which added a new binary variable for each category.

After transformation, we compared the mean value of attributes between transition and non-transition group at 24M to validate that our features are capable of distinguishing MCI and AD. 27 out of 41 attributes showed significant differences between the two groups as shown in Figure 4.

4.4 Evaluation Criteria

We used sensitivity and specificity to evaluate our model performance. Sensitivity and specificity are two important measures in medical problems. In our context, sensitivity, the true positive rate, refers to the probability of being correct when the patient actually transitions. Specificity, the true negative rate, refers to the probability of being correct when the patient actually does not transition.

For disease warning, not predicting a patient might transition is more costly than predicting a patient to transition when they does not. Therefore, our first goal is to maintain high sensitivity and then to improve specificity.

5. Results

We mainly have two types of results, those from building our prediction model and those from predicting on new patients. These are described in the subsections below.

5.1 Model Result

Four algorithms were tried to predict the transition from MCI to AD during our analysis as described earlier: Random Forest, Decision Tree, Logistic Regression and Gradient Boosting. By comparing the ROC curves and the corresponding AUC values for all models shown in Figure 5, Random Forest method was picked as our best model, which had all AUC values higher than 0.80 at all prediction time points.

Based on Table 1, sensitivity of the Random Forest model for 6M, 12M and 24M predictions is 0.824, 0.863 and 0.958 respectively. This means that 82.4% of the patient we predicted to transition at 6M will in reality do so, and 96% of those at 24M will do so as well. specificity is 0.627, 0.654 and 0.518 respectively. False negative rate is also tuned to be minimal, at 0.177 for 6M and 0.0421 for 24M, which means that we have 17.7% chance at 6M and 4.21% chance at 24M of failing to predict the patient will transition if the patient in reality do transition. The confusion matrix is shown below:

		True Class-6M		True Class-12M		True Class-24M	
		Positive	Negative	Positive	Negative	Positive	Negative
Predicted Class	Positive	28	230	44	196	91	227
	Negative	6	387	7	370	4	244

Table 1: Confusion Matrix of Random Forest

To better understand the random forest model, the top 20 features that make the greatest contribution to the final prediction result were extracted from all predictors (Figure 6). For all three periods, the top 3 important predictors are cognitive tests. The modified Preclinical Alzheimer Cognitive Composite (mPACC) digits and trails tests are most effective in predicting transition at early time points (12M, 24M), whereas Alzheimer's Disease Assessment Scale (ADAS) and Functional Activities Questionnaire Index (FAQ) are most effective for predicting transition at immediate time points(6M, 12M). Imaging data is the next most important predictor type; however, its effectiveness drops as prediction time extends to 24M. In immediate prediction, MRI volumetric data of hippocampus and entorhinal cortex and PET imaging with F-fluorodeoxyglucose (FDG) are helpful in predicting transition. Imaging data's importance in early prediction is replaced by levels of amyloid-beta and phosphorylated tau in cerebral spinal fluid (CSF). Demographic data is not important in predicting transition as only age showed up in the top 20 important features list for all three time-points. This suggests that older patients have higher chances of progression.

To further evaluate our modeling results, we picked random forest, the best performing model, and drew calibration plot on its 6 month, 6-12 month and 12-24 month predictions (Figure 7). X-axis represents the average predicted probability of having MCI to AD tran-

sition and y-axis represents the percentage of patients having transitioned in that cohort. From initial visit to the six month period, our model managed to classify the transition risk perfectly with a threshold of 0.212. From the sixth month to twelfth month, the model still showed a strong classification power with a higher threshold, 0.641. Finally, when the model predicts the probability of transition from 12 to 24 month, the threshold moved to 0.677 with more predicting bias in comparison with the actual probability.

5.2 Sample New Patient Result

We included risk analysis for three new patients in our R Shiny application to visualize the transition risk change over time. These three patients were randomly sampled from those that did not have labels at 6, 12 and 24M. We were able to show the distribution of predicted risk under three time points and the corresponding risks of these three new patients (Figure 8). We also summarized the risk trajectory of new patients over time in a boxplot (Figure 9). Based on the results, most people fall under 0.05 transition risk and the distribution is right skewed. The overall risk increased when looking at the 6-12 month risk distribution and the 12-24 distribution. More patients have higher risks at later time points, which is logical since the risk of progression increases over time.

