Summer 8-8-2017

Predicting Alzheimer Disease Status Using High-Dimensional MRI Data Based on LASSO Constrained Generalized Linear Models

Zainab Salah

ABSTRACT

PREDICTING ALZHEIMER DISEASE STATUS USING HIGH-DIMENSIONAL MRI DATA BASED ON LASSO CONSTRAINED GENERALIZED LINEAR MODELS

BY

ZAINAB SALAH

MAY 22, 2017

INTRODUCTION: Alzheimer's disease is an irreversible brain disorder characterized by distortion of memory and other mental functions. Although, several psychometric tests are available for diagnosis of Alzheimer's, there is a great concern about the validity of these tests at recognizing the early onset of the disease. Currently, brain magnetic resonance imaging is not commonly utilized in the diagnosis of Alzheimer's, because researchers are still puzzled by the association of brain regions with the disease status and its progress. Moreover, MRI data tend to be of high dimensional nature requiring advanced statistical methods to accurately analyze them. In the past decade, the application of *Least Absolute Shrinkage and Selection Operator* (LASSO) has become increasingly popular in the analysis of high dimensional data. With LASSO, only a small number of the regression coefficients are believed to have a non-zero value, and therefore allowed to enter the model; other coefficients are while others are shrunk to zero.

AIM: Determine the non-zero regression coefficients in models predicting patients' classification (Normal, mild cognitive impairment (MCI), or Alzheimer's) using both non-ordinal and ordinal LASSO.

METHODS: Pre-processed high dimensional MRI data of the Alzheimer's Disease Neuroimaging Initiative was analyzed. Predictors of the following model were differentiated: Alzheimer's vs. normal, Alzheimer's vs. normal and MCI, Alzheimer's and MCI vs. Normal. Cross-validation followed by ordinal LASSO was executed on these same sets of models.

RESULTS: Results were inconclusive. Two brain regions, frontal lobe and putamen, appeared more frequently in the models than any other region. Non-ordinal multinomial models performed better than ordinal multinomial models with higher accuracy, sensitivity, and specificity rates. It was determined that majority of the models were best suited to predict MCI status than the other two statues.

DISCUSSION: In future research, the other stages of the disease, different statistical analysis methods, such as elastic net, and larger samples sizes should be explored when using brain MRI for Alzheimer's disease classification.

PREDICTING ALZHEIMER DISEASE STATUS USING HIGH-DIMENSIONAL MRI DATA BASED ON LASSO CONSTRAINED GENERALIZED LINEAR MODELS

by

ZAINAB SALAH

B.S., GEORGIA STATE UNIVERSITY

A Thesis Submitted to the Graduate Faculty

of Georgia State University in Partial Fulfillment

of the

Requirements for the Degree

MASTER OF PUBLIC HEALTH

ATLANTA, GEORGIA

30303

APPROVAL PAGE

PREDICTING ALZHEIMER DISEASE STATUS USING HIGH-DIMENSIONAL MRI DATA BASED ON LASSO CONSTRAINED GENERALIZED LINEAR MODELS

BY

ZAINAB SALAH

APPROVED:

DR. RUIYAN LUO

SIGNATURE:

DR. XIN QI

SIGNATURE:

DATE OF DEFENSE:

MAY 22, 2017

DEDICATION

This paper is dedicated to my son, Jafar, the apple of my eye, and to my mother, to whom I owe all of my success to.

ACKNOWLEDGMENTS

First and foremost, I would like to express my sincere gratitude to my advisor, Dr. Ruiyan Luo, who guided me with indispensable instructions and valuable suggestions throughout my research period. This thesis would not have been possible without her assistance and advice. Secondly, I would like to thank Dr. Xin Qi, for his much appreciated help and guidance in this process.

Additionally, I would like to thank Dr. Matt Hayat for instilling the love of *biostatistics* in me, and for offering me the graduate teaching assistant position, through which I gained a wealth of knowledge both academically and professionally.

I would like to express my appreciation to Dr. Dora Il'yasova for the constant reminder that I can accomplish anything I set my mind to; thank you for believing in me.

I would like to acknowledge all the dedicated faculty and staff of Georgia State University School of Public Health, who have created an excellent platform, which has enabled me to enhance my public health knowledge and refine my biostatistics skillset.

Last but not least, I would like to thank my family: my mother, Miameen, for her relentless sacrifice, support, and faith in me; my sister, Zahra'a, for her financial support without which I would not have been able to complete my degree, my sister, Fatima, for proof reading my assignments, sometimes hours before their deadlines; and finally, my son, Jafar, who gave me all the strength I needed in the past two years.

AUTHOR'S STATEMENT PAGE

In presenting this thesis as a partial fulfillment of the requirements for an advanced degree from Georgia State University, I agree that the Library of the University shall make it available for inspection and circulation in accordance with its regulations governing materials of this type. I agree that permission to quote from, to copy from, or to publish this thesis may be granted by the author or, in his/her absence, by the professor under whose direction it was written, or in his/her absence, by the Associate Dean, School of Public Health. Such quoting, copying, or publishing must be solely for scholarly purposes and will not involve potential financial gain. It is understood that any copying from or publication of this dissertation which involves potential financial gain will not be allowed without written permission of the author.

Zainab Salah

Signature of Author

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. INTRODUCTION

1.1 CURRENT STATISTICS OF ALZHEIMER'S DISEASE:

Alzheimer's disease is the most common form of dementia among adults age 65 and above. It's ranked the sixth leading cause of death in the United States, causing more deaths than both breast and prostate cancer combined [1]. According to Alzheimer's Association and Alzheimer's Disease International, over 5 million American [1] and around 47 million worldwide [2] seniors are diagnosed with Alzheimer's or other forms of dementia. And while only 1 in every 10 seniors is believed to have some form of dementia in the United States, 1 in every three is thought to be terminal with either Alzheimer's or dementia [1]. These statistics may appear startling, but they are expected to rise in the near future as the population of older adults is rapidly increasing. Dementia, including Alzheimer's, has a severe impact on the nation's health care system and worldwide economy. It is believed that the burden of these brain disorders will cost over quarter of trillion dollars for the United States alone [1].

Alzheimer's disease was first discovered by German psychiatrist Alois Alzheimer in the early 20th century and was named after him [3,4]. In his research, Alzheimer noticed distinctive plaques in the brain histology of deceased women. He discovered that this woman had died of an unknown mental illness after studying her for five years [4]. The main noted symptoms of the disease included loss of cognitive functioning, namely thinking, remembering, reasoning, and a loss of behavioral abilities, which hindered a person's ability to accomplish activities of daily living [3]. Today, this irreversible disease is known to progress over time by destroying memory cells in the brain and results in memory decline, and notably deteriorated mental functioning. The current Alzheimer's treatments are incapable of stopping the progress of this disease and continue to baffle healthcare providers and researchers. Current scholarship is dedicated to

designing treatment which will slow down the process of the disease in order to improve lives of all those severely impacted by the disease [1]. Now more than ever it is pivotal to understand the stages of this disease and develop clear scheme of how to predict and detect its prevalence as early as possible.

1.2 STAGES OF THE DISEASE:

In order to understand the disease one must first understand the physiology of the brain. Human brains start shrinking as they begin to age. In this aging process a human brain also begins to lose weight. In Alzheimer's patients, an aging adult brain not only shrinks but also begins to develop knots. These knots are due to twisted protein fibers which release chemicals, damaging the nerve cells in the brain [5]. Although these chemicals alter the brain silently, they are lethal. Within a few years Alzheimer's symptoms are evident. The disease progresses from mild to moderate to severe cognitive impairment [1]. Currently, there are no definite molecular pathways used to define these specific stages. There are no reliable biomarkers used to diagnose the early stages of this disease. Many factors have been hypothetically attributed to the onset of Alzheimer's which include: age, genetics, dysfunctional immune system, and infectious disease [6]. Researchers are systematically attempting to validate these hypotheses by also examining the progression of the disease from normal to mild cognitive impairment to Alzheimer's patients [1,2], with the hope of developing effective drugs and therapies to slow the progression of the disease.

1.3 CLINICAL DIAGNOSIS:

Physicians depends the knowledge of family member's when consulting a new patient regarding his/her overall health. Tests such as: memory retention, problem solving, counting, attention, and language tests are used, and immediately by blood and urine tests, and magnetic

resonance imaging (MRI) scans or positron emission tomography (PET) scans [3]. While the neurological tests are useful tools in distinguishing demented from non-demented seniors, they lack the ability to detect early signs of cognitive impairment or the current stage or progression of the disease. At times, these assessments may also fail to detect the problems the patient is experiencing and completely misdiagnose the main health issue [8]. At this time there are no known and standardized biomarkers used to detect the early signs of Alzheimer's disease other than post-mortem brain histological examinations [2,3]. Therefore, conducting brain scans are only useful methods which should be used to rule out the possibility of other diseases, and not to diagnose Alzheimer's disease. Current research focuses on diagnostic biomarkers based on brain regions, cerebrospinal fluid, and blood content [3,8].

1.4 REGIONS OF INTEREST:

Several researchers have conducted studies [9-10, 12-13, 17-19] using the ADNI database in hopes of finding regions of interest (ROI) in the brain which should be the most predictive of early onset dementia, while others researchers considered regions related to conversion from normal to mild cognitive impairment (MCI), normal to Alzheimer's (AD), and/or MCI to AD. Other studies recruited volunteer participants of memory centers in Germany [11, 16] and in France [14]. The following ROI were cited in the above studies: left entorhinal cortex [10], left and/or right hippocampus [10,12,14,16,18, 19], parahippocampus [16], frontal lobe [11, 16, 18], temporal lobe [11, 16, 17, 18, 19], right inferior [12], parietal lobule [12, 14], anterior cingulate [12], amygdala [18, 15, 17], and occipital [14]. Other studies had contradictory finding regarding the association of frontal and parietal lobe with disease classification [13], there two regions were found be insignificant predictors of the disease classification.

Classification of MCI were best achieved by inclusion of amygdala and hippocampus in the model [13, 16, 17], while inclusion of ganglia and insula were insignificant for both MCI and AD [11]. Classification accuracy rates were reported from most of the above literatures, which ranged from 85% to 95%, with highest accuracy rate reported with grey matter region included. For MCI classification, accuracy rates were lower ranging from 70% to low 90's%. Classification accuracy rates seemed to be lower for studies that used smaller sample sizes.

No associations were found between Alzheimer's disease with age or gender [10,11,20]. However, statistics of the disease suggests otherwise; women are twice as likely to be diagnosed with dementia when compared to the number of men diagnosed with dementia. A review of the current research suggests one explanation for women having higher likeligood of being diagnosed: it is believed than men are more ikely to die from competing causes, such as auto accidents, gun violence, etc., which women are less likely to encounter [21]. Studies have also suggested an insignificant association between the disease and other demographic parameters such as race, educational level, and economical factors [20]. Based on these findings from previous studies and the time constraint of this study, gender, age, and demographic factors were excluded from our study.

1.5 ALZHEIMER'S DISEASE NEUROIMAGING INITIATIVE:

The data used for this study is from the Alzheimer's Disease Neuroimaging Initiative (ADNI), which is comprised of images collected by scientists of both health and cognitively impaired adult brains to develop biomarkers for tracking disease progression. ADNI, which funded by the National Institute of Health, collects demographic, clinical, neuroimaging, genetic, as well as biospecimens data from volunteer participants [22]. Researchers from around the world may access the data upon registration. Seven clinics in the United States and Canada

currently contribute to these frequently updated data archive. The initiative consists of three cohorts of data participants: ADNI started in 2004 and followed participants for 5 years, ADNI-GO started in 2009 and followed participants for two years, and ADNI2 started in 2011 with a follow-up period of 5 years. Recently, ADNI began its fourth phase called, ADNI3 which began in September of 2016 [22].

2. A Summary of The Statistical Method

2.1 Ordinary Least Square Regression:

In the basic regression, or ordinary least square (OLS) regression, a response variable is graphed against an explanatory variable, where the variability in the later explains the variability of the former. A linear curve is then fit through the data points. This linear curve quantitatively minimizes the sum of the square error and hence the nomenclature of Best Linear Unbiased Estimator (BLUE) [25, 28, 30].

$$\min_{\vec{\beta}} \| \vec{y} - x\vec{\beta} \|_2^2 \qquad (1)$$

OLS performs well when the number of variables in the models (x's) is small. However, as the number of variables increase, the possibility of correlation between these variables increases as well, which in turn upsurges the variance of the βs both drastically and erroneously [23,24,26,29]. Therefore, it seems plausible to introduce a constraining term to minimize how large a β can approach [24].

2.2 Regularization And Penalty:

There are several forms of regularization, such as forward and backward selection, best-subset, ridge regression, and LASSO [24,25,27]. In forward selection, features are added one at a time and stop only when an overfitting is detected. In backward selection, features are

permanently removed one at a time so long as they are insignificant. However, a dropped feature at an earlier step could be significant when added to the final reduced model. Best-subset selection finds the set of variables that fits the data the best. There feature selection methods are optimal when the number of variables is small [24,26,29]. Ridge regression and LASSO regression are best suited for high dimensional data [23,24,30]. Both methods use a form of penalty to limit the number of variables entering the model.

Many constraining or penalty terms have been researched, but perhaps the two most well know ones are the L1 and L2 norm, with L2 norm being more commonly used than L1 norm [25,28]. Both L1 and L2 serve the same purpose, which is to limit the size of the βs to control the erroneous variation caused by their correlation. Both the L1 and L2 end up constructing around the center of the origin. The main difference between the two is the geometric shape of the constraint. Typically, the L2 norm is circular and the L1 norm is square [25,28,29]. The square shape constraint of the L1 norm shrinks many βs to exactly zero, reducing the number of non-zero coefficients [23,24,30]. This advantageous property of L1 norm is termed feature selection, because it allows the model to select only feature or variables that have true impact on the response [23,24,27]. LASSO developed by Tibshirani in 1996 utilizes L1 norm [23-30].

2.3 Least Absolute Shrinkage And Selection Operator:

In LASSO, the goal is to minimize the sum of the square error, but within a constraining limit.

$$\min_{\vec{\beta}} \| \vec{y} - x\vec{\beta} \|_2^2 + \lambda \| \vec{\beta} \|_1 \qquad (2)$$

Where λ is the shrinkage estimator or the tuning parameter, which controls the strength of the constraint. It takes any value greater than zero ($\lambda > 0$). When λ is close to zero, equation (2) is equal to equation (1), or LASSO would produce similar parameter estimates as OLS. On the

other hand, when λ is large, then more parameter estimates approach a minimal value or zero and are therefore removed from the model [23,24,28-30]. In practice, several λs are used to calculate an array of parameters, and the right choice of λ should ultimately reduce the square error [23,26,29]. This however, does come at a price. The generated estimates are not unbiased but have reduced variance. Consequently, if the βs are all large relative to their variances, a small value of λ is selected, whereas if several β's in the model are small, a large λ is preferred to pull these parameters close to zero [24,25,29].

Similarly, a regularization term may also be added to a model with discrete outcome to avoid overfitting using the following model;

$$\min_{\vec{\beta}} -2 \ln \vec{\beta} + \lambda \| \vec{\beta} \|_1 \qquad (3)$$

Where if lambda is set too high an underfitting will occur, and if set too low, close to zero, the generated estimates will approximate the unpenalized logistic regression estimates.

3. METHODS

3.1 DATA:

For this study, ADNI UA-MRI SPM VBM [22] imaging data, and ADNI DXSUM PDXCONV ADNIALL [22] diagnosis data were merged using the unique identifier variable (ID). Both datasets were accessed and downloaded on February 15, 2017 (permission granted in January 7, 2017). Only images for the 6-month follow-up visit were included. Unrelated MRI imaging variables, such as image quality and different diagnosis statuses and conversions, were excluded, and only participants with known clinical diagnosis were included. The Final dataset included 136 variables excluding the response variables (diagnosis), and 343 observations.

3.2 STATISTICAL MODELS:

For this study three disease classifications (diagnoses) were considered: normal (control), mild cognitive impairment (MCI), and clinically diagnosed with Alzheimer's disease (AD). Initially, we only examined prediction models with binomial responses, which included only Alzheimer's and the control individual. To check whether MCI's MRIs were more to the control group or AD group, additional binomial models were fitted by either combining MCI with the control group or with the AD group. Following the binomial models, a multinomial model was fitted using the three disease classes as the dependent variable. Both ungrouped and grouped multinomial models were analyzed. The ungrouped models allowed different predictors with different response values, while grouped models were considered either all-in or all-out models. In other words, a predictor is either applicable for all three disease classifications or left out of the model entirely.

To estimate the prediction, cross validation was carried out on these models. Finally, to account for the ordering characteristic of our dependent variable, the data was fitted in an ordinal multinomial logistic regression model. All models were induced with LASSO penalty.