When looking at the individual level, patient 3 has the lowest transition probability from the initial visit to the end of 24 month. Patient 1 and 2 both progress to higher risks over time and Patient 1 is the most likely person to move into AD stage among all new patients.

6. Discussion

Our work has significant implications on clinical decision making. The clinical decision support tool we built provides one step further toward evidence-based medicine. Using the predicted probabilities generated by our models, physicians could better decide on whether if to use use more aggressive therapies on some patients over the others. Moreover, we also provide physicians with a list of the top effective predictors, thereby helping with the decision of choosing certain cognitive tests over the others.

6.1 Result Analysis

From analyzing sensitivity, false negative rate and AUC shown earlier, our forest model showed better performance for early than immediate prediction. However, this could be partially caused by our 24M cohort having the least amount of imbalanced labels (16.8%) whereas the 6M cohort had the highest imbalance (5.2%).

Because our model showed higher efficacy of cognitive tests in predicting transition, clinicians should place more emphasis on results from these tests in comparison with those generated from more costly MRI or PET scans as well as biomarkers from labs. Moreover, out of the myriad of possible cognitive tests, our random forest model showed that mPACC Digit and Trails tests, which are both individual and easy-to-administrate tests, are the most effective ones in early prediction. Therefore, when faced with time or cost constraints, these tests should be chosen over more complicated ones like CDRSB and RAVLT. ADAS and FAQ are more important for immediate prediction; however, ADAS contained 13 tests

and FAQ is a subjective in nature. The shrinkage of hippocampus region, which is essential for forming new memories, and entorhinal cortex, which provides sensory information to the hippocampus for memory and learning, are shown to be important predictors for AD transition prediction in addition to being seen as hallmarks of CN transition to MCI (Zhou et al. (2016), Mueller et al. (2010)). To summarize, our model provided additional insights from a modeling perspective to existing biomedical researches about transitions from MCI to AD.

6.2 Use Case

Our R Shiny application, AD Predictor, is our user interface that assists physician in AD transition forecasting of new patients. The goal of AD Predictor is to help physicians understand MCI patients' predicted risks of transition to AD within 6, 12 and 24 months after the initial visit.

Our tool first provides physicians with an understanding our database through data view and summary as well as how our model is calibrated to set credibility. Physicians can then upload their own data and see how their patient's risk variables compare with our database of over 600 patients. This feature helps physicians to visualize where a patient stands with other MCI patients without getting to the predicted probabilities. We then provide multiple charts and plots to let them have a clear view of the risk trajectory of their patients and how they compare with others. Physicians would be able to identify the high-risk patients by referring to the risk stratification and risk change over time dashboards and set up rigorous intervention treatments.

6.3 Limitations

The limitation of this project mainly lies in data collection and data possessing. We did not exploit the full potential of the ADNI database by excluding features like current medications, prior conditions from clinical notes, and raw brain images. Future work could extend the current model to employ multi-modal, unstructured data and further increase the model's predictive power.

In terms of data prepossessing, we did not consider the time series aspect and instead only used the baseline features collected during patients' first visit. For patients with missing diagnosis records at any time points, we simply imputed it with patients' last or most recent record and deleted those patients with missing diagnosis that are not imputable. For patients with few follow-ups, we were unable to identify if the patient were dead or dropped out. Thus, our data might be biased toward those who are alive or more active with visiting clinicians. All of these may have impacted our models.

6.4 Future Directions

Our future works could be done in three aspects, pipeline automation, full range transition risk prediction and use case expansion.

Currently our model lacks flexibility in feature input. A practical improvement is to allow missing inputs from clinicians. For example, clinician may only conduct one cognitive

test for the patient, and the tool should still be able to leverage only this feature and provide risk estimates. An automated pipeline including model adaptations is required in this case.

The tool could also be expanded to prediction the full range of transition status from CN to MCI in additional MCI to AD. This feature will expand the affected population and allow physicians to visualize the progression of each Alzheimer Disease patient from the early onset.