3.3 STATISTICAL PACKAGES:

Data were merged using SAS 9.3 software [31]. The same software was also used to generate the demographic statistics. The remaining statistical analysis were performed in R [32] using mainly two packages for non-ordinal and ordinal LASSO logistic regression, respectively. The first package was glmnet package [33] with glmnet () function for model fitting, predict () function for coefficient prediction, and cv.glmnet () for cross validation [34]. The second package was glmnetcr package [35] with glmnet.cr () function for ordinal logistic modeling fitting, and nonzero.glmnet.cr () for displaying non-zero coefficients [36].

4. RESULTS AND DISCUSSION

4.1 DATASET:

In this study, the task was to predict the diagnosis status of individuals in the selected ADNI study dataset based on a spectrum of MRI data. A total of 343 individuals' MRI were examined within a data matrix containing 174 total number variables. The number of covariates considered large rendering it a good candidate for LASSO regression model fitting technique. Lasso allows for penalized regression techniques that are flexible enough for high-dimensional data and can be applied directly without the need of dimension reduction [23-30]

The dataset was examined and variables not related to MRI imaging, which included image quality parameters, several recoded diagnoses, recoded date of visits, and several subject identifications, were excluded. The final data set included 83 AD, 146 MCI, and 114 control individuals with total of 137 predictors (see Appendix A for detailed list of predictors). Number of predictors included in the models, subsequent to model convergence was 109 for the AD vs control. While number of predictors entered in the model, after model convergence, was 117 for the AD vs. MCI&normal, AD&MCI vs. normal, and the multinomial models.

4.2 DEMOGRAPHICS:

The overall mean age of the participants in the sample was 72.1 (±6.7). Majority of the sample have college degrees. Males composed about 58% of the sample, and the rest were female seniors. Majority of the sample were married (74.3%). The rest were either widowed, divorced, never married, or with unknown marital statuses, 15%, 7%, 3.2%, and 1%, respectively. Majority of the participants were retired (84.0%), and only a small fraction of them were currently working (9.9%). Finally English was the primary language of the participants (98.3%). Detail demographic statistics appear in Table 1.

4.3 LASSO Model Fitting:

The coefficients for the converged model predictors were plotted against two different scales: the fraction deviance explained, and the L1 norm for each set of comparison (Figure 1-4). In these plots, each curve represents a coefficient in the model. In these plots, each colored line summarizes the path of a different coefficient in the model [33-34]. Lambda is used as the regularization term, and thus as lambda decreased or approaches zero, the more coefficients enter the model, and the model approaches the OLS solution in which all coefficients are allowed in the model. Likewise, when lambda is allowed to increased, the regularization term is allowed to have a greater effect by allowing as few variables into the model, leaving more coefficients with a zero value. While variables that enter the model early on are considered the most predictive of the outcome of interest, the variables that enter the model later are less predictive. In other words, the predictors go into the model in the order of their magnitude of predictive effect. As predictors enter into the model, the slope of the coefficient path of their predictors already in the model is affected [33-34].

The performance of LASSO algorithm was then analyzed by examining the specificity, sensitivity, and the accuracy rates of prediction using several lambda values. The results for the specificity, sensitivity, and accuracy rates are summarized in Figure 6. As the lambda values increased, the accuracy of the model simultaneously decreased. At smaller lambda values when no regularization term was added to the model and all predictors were in the model, the sensitivity was found to be zero, with specificity at its peak at 100%, with accuracy fluctuation between the models. As the lambda values increased, and on average all models' accuracies rates decreased, some of the models did outperform others under different lambda values. For instance, while the accuracy rates for the binomial models were low 70% at $\lambda 10$, the multinomial

models were barely over 50% at the same lambda size. Furthermore, the AD vs. control model outperformed all model at low lambda values.

4.4 Cross Validation:

Cross validation is a useful technique that estimates prediction error of a model. In principle, once a model is fitted, it is trained on an external sample to determine the accuracy of the prediction. However, practically, cross-validation is accomplished by splitting the data into K sets, and a model is fitted using one of the sets and then trained on the rest. Hence, the nomenclature K-fold cross-validation [37]. For this study, 10-fold cross-validation was carried to choose the tuning parameter, λ, with the smallest prediction error. Figures 7-11 depict the cross-validated error curves for binomial (AD vs. Control, AD&MCI vs. Control, AD vs. MCI&Control), multinomial, and grouped multinomial logistic regression models. All model generated similar lambda min, λ_{min}, and lambda plus 1 standard error, λ_{1se}, values, except for the multinomial model.

4.5 Model Coefficients:

The multinomial model was the only model that the contained non-zero coefficients (Table 2). Non-zero coefficients for different lambda values were determined based on the different models (Table 3). The results suggested an unstable prediction error for all the models except for the Grouped-Multinomial models and possibly for the binomial AD&MCI vs. Control. When different lambda values were trained, the number of coefficients increased as the lambda value increased (Table 3). The results fluctuated with a high level of uncertainty. In certain brain regions there seemed to be an association with disease classification regardless of lambda values. For instance, putamen appeared in many of the models and in all lambda values. However, for the multinomial models, different brain regions appeared to be associated with different disease

classification. When the grouped multinomial logistic regression is applied, frontal lube and putamen appear to have the most effect on disease classification.

Based on the prediction error curves the following values of lambda seemed to generate the least mean prediction error: binomial model (AD&MCI vs. Control) s = 0.3-0.07, binomial model (AD vs. MCI&Control), s = 0.015-0.3, and grouped-multinomial s = 0.02. The lambda values were used to determine the coefficients most predictive of disease classification (Table 4). Disease classification seemed to be related to different brain regions depending on classification criteria. When AD is combined with MCI and compared to control individuals, frontal lobe and supplementary motor, and vermis regions are the best predictors. On the other hand, when AD is compared to the combined MCI and control individuals, many regions are included in the model which suggested ambiguity of this type of classification. This is also true when classification is done on grouped level multinomial.

Frontal lobe was found to be one of the ROI in earlier studies of Alzheimer's disease [3, 16, 18], whereas the vermis, supplementary motor, or putamen region were not cited in previous literature as one of the ROI. Other cited ROI such as hippocampus [10, 2, 4, 6, 8,19] did not appear in any of the model as good predictor of disease classification. Amygdala [15,18, 9] was only good predictor in the grouped-multinomial model. Temporal lobe [10,16,17,18,19] predicted disease classification only at small tuning parameter values for the binomial AD vs. MCI&Control and grouped-multinomial models. Parietal lobe [12,14] also, only predicted at small lambda values, binomial AD&MCI vs. Control and grouped-multinomial models.

4.6 ORDINAL LASSO

Dementia progresses though several stages of cognitive impairment before developing into late stage dementia or Alzheimer's. For that reason, there is a natural ordering proper

disease classification using variables which start with the control followed by mild cognitive impairment to an official Alzheimer's diagnosis among participants. Accordingly, this study sought to find a model that took into consideration the ordinal property of the response variable into account, and that is an ordinal LASSO logistic regression model.

Figures 12 and 13 depicts the coefficient entry path into the models by betas, lambda, and steps for backward and forward selection method, respectively [35-36]. Figure 14, 16, and 18 presents the trends of the Bayesian Information Criterion (BIC) and Akaike Information Criterion (AIC) over steps, lambda, and betas for backward selection. Likewise, figures 15, 17, and 19 present the trends of both BIC and AIC over steps, lambda, and betas for forward selection. Based on these plots, models 2 through 6 were determined to be indifferent and have the lowest BIC and AIC values. Non-zero coefficients for the models used during these steps are seen in Table 5.

At step one, the empty model was fitted. In step two, two and four predictors entered the model in the backward and forward selection, respectively. Both supplementary motor and putamen regions of the brain were good predictors at this step. The backward selection gradually increased the number of predictors as the number of steps increased. In contrast to the forward selection, the models at step 4-6 remained stable with the same nine predictors. Moreover, certain brains regions were found to be consistently good predictors in both methods of selection. These regions included: supplementary motor, the putamen, the pallidum, the cerebral, and the vermis regions. These results are not in agreement with earlier studies and have not been previous cited as brain regions which are typically associated with the disease. However, these findings are congruent with the non-ordinal results presented earlier in this study.

The different models' sensitivity, specificity, and accuracy rates are displayed in Figure 20. These rates demonstrate the underperformance of these models. While MCI were accurately predicted in all models, the models failed to predict both normal and AD classes. These rates only improved with the increased BIC and AIC levels. More specifically, the accuracy rates improved around the midpoint of BIC and AIC, or at step 25 and higher.