Lastly, another use case of this product is to evaluate the effects of medications and suggest optimized intervention regimens to patients, provided that we can access and aggregate medication data into our platform. By tracking the risk trajectory among patient with similar starting conditions but different treatments and outcomes, we would be able to compare treatment effects and provide personalized regimen for each patient.

7. Conclusion

Because early intervention is a treatment goal for AD patients, we built machine learning models to predict transition probabilities in the near and distant future. By uploading feature values gathered at a typical first visit to neurologists or geriatricians to our Shiny tool, we can provide physicians with results that are highly sensitive. The value of adding machine learning models in our task is to integrate clinician expertise and patient experience with the best available scientific information, which together can guide the decision making process for each patient. Because our model leverages data from hundreds of previous patients and considers a wide range of attributes, it provides an unbiased analysis of incoming patient. We hope that automated tools like our AD Predictor could be deployed in hospitals, thereby improving patient outcomes through evidence based medicine.

8. Acknowledgement

We are very thankful for the guidance that Professor Jeremy Weiss had given us over the course of this project. We also would like to express our gratitude toward the researchers who have continued to contribute to and maintain the ADNI database.

References

Paul S. Aisen, Jeffrey Cummings, Clifford R. Jack, John C. Morris, Reisa Sperling, Lutz Frölich, Roy W. Jones, Sherie A. Dowsett, Brandy R. Matthews, Joel Raskin, Philip Scheltens, and Bruno Dubois. On the path to 2025: understanding the alzheimer's disease continuum. (1):60, 2017. doi: 10.1186/s13195-017-0283-5.

Juan Felipe Beltrán, Brandon Malik Wahba, Nicole Hose, Dennis Shasha, Richard P. Kline, and For the Alzheimer's Disease Neuroimaging Initiative. Inexpensive, non-invasive biomarkers predict alzheimer transition using machine learning analysis of the alzheimer's disease neuroimaging (adni) database. *PLOS ONE*, 15(7):1–26, 07 2020. doi: 10.1371/journal.pone.0235663. URL https://doi.org/10.1371/journal.pone.0235663.

Yubraj Gupta, Ramesh Kumar Lama, and Goo-Rak Kwon. Prediction and classification of alzheimer's disease based on combined features from apolipoprotein-e genotype, cerebrospinal fluid, mr, and fdg-pet imaging biomarkers. *Frontiers in Computational Neuroscience*, 13:72, 2019. ISSN 1662-5188. doi: 10.3389/fncom.2019.00072. URL https://www.frontiersin.org/article/10.3389/fncom.2019.00072.

Jessica Kent. Machine learning uses EHR data to predict alzheimer's risk, 2020. URL https://healthitanalytics.com/news/machine-learning-uses-ehr-data-to-predict-alzheimers-risk.

Max Kuhn and Hadley Wickham. *Tidymodels: a collection of packages for modeling and machine learning using tidyverse principles.*, 2020. URL https://www.tidymodels.org.

Garam Lee, Kwangsik Nho, Byungkon Kang, Kyung Ah Sohn, and Dokyoon Kim. Predicting alzheimer's disease progression using multi-modal deep learning approach. *Scientific Reports*, 9(1), 12 2019. doi: 10.1038/s41598-018-37769-z.

Susanne G Mueller, Norbert Schuff, Kristine Yaffe, Catherine Madison, Bruce Miller, and Michael W Weiner. Hippocampal atrophy patterns in mild cognitive impairment and alzheimer's disease. *Human brain mapping*, 31(9):1339–1347, 2010.

National Institute on Aging. Alzheimer's disease fact sheet, 2019. URL http://www.nia.nih.gov/health/alzheimers-disease-fact-sheet.

Ji Hwan Park, Han Eol Cho, Jong Hun Kim, Melanie M. Wall, Yaakov Stern, Hyunsun Lim, Shinjae Yoo, Hyoung Seop Kim, and Jiook Cha. Machine learning prediction of incidence of alzheimer's disease using large-scale administrative health data. 3(1):1–7, 2020. ISSN 2398-6352. URL https://www.nature.com/articles/s41746-020-0256-0.