5. Limitations And Direction For The Future

Although two studies [20,21] documented lack of association between demographic characteristics, several others found significant association between certain demographics and Alzheimer's disease. Among demographics found to be related to the disease are age [38], gender [39], education [40], and race [41]. Although majority of these studies date back to 1980s, 1990s, and early 2000, we believe the decision to exclude these factors from the current study limits the results and conclusion.

One of LASSO's disadvantages is its inability to distinguish between correlated data. With brain MRI unfortunately, all of the brain regions are somehow correlated. LASSO just picks one over the other, indeed, we do not know which variable is chosen. Modeling with a different constraining penalty may ameliorate this problem. Elastic net is a regularization introduced to the model that uses a mix of L1 and L2 norms. It simultaneously shrinks the coefficients and accomplishes sparse selection. This method could possibly be a better method to answer the research question at hand.

As a final note, in the era of fast growing technology, the acquisition of brains images is becoming readily available. At first glance, these images provide insight into the brain regions' structure. These is images can further be scrutinized and transferred to pure data than can be

further analyzed. Having prior set models for disease classification can help health providers easily diagnose patients. It is therefore imperative to explore this field further.

REFERENCES:

1. Alzheimer's Association. (2017). 2017 Alzheimer's disease facts and figures full report. Retrieved from Alz's website; http://www.alz.org/documents_custom/2017-facts-and-figures.pdf
2. Alzheimer's Disease International. (2016) World Alzheimer report 2016 improving healthcare for people living with dementia, coverage, quality and cost now and in the future. Retrieved from ADI website; www.alz.co.uk/worldreport2016
3. National Institute of Aging. (2016). Alzheimer's Disease Fact Sheet. Retrieved from Alzheimer's Disease Education and Referral Center website; file:///C:/Users/Zainab/Downloads/ad_fact_sheet-2016_reprint-final.pdf
4. Hippius, H. (2003). The discovery of Alzheimer's disease. *Dialogues in Clinical Neuroscience*, 5(1), 101-108.
5. Nazarko, Linda. (2011). Understanding dementia: diagnosis and development. *British Journal of Healthcare Assistants*, 5(5), 216-220.
6. Armstrong, R. A. (2013). What causes Alzheimer's disease? *Folia Neuropathology*, 51(3), 169-188. DOI: 10.5114/fn.2013.37702
7. Mayo Clinic. (2017). Diagnosing Alzheimer's: How Alzheimer's is diagnosed. Retrieved from Mayo Clinic website; http://www.mayoclinic.org/diseases-conditions/alzheimers-disease/in-depth/alzheimers/art-20048075
8. Duchesne, S., Valdivia, F., Mouiha, Ab., & Robitaille, N. (2016). High-dimensional morphometry longitudinal analysis in Alzheimer's disease. *Neurobiology Aging*, 36(1), 475-486.
9. Ewers, M., Walsh, C., Trojanowski, J. Q., Shaw, L. M., Petersen, R. C., Clifford, J. Jr. , …, & Hampel, H. (2012). Prediction of conversion from mild cognitive impairment to Alzheimer's disease dementia based upon biomarkers and neuropsychological test performance. *Neurobiology Aging*, 33(7), 1203-1214.
10. Meyer, S., Mueller, K., Stuke, K., Bisenius, S., Diehl-Schmid, J., Jessen, F., …, & Schroter, M. (2017). Predicting behavioral variant frontotemporal dementia with pattern classification in multi-center structural MRI data. *NeuroImage: Clinical*, 14, 656-662.
11. Beheshti, I., Demirel, H., & Matsuda, H. (2017). Classification of Alzheimer's disease and prediction of mild cognitive impairment-to-Alzheimer's conversion from structural magnetic resonance imaging using feature ranking and a genetic algorithm. *Computers in Biology and Medicine*. 83, 109-119.
12. Long, X., Chen, L., Jiang, C., & Zhang, L. (2017). Prediction and classification of Alzheimer disease based on quantification of MRI deformation. *PLOS ONE*, DOI:10.1371/journal.pone.0173372.
13. Mesrob, L., Sarazin, M., Hahn-Barma, V., de Souza, L. C., Dubois, B., Gallinari, P., & Kinkingnehun, S.. (2012). DTI and structural MRI classification in Alzheimer' disease. *Advances in Molecular Imaging*, 2, 12-20.

14. Vemuri, P., Wiste, H.J., Weigand, S.D., Shaw, L. M., & Trojanowski, J.Q., Weiner, M.W., Knopman, D.S., Petersen, R.C., Jack, Jr., C.R. (2009). MRI and CSF biomarker in normal, MCI, and AD subjects: diagnostic discrimination and cognitive correlations. *Neurology*, 73(4), 287-293.
15. Goerlich, K. S., Votinov, M., Dicks, E., Ellendt, S., Csukly, G., & Habel, U. (2017). Neuroanatomical and Neuropsychological Markers of Amnestic MCI: A Three-Year Longitudinal Study in Individuals Unaware of Cognitive Decline. *Frontiers in Aging Neuroscience*, 9(34).
16. Hua, X., Lee, S., Hibar, D. P., Yanovsky, I., Leow, A., Toga, A. W., …, & Thompson, P. M. (2010). Mapping Alzehimer's disease progression in 1309 MRI scans: Power estimates for different inter-scan intervals. *Neuroimage*, 51(1), 63-75.
17. Qiu, A., Fennema-Notestine, C., Dale, A. M., & Miller, M. I. (2009). Regional shape abnormalities in mild cognitive impairment and Alzheimer's disease. *Neuroimage*, 45(3), 656-661. ADNI.
18. Zhang, D., Wang, Y., Zhou, L., Yuan, H., & Shen, D. Multimodal classification of Alzheimer's disease an mild cognitive impairment. (2011). *Neuroimage*, 55(3), 856-867. ADNI
19. Molero, A. E., Pino-Ramirez, G., Maestre, G. E. (2001). Modulation by age and gender of risk for Alzhimer's disease and vascular dementia associated with apoliopoprotien E-epsilon4 allele in Latin Americans: finding from Maracaibo Ageing Study. *Nueroscience Lett*, 307 (1), 5-8.
20. Ballard, C., Gautheir, S., Corbett, A., Brayne, C., Aarslan, D., & Jones, E. (2011). Alzheimer's disease. *The Lancet*, 377(9770), 1019-1031.
21. Mielke, M. M., Vemuri, P., & Rocca, W. A. (2014). Clinical epidemiology of Alzheimer's disease: assessing sex and gender differences. *Clinical Epidemiology*, 6, 37-48.
22. ADNI. (2017). Alzheimer's Disease Neuroimaging Initiative. Retrieved from ADNI website; http://adni.loni.usc.edu/
23. Tibshirani, R. (1996). Regression shrinkage and selection via lasso. *Journal of Royal Statistical Society. Series B (Methodological)*, 58(1), 267-288.
24. Tibshirani, R. (2011). Regression shrinkage and selection via lasso: a retrospective. *Journal of Royal Statistical Society. Series B (Methodological)*, 72(3), 273-282.
25. Fu, W. J. (1998). Penalized regressions: The bridge versus the lasso. *Journal of Computational and Graphical Statistics*, 7(3), 397-416.
26. Hans, C. (2009). Bayesian lasso regression. *Biometrika*, 96(4), 835-845.
27. Reid, S., Tibshirani, R., & Friedman, J. (2014). A study of error variance estimation in lasso regression. Statistica Sinica, 26(1).
28. Meier, L, Geer, S. V. D., & Buhlmann, P. (2008). The group lasso for logistic regression. *Journal of Royal Statistical Society. Series B (Methodological)*, 70(1), 53-71.
29. Zhang, C-H., & Huang, J. (2008). The sparsity and bias of the lasso selection in high-dimensional linear regression. The Annals of Statistics, 36(4), 1567-1594.
30. Huang, J., Ma, S., Zhang, C-H. (2008). Adaptive lasso for sparse high-dimensional regression models. Statistica Sinica, 18(4), 1603-1618.