Sabrina Pospich and Stefan Raunser. The molecular basis of alzheimer's plaques. *Science*, 358(6359):45–46, 2017. ISSN 0036-8075. doi: 10.1126/science.aap8002. URL https://science.sciencemag.org/content/358/6359/45.

Jason Weller and Andrew Budson. Current understanding of alzheimer's disease diagnosis and treatment. 7, 2018. ISSN 2046-1402. doi: 10.12688/f1000research.14506.1. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6073093/.

Mengxi Zhou, Feng Zhang, Li Zhao, Jin Qian, and Chunbo Dong. Entorhinal cortex: a good biomarker of mild cognitive impairment and mild alzheimer's disease. *Reviews in the Neurosciences*, 27(2):185–195, 2016.

Appendix A: Tables and Figures

		True Class-6M		True Class-12M		True Class-24M	
		Positive	Negative	Positive	Negative	Positive	Negative
Random Forest	Positive	28	230	44	196	91	227
	Negative	6	387	7	370	4	244
Decision Tree	Positive	25	501	40	308	72	282
	Negative	9	116	11	258	23	189
Logistic Regression	Positive	31	336	46	330	84	164
	Negative	3	281	5	236	11	307
Gradient Boosting	Positive	30	217	40	201	89	162
	Negative	4	400	11	365	6	309

Table 2: Model Comparison - Confusion Matrix

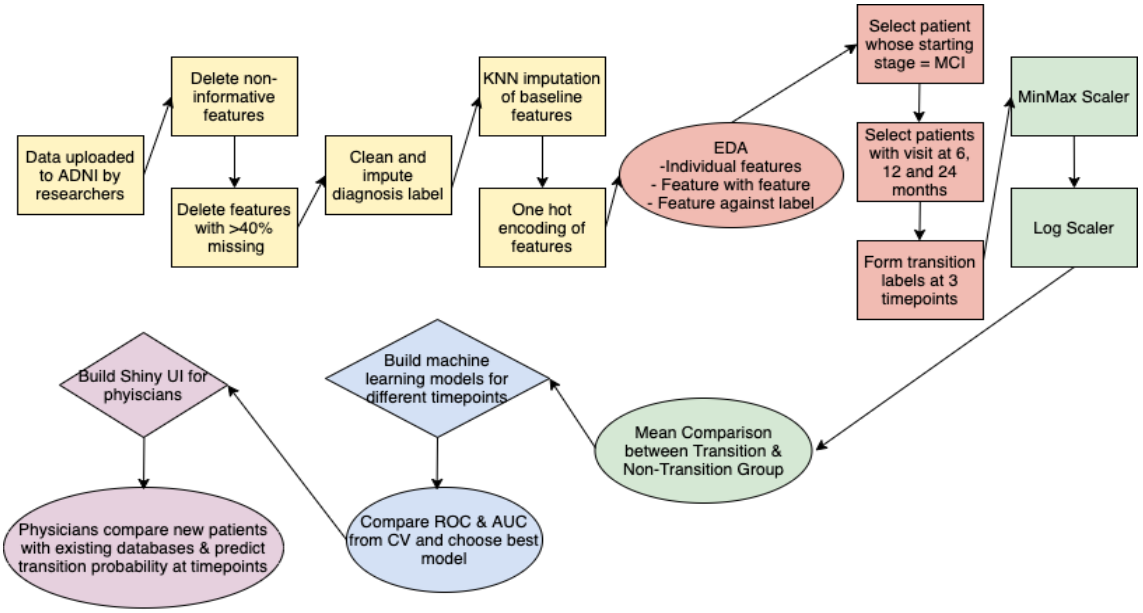
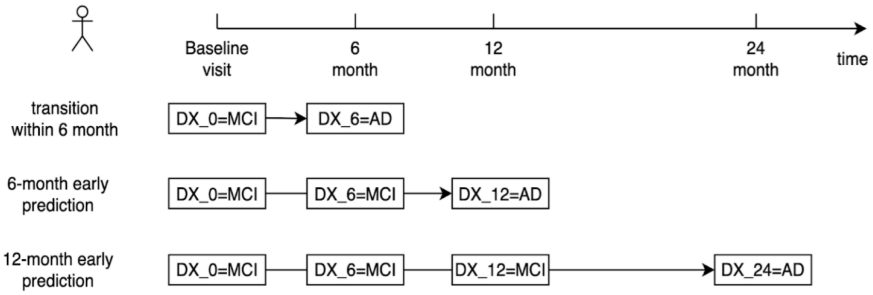