31. SAS Institute Inc., 2011. Base SAS® 9.3 Procedures Guide. Cary, NC: SAS Institute Inc.
32. R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, North Carloina, USA. URL http://www.R-project.org/.
33. Friedman, J., Hastie, T., Simon, N., Tibshirani, R. (2017). Glment: lasso and elastic-net regularized generalized linear models. R Package version 2.0-10.
34. Hastie, T., Qian, J. (2016). Glmnet Vignette.
35. Archer, K. J. (2015). Glmnetcr: fit a penalized constrained continuation ratio model for predicting an ordinal response. R package version 1.0.2.
36. Archer, K. J. (2015). Glmnetr: an R package for ordinal response prediction in high-dimensional data setting.
37. Kohavi, R. (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. *International Joint Conference on Artifical Intelligence, 14*(2).
38. Kawarabayashi, T., Younkin, L. H., Saido, T. C., Shoji, M., Ashe, K. H., & Younkin, S. G. (2001), Age-dependent changes in brain, CSF, and plasma amyloid β protein in the Tg2576 transgenic mouse model of Alzheimer's disease. *Journal of Neuroscience, 21*(2), 372-281.
39. Heyman, A., Willkinson, W. E., Stafford, J. A., Helms, M. J., Sigmon, A. H., & Weinberg, T. (1984). Alzheimer's disease: A study of epidemiological aspects. *Annals of Neurology, 14*(4), 335-341.
40. Stern, Y., Gurland, B., & Tatemichi, T. (1994). Influence of Education and Occupation on the Incidence of Alzheimer's Disease. *Journal of American Medical Association, 271*(13), 1004-1010.
41. Green, R. C., Cupples, A., Go, R. (2002). Risk of dementia among White and African American of patients with Alzheimer disease. *Journal of American Medical Association, 287*(3), 329-336.

APPENDIX A: LIST OF VARIABLES

PRECENTL: Left Percentile
PRECENTR: Right Percentile
FRONTSUPL: Left Frontal Supplementary
FRONTSUPR: Right Frontal Supplementary
FRONTSORBL: Left Frontal sorbitol
FRONTSORBR: Right Frontal sorbitol
FRONTMIDL: Left Frontal Medial
FRONTMIDR: Right Frontal Medial
FRTMIDORBL: Left Frontal Medial Orbital
FRTMIDORBR: Right Frontal Medial Orbital
FRONTINOPL: Left Frontal Inferior Opercula
FRONTINOPR: Right Frontal Inferior Opercula
FRONTINTRL: Left Frontal Inferior
FRONTINTRR: Right Frontal Inferior
FRONTINOBL: Left Frontal Inferior Orbital
FRONTINOBR: Right Frontal Inferior Orbital
ROLANDOPL: Left Rolandic Operator
ROLANDOPR: Right Rolandic Operator
SUPMOTORL: Left Supplementary Motor
SUPMOTORR: Right Supplementary Motor
OLFACTL: left Olfactory
OLFACTR: Right Olfactory
FRONTSMEDL: Left Frontal Supplementary Medial
FRONTSMEDR: Right Supplementary Medial
FRTMEDORBL: Left Frontal Medial Orbital
FRTMEDORBR: Right Frontal Medial Orbital
RECTUSL: Left Rectus
RECTUSR: Right Rectus
INSULAL: Left Insula
INSULAR: Right Insula
CINGANTL: Left Cingulum Angular Bundle
CINGANTR: Right Cingulum Angular Bundle
CINGMIDL: Left Cingulum Medial
CINGMIDR: Right Cingulum Medial
CINGPOSTL: Left Cingulum Posterior
CINGPOSTR: Right Cingulum Posterior
HIPPL: Left Hippocampus
HIPPR: Right Hippocampus
PARAHIPPL: Left Parahippocampus
PARAHIPPR: Right Parahipocampus
AMYGDL: Left Amygdala
AMYGDR: Right Amygdala
CALCARINEL: Left Calcarine
CALCARINER: Right Calcarine
CUNEUSL: Left Cuneus

CUNEUSR: Right Cuneus
LINGUALL: Left Lingual
LINGUALR: Right Lingual
OCCSUPL: Left Supplementary Occipital
OCCSUPR: Right Supplementary Occipital
OCCMIDL: left Occipital Medial
OCCMIDR: Right Occipital Medial
OCCINFL: Left Occipital Inferior
OCCINFR: Right Occipital Inferior
FUSIFORML: Left Fusiform
FUSIFORMR: Right Fusiform
POSTCENTL: Left Posterior Central
POSTCENTR: Right Posterior Central
PARIETSUPL: Left Supplementary Parietal
PARIETSURP: Right Supplementary Parietal
PARIETINFL: Left Inferior Parietal
PARIETINFR: Right Inferior Parietal
SUPRAMARGL: Left Supplementary amygdala
SUPRAMARGR: Right Supplementary amygdala
ANGULARL: Left Angular
ANGULARR: Right Angular
PRECUNEUSL: Left PreCuneus
PRECUNEUSR: Right PreCuneus
PARCENTLBL: Left Paracentral Lobe
PARCENTLBR: Right Paracentral Lobe
CAUDATEL: Left Caudate
CAUDATER: Right Caudate
PUTAMENL: Left Putamen
PUTAMENR: Right Putamen
PALLIDUML: Left Pallidum
PALLIDUMR: Right Pallidum
THALAMUSL: Left Thalamus
THALAMUSR: Right Thalamus
HESCHLL: Left Heschel
HESCHLR: Right Heschel
TEMPSUPL: Left Supplementary Temporal
TEMPSURP: Right Supplementary Temporal
TEMPPLSUPL: Left Supplementary Temporal Plane
TEMPPLSUPR: Right Supplementary Temporal Plane
TEMPMIDL: Left Medial Temporal
TEMPMIDR: Right Medial Temporal
TEMPPLMIDL: Left Medial Temporal Plane
TEMPPLMIDR: Right Medial Temporal Plane
TEMPINFL: Left Inferior Temporal
TEMPINFR: Right Inferior Temporal
CEREBCR1L: Left Cerebral Cortex Region 1
CEREBCR1R: Right Cerebral Cortex Region 1

CEREBCR2L: Left Cerebral Cortex Region 2
CEREBCR2R: Right Cerebral Cortex Region 2
CEREBCR3L: Left Cerebral Cortex Region 3
CEREBCR3R: Right Cerebral Cortex Region 3
CEREB45L: Left Cerebral Region 4-5
CEREB45R: Right Cerebral Region 4-5
CEREB6L: Left Cerebral Region 6
CEREB6R: Right Cerebral Region 6
CEREB7BL: Left Cerebral Blood
CEREB7BR: Right Cerebral Blood
CEREB8L: Left Cerebral Region 8
CEREB8R: Right Cerebral Region 8
CEREB9L: Left Cerebral Region 9
CEREB9R: Right Cerebral Region 9
CEREB10L: Left Cerebral Region 10
CEREB10R: Right Cerebral Region 10
VERMIS12: Vermis Region 1-2
VERMIS3: Vermis Region 3
VERMIS45: Vermis Region 4-5
VERMIS6: Vermis Region 6
VERMIS7: Vermis Region 7
VERMIS8: Vermis Region 8
VERMIS10: Vermis Region 10
ETIV: Estimated Intracranial Volume
DXCURREN: Current Diagnosis (1: Normal, 2: MCI, 3: AD)

APPENDIX B: TABLES

Table 1: Descriptive Statistics of the Participants

Characteristic	N(%)
Age	75.10 (±6.72)
Education	16 (14-18))
Male	208 (58.26)
Right-Handed	323 (90.48)
Marital Status	
Married	255 (74.34)
Widowed	52 (15.16)
Divorced	24 (7.0)
Never Married	11 (3.21)
Unknown	1 (0.29%)
Retired	
Yes	288 (83.97)
No	34(9.91)
Missing	21 (6.12)
Primary language	
English	337 (98.25)
Spanish	2 (0.58%)
Unknown	28 (0.82)

Age: mean (±SD), Year of education: median (IQR)

Table 2: Non-Zero Coefficients Cross-Validation Multinomial LASSO – AD&MCI vs. Control

Control		MCI		AD	
Coefficient	Coefficient Estimates	Coefficient	Coefficient Estimates	Coefficient	Coefficient Estimates
FRONTINOPR	4.4076812	FRONTINOPL	5.2001940	FRONTSUPR	1.2870474
SUPMOTORL	2.2581351	RECTUSL	4.4316612	FRONTMIDR	3.3329974
CUNEUSR	-2.3267197	CINGMIDL	0.6565016	FRTMIDORBR	1.6357613
PARIETSUPR	0.1030343	CINGPOSTR	-0.8786198	FRTMEDORBR	4.0945468
PARIETINFL	2.3798372	AMYGDL	-0.3010206	OCCINFL	-2.0094674
PUTAMENR	2.9849488	CALCARINER	1.7689149	POSTCENTR	-2.9393642
HESCHLR	-0.2553623	OCCMIDR	-2.7842050	SUPRAMARGR	-0.3856534
TEMPPLMIDL	0.1252874	OCCINFL	0.2076641	PARCENTLBL	-1.5763218
TEMPINFL	3.8724262	SUPRAMARGL	1.2655111	PARCENTLBR	-1.2712380
CEREBCR2L	0.5981673	THALAMUSL	-0.2908878	THALAMUSL	0.6101122
CEREBCR2R	2.2958389	TEMPMIDL	0.6243081	HESCHLR	0.7034065
CEREB45L	1.0101785	CEREB3L	-1.6049941	CEREB7BR	-0.2797097
CEREB7BR	0.2986538	CEREB10L	-0.6472638	CEREB10R	0.5352596
CEREB10R	-0.4135493			VERMIS6	1.0312296
VERMIS12	-1.7909975			ETIV	-0.3119901
VERMIS10	-0.3605663				

Table 3: Cross-Validation Coefficients Under Different Lambda Values.