Figure 1: Pipeline Flowchart of 5 Steps



Label	Transition within 6 month (n=651)	6 month early prediction (n=617)	12 month early prediction (n=566)
True	34 (5.2%)	51 (8.3%)	95 (16.8%)
False	617 (94.8%)	566 (91.7%)	471 (83.2%)

Figure 2: Cohort Breakdown

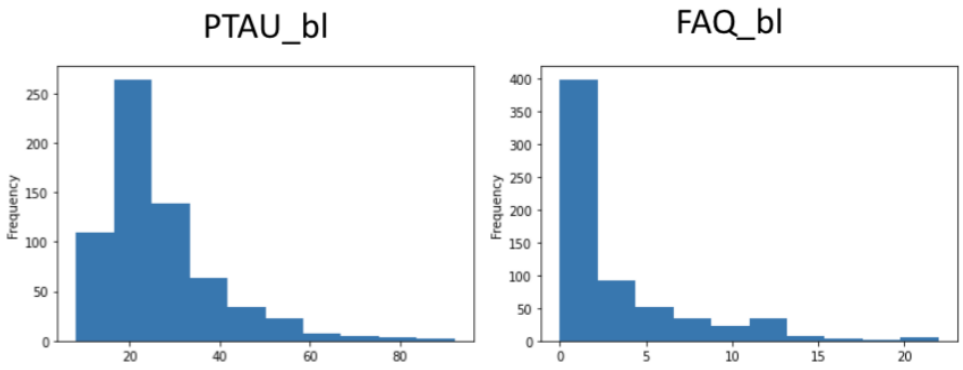


Figure 3: Skewed Distributions Requiring Transformation

attribute	p_value	no_transition_mean	transition_mean	category
Modified Preclinical Alzheimer Cognitive Composite with Trails test	5.48E-44	0.612	0.376	cognitive
Modified Preclinical Alzheimer Cognitive Composite with Digit test	2.54E-38	0.625	0.402	cognitive
Alzheimer's Disease Assessment Scale -13 Items	1.06E-32	0.384	0.592	cognitive
Alzheimer's Disease Assessment Scale -11 Items	9.60E-29	0.358	0.544	cognitive
Functional Assessment Questionaire	2.54E-28	0.092	0.267	cognitive
Task 4 of the Cognitive Subscale - Constructive Praxis	1.81E-27	0.482	0.719	cognitive
Rey Auditory Verbal Learning Test -Immediate Recall	2.14E-26	0.460	0.289	cognitive
Clinical Dementia Rating - Sum of Boxes	3.07E-22	0.156	0.302	cognitive
Rey Auditory Verbal Learning Test -Precent Forgetting	2.45E-20	0.611	0.829	cognitive
Trail Making Test - Time to Complete	7.79E-19	0.245	0.425	cognitive
MRI Volumetric Data of Hippocampus	8.19E-17	0.537	0.400	imaging
Mini-Mental State Exam	9.94E-17	0.720	0.535	cognitive
Rey Auditory Verbal Learning Test - Rate of Learning	3.35E-14	0.470	0.345	cognitive
Beta-amyloid plaques	**6.51E-14**	**0.532**	**0.349**	**biomarker**
MRI Volumetric Data of Entorhinal	4.58E-13	0.491	0.379	imaging
MRI Volumetric Data of Mid-Temporal Lobe	6.00E-13	0.487	0.384	imaging
MRI Volumetric Data of Fusiform Gyrus	1.69E-10	0.481	0.385	imaging
Number of APOE4 alleles	1.87E-10	0.261	0.455	biomarker
Hyperphosphorylation of tau	1.01E-07	0.204	0.279	biomarker
Neurofibrillary tau tangles	3.83E-07	0.220	0.290	biomarker
MRI Volumetric Data of Whole Brain	3.22E-06	0.418	0.351	imaging
MRI Volumetric Data of CSF Ventricles	4.43E-05	0.227	0.284	imaging
Fluorodeoxyglucose (FDG)-positron emission tomography (PET)	4.52E-05	0.504	0.442	imaging
Rey Auditory Verbal Learning Test - Delayed Recall Score - Immediate Recall Score	0.0038	0.434	0.475	cognitive
Divorced	0.0289	0.104	0.048	demographic
Age	0.0367	0.476	0.514	demographic
Married	0.0461	0.764	0.838	demographic

Figure 4: Mean Comparison between Transition and Non-Transition Group

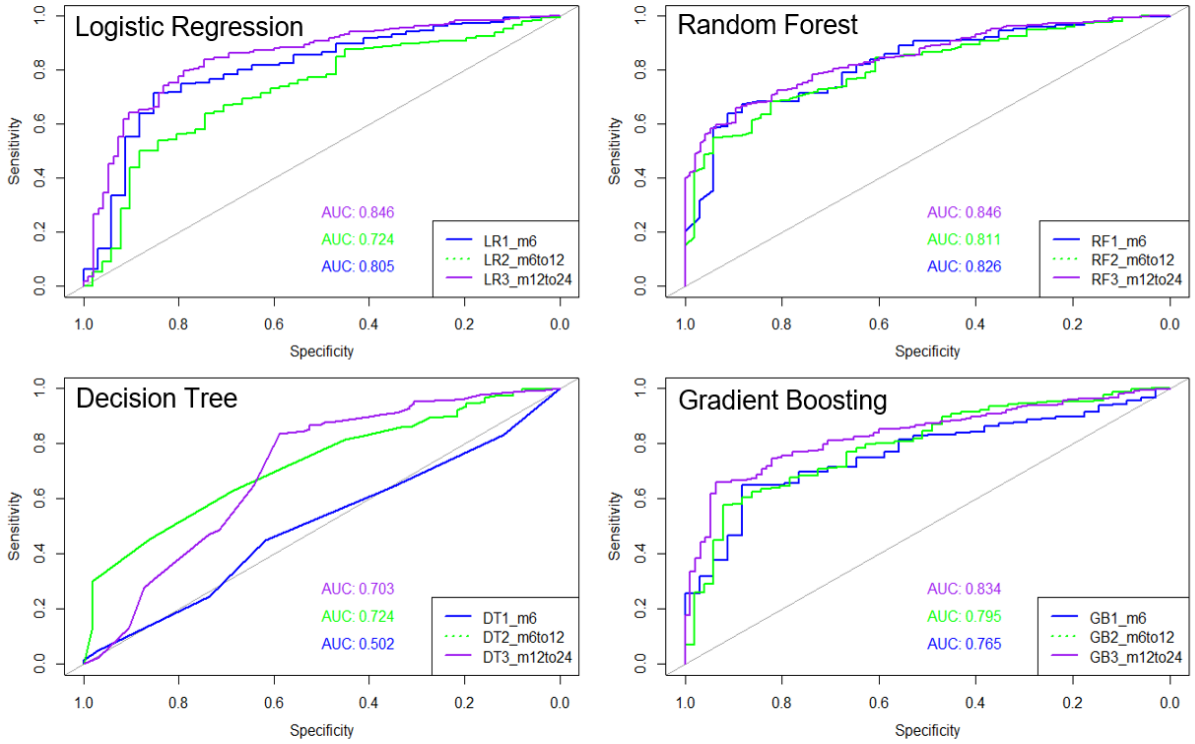


Figure 5: Model Comparison: ROC curve, CV = 5