Class	λ_1	λ_2	λ_5	λ_7	λ_{10}
Binomial (AD vs. Control)	None	PARCENTLBR PUTAMENR	SUPMOTORR FRTMEDORBR POSTCENTR PARCENTLBL PARCENTLBR PUTAMENR CEREB7BR VERMIS12	SUPMOTORL FRTMEDORBR OCCINFL POSTCENTR PARCENTLBR PUTAMENR PALLIDUML CEREB7BR CEREB10R VERMIS12 VERMIS6	FRONTSURP FRTMIDORBR FRONTINOPR SUPMOTORL FRTMEDORBR OCCINFR POSTCENTR PARCENTLBL PARCENTLBR PUAMENR HESCHLR TEMPPLMIDL TEMPINFL CEREBCR2L CEREBCR2R CEREB7BR CEREB10R VERMIS45 VERMIS6
Binomial (AD&MCI vs. Control)	None	PUTAMENR	SUPMOTORL PUTAMENR VERMIS12	FRONTINOPL SUPMPTPRL RECTUSL PUTAMENR CEREBCR2R CEREB7BR VERMIS12	FRONTINOPL FRONTINOPR SUPMOTORL RECTUSL CINGANTR CUNEUSR PARIETINFL PUTAMENR HESCHLR TEMPINFL CEREBCR2R CEREB7BR VERMIS12
Binomial (AD vs. MCI&Control)	None	PARCENTLBL	FRTMEDORBR POSTCENTR PARCENTLBL PARCENTLBR CEREB8R	FRTMEDORBR OCCINFL POSTCENTR SUPRAMARGR PARCENTLBL PARCENTLBR CEREB7BR VERMIS12 VERMIS6	FRONTMIDR FRIMIDORBR FRTMEDORBR OCCINFR POSTCENTR SUPRAMARGR PARCENTLBL PARCENTLBR THALAMUSL HESCHLR TEMPSURL CEREB7BR

					CEREB10R VERMIS12 VERMIS6 ETIV
Multinomial (AD vs. MCI vs. Control)					
Control	None	PUTAMENR	SUPMOTORL PUTAMENR CEREBCR2R	SUPMOTORL PUTAMENR CEREBCR2R CEREB5BR VERMIS12	FRONTINOPR SUPMOTORL CUNEUSR PARIETINFL PUTAMENR TEMPINFL CEREBCR2R CEREBCR7R VERMIS12 VERMIS10
MCI	None	FRONTINOPL	FRONTINOPL	FRONTINOPL RECTUSL SUPRAMARGL CEREB3L	FRONTINOPL RECTUSR CALCARINER OCCMIDR SUPRAMARGR CEREB3R
AD	None	None	FRTMEDORBR PARCNTLBL	FRTMEDORBR PARCENTLBL PARCENTLBR	FRONTMIDR FRTMIDORBR FRTMEDORBR OCCINFL POSTCENTR PARCENTLBL PARCENTLBR THALAMUSL HESCHLR CEREB7BR CEREB10R VERMIS6
Grouped-Multinomial (AD vs. MCI vs. Control)		FRONTINOPL PUTAMENR	FRONTINOPL SUPMOTORL FRTMEDORBR RECTUSL PARCENTLBL PUTAMENR CEREBCR2R CEREB7BR VERMIS12	FRONTINOPL FRONTINOPR SUPMOTORL FRTMEDORBR RECTUSL SUPRAMARGL PARCENTLBL PARCENTLBR PUTAMENR CEREBCR2R CEREB3L CEREB7BR	FRONTMIDR FRTMIDORBR FRONTINOPL FRONTINOPR SUPMOTORL FRTMEDORBR RECTUSL CALCARINER CUNEUSR OCCMIDR OCCINFL POSTCENTR

	VERMIS12	PARIETINFL SUPRAMARGL PARCENTLBL PARCENTLBR PUTAMENR THALAMUSL HESCHLR TEMPINFL CEREBCR2R CEREB3L CEREB45L CEREB7BR CEREB10R VERMIS12 VERMIS6

Table 4: Predictor of Cross-Validated Models Based on Prediction Error Suggested Lambda Values.

Class	Penalty Value (*s*)	Non-Zero Coefficients
Binomial (AD&MCI vs. Control)	*s* = 0.03	FRONTINOPL, SUPMOTORL, RECTUSL, CUNEUSR, PARIETINFL, PUTALMENR, HESCHLR, CEREBCR2R, CEREB7BR, VERMIS12
	s = 0.04	SUPMOTORL, PUTAMENR, VERMIS12
	s = 0.05	FRONTINOPL, SUPMOTORL, FRTMEDORBR, RECTUSL, PARCENTLBL, PUTAMENR, CEREBCR2R, CEREB7BR, VERMIS12
	s = 0.055	FRONTINOPL, SUPMOTORL, PARACENLBL, PUTAMENL, VERMIS12
	s = 0.06	FRONTINOPL, SUPMOTORL, PARCENTLBL, PUTAMENR
	s = 0.07	FRONTINOPL, PUTAMENR
Binomial (AD vs. MCI&Control)	*s* = 0.015	FRONTSUPR, FRONTMIDR, FRTMIDORBR, FRONTINOPL, FRONTSMEDL, FRTMEDORBR, RECTUSL, OCCSUPR, OCCINFL, POSTCENTR, SUPRAMARGR, PARCENTLBL, PARCENTLBR, CAUDATEL, PALLIDUML, THALAMUSL, HESCHLR, TEMPSUPL, TEMPPLMIDL, CEREB3L, CEREB7BR, CEREB10R, VERMIS12, VERMIS6, VERMIS8, ETIV
	s = 0.03	FRTMEDORBR, OCCINFL, POSTCENTR, SUPRAMARGR, PARCENTLBL, PARCENTLBR, CEREB7BR, VERMIS12, VERMIS6
Grouped-Multinomial (AD vs. MCI vs. Control)	*s* = 0.02	FRONTSUPR, FRONTMIDR, FRTMIDORBR, FRONTINOPL, FRONTINOPR, FRONTINOBR, SUPMOTORL, FRTMEDORBR, RECUSL, CINGANTR, CINGMIDL, CINGPOSTR, AMYGDL, CALCARINER, CUNEUSR, OCCMIDR, OCCINFL, POSTCENTR, PARIETSUPR, PARITINFL, SUPRAMARGL, SUPRAMARGR, PARCENTLBL, PARCENTLBR, CAUDATEL, PUTAMENR, THALAMUSL, HESCHLR, TEMPSUPL, TEMPMIDL, TEMPPLMIDL, TEMPINFL, CEREBCR2L, CEREBCR2R, CEREB3L, CEREB3R, CEREB45L, CEREB6L, CEREB7BR, CEREB10L, CEREB10R, VERMIS12, VERMIS6, VERMIS8, VERMIS10, ETIV

Table 5: Non-Zero Coefficients for Backward and Forward Models.