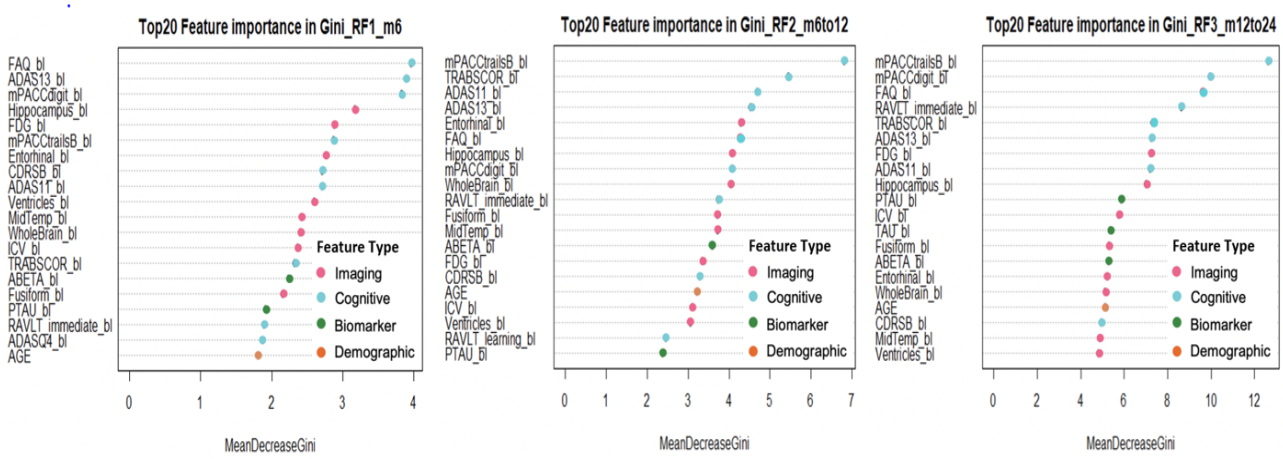


Figure 6: Top20 Feature Importance from Random Forest

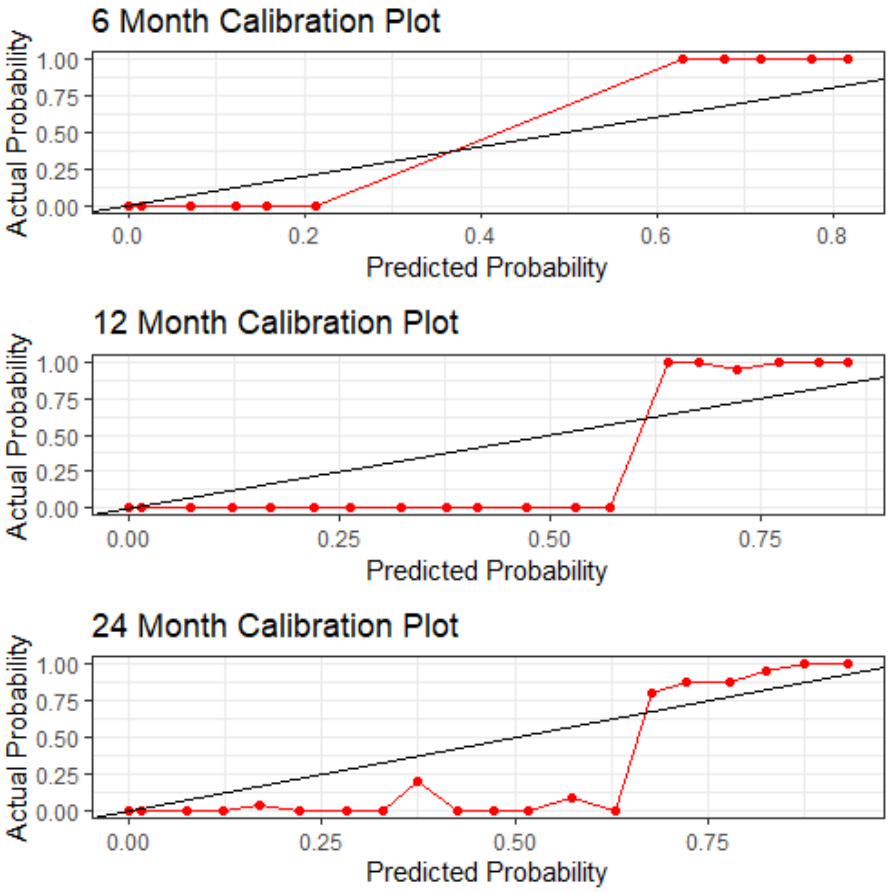


Figure 7: Model Comparison: Calibration Plot