Step 1	Step 2	Step 3	Step 4	Step 5	Step 6
			Backward Models		
None	SUPMOTORL	SUPMOTORL	SUPMOTORL	SUPMOTORL	SUPMOTORL
	PUTAMENR	PUTAMENR	PUTAMENR	FREMEDORBR	FREMEDORBR
		VERMIS12	CEREB7BR	PUTAMENR	PUTAMENR
			VERMIS12	CEREB7BR	PALLIDUML
				VERMIS12	CEREB7BR
					VERMIS12
			Forward Models		
None	SUPMOTORL	SUPMOTORL	SUPMOTORL	SUPMOTORL	SUPMOTORL
	PARCENTLBL	POSTCENR	FRTMEDORBR	FRTMEDORBR	FRTMEDORBR
	PARCENTLBR	PARCENTLBL	POSTCENR	POSTCENR	POSTCENR
	VERMIS12	PARCENTLBR	PARCENTLBL	PARCENTLBL	PARCENTLBL
		PUTAMENR	PARCENTLBR	PARCENTLBR	PARCENTLBR
		PALLIDUML	PUTAMENR	PUTAMENR	PUTAMENR
		CEREB7BR	PALLIDUML	PALLIDUML	PALLIDUML
		VERMIS12	CEREB7BR	CEREB7BR	CEREB7BR
			VERMIS12	VERMIS12	VERMIS12

APPENDIX C: FIGURES

Figure 1: Trace Plot of Coefficients Fit by Binomial LASSO – AD vs. Control (Left: Coefficients are plotted against the Fraction Deviance Explained; Right: coefficients are plotted against the L1 Norm)

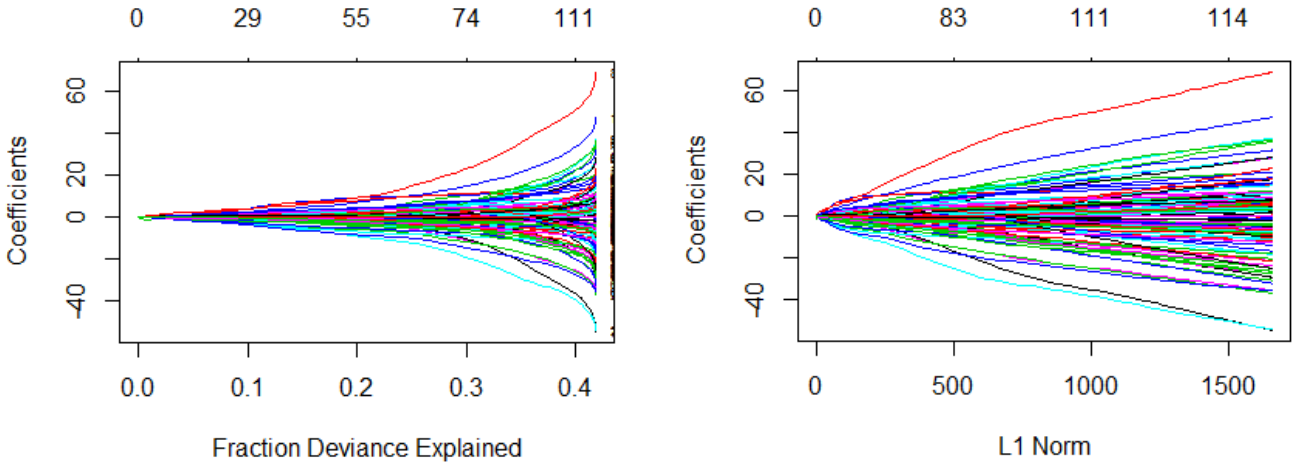


Figure 2: Trace Plot of Coefficients Fit by Binomial LASSO – AD and MCI vs. Control (Left: Coefficients are plotted against the Fraction Deviance Explained; Right: coefficients are plotted against the L1 Norm)

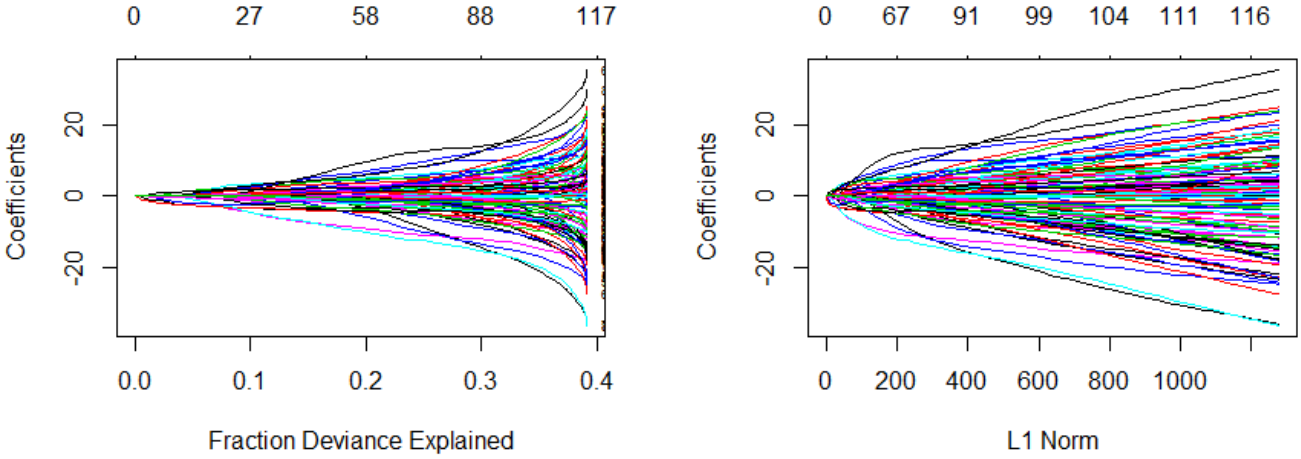


Figure 3: Trace Plot of Coefficients Fit by Binomial LASSO – AD vs. MCI and Control (Left: Coefficients are plotted against the Fraction Deviance Explained; Right: coefficients are plotted against the L1 Norm)

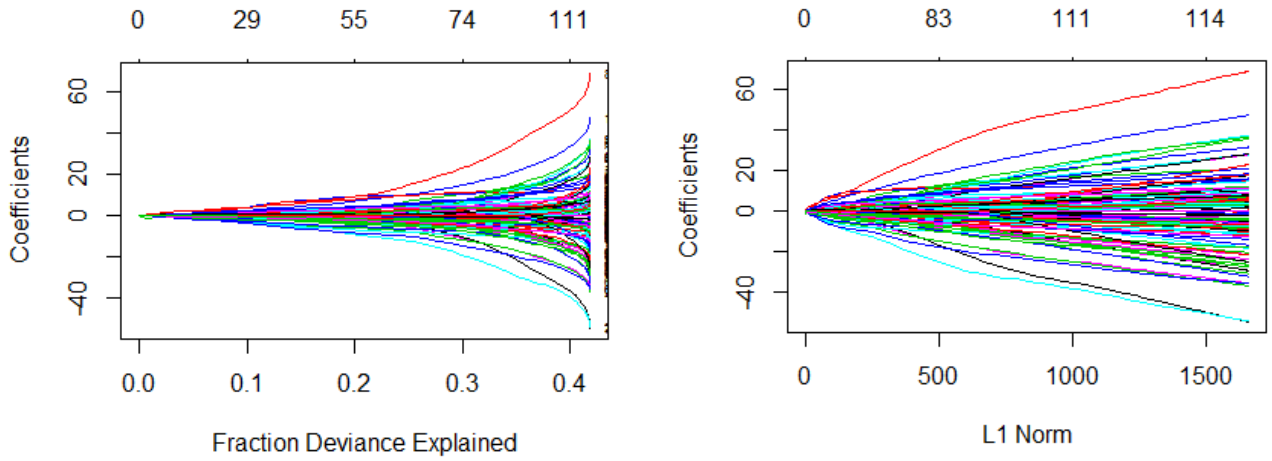


Figure 4: Coefficients Fit by Multinomial LASSO – AD vs. MCI vs. Control (Left: Coefficients are plotted against the Fraction Deviance Explained; Right: coefficients are plotted against the L1 Norm)

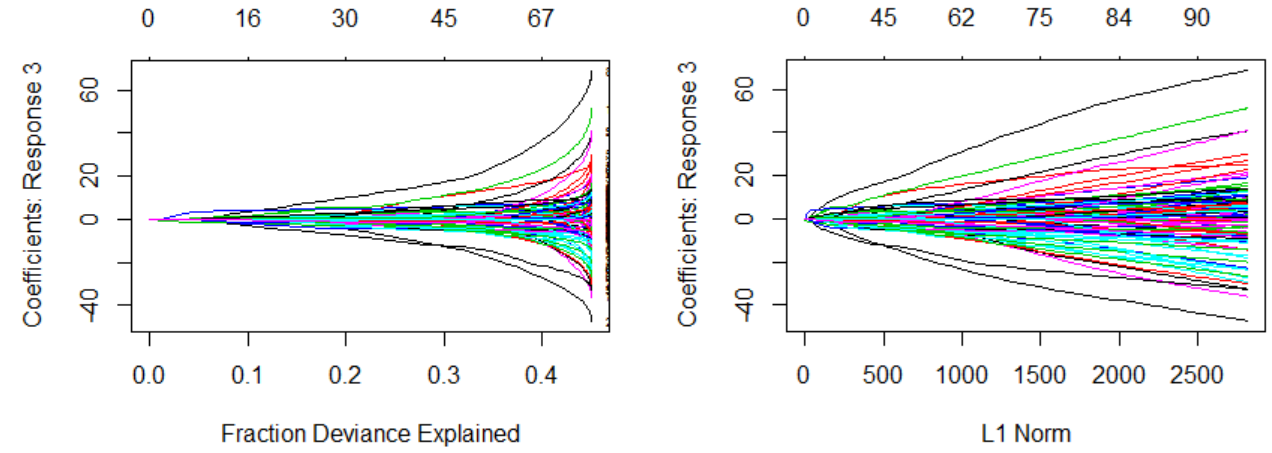


Figure 5: Coefficients Fit by Grouped Multinomial LASSO – AD vs. MCI vs. Control (Left: Coefficients are plotted against the Fraction Deviance Explained; Right: coefficients are plotted against the L1 Norm)