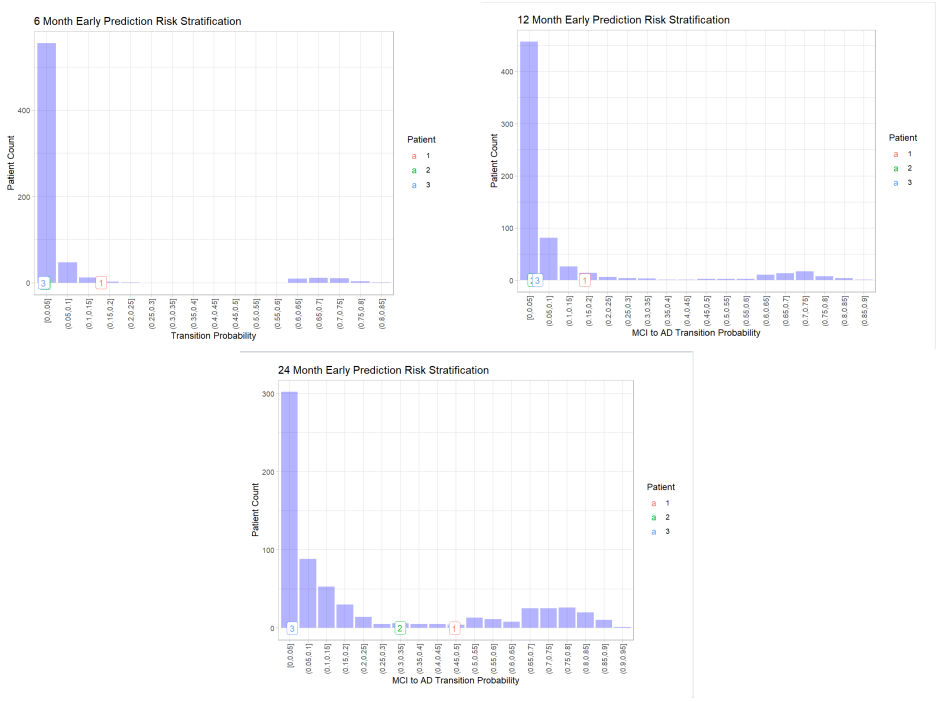


Figure 8: Risk Stratification in Each Period

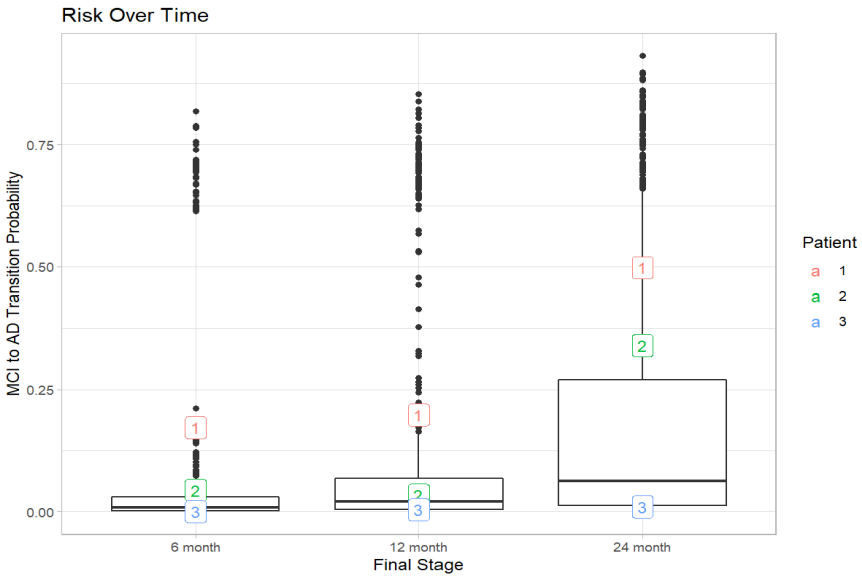


Figure 9: Three New Patients' Risk Change over Time