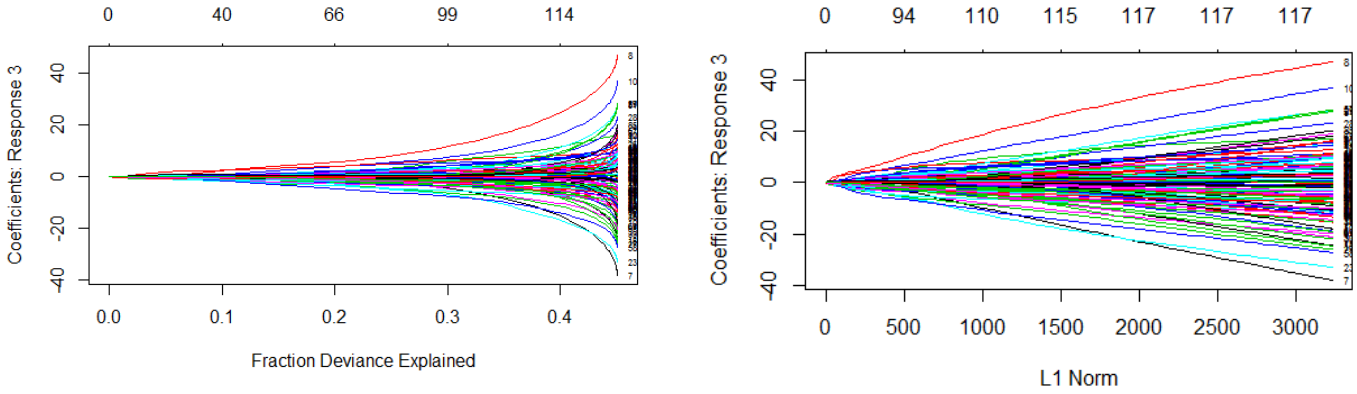


Figure 6: Specificity, Sensitivity, and Accuracy Rates of LASSO Fitted Model Prediction

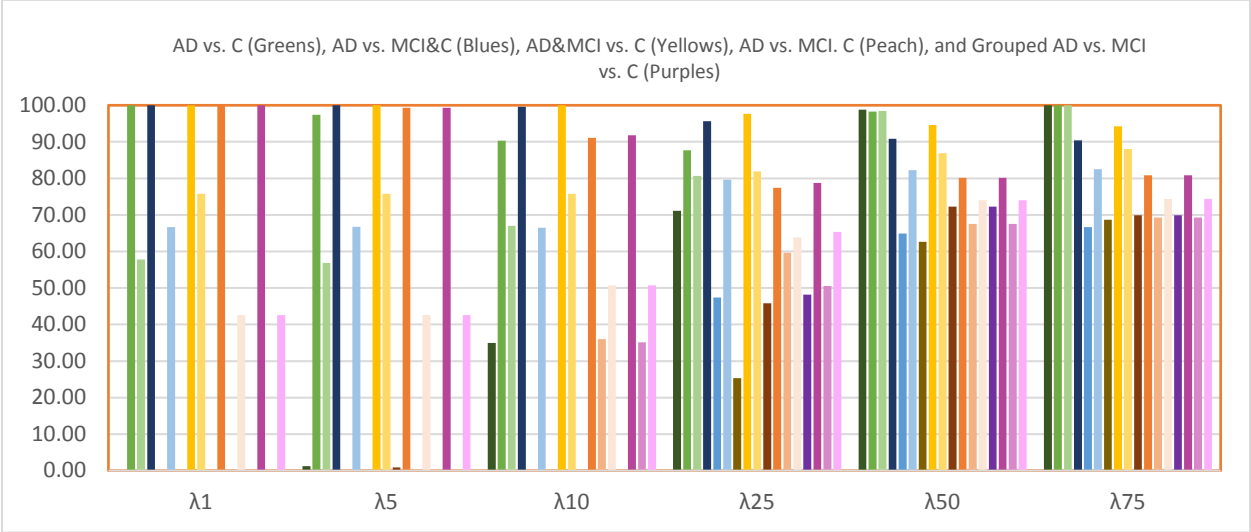


Figure 7: Cross-Validation Error (& Mean Error – Left) Curve for Binomial LASSO – AD vs. Control (λ_{min} = λ_{1se} = 0.07003043 with zero coefficients)

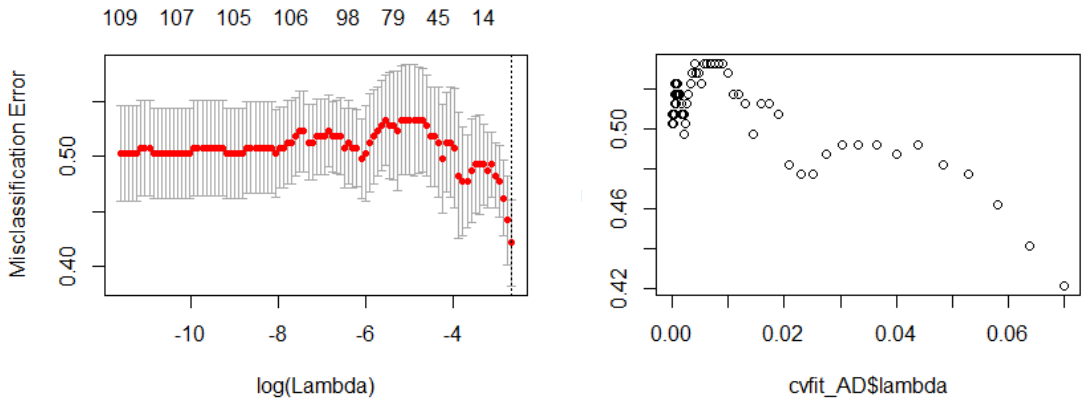


Figure 8: Cross-Validation Error (& Mean Error – Left) Curve for Binomial LASSO – AD&MCI vs. Control (λ_{min} = λ_{1se} = 0.05800169 with zero coefficients)

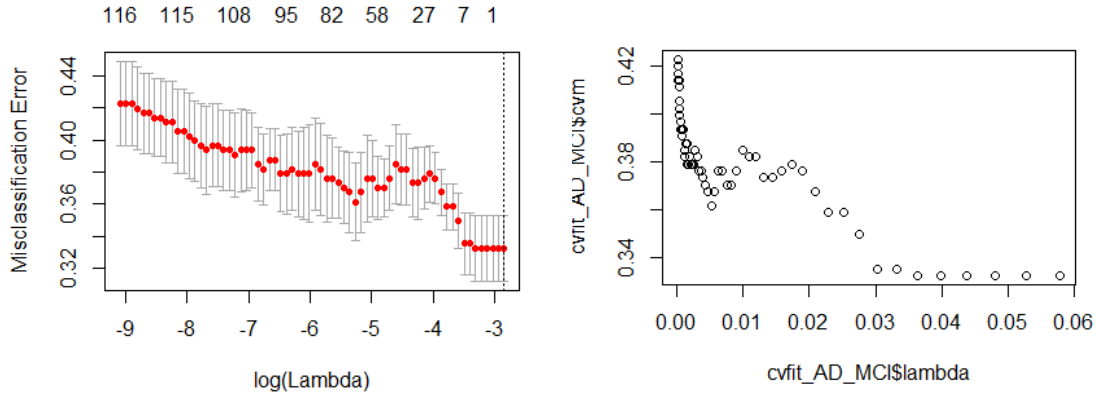


Figure 9: Cross-Validation Error (& Mean Error – Left) Curve for Binomial LASSO – AD vs. MCI&Control (λ_{min} = λ_{1se} = 0.0501518 with zero coefficients)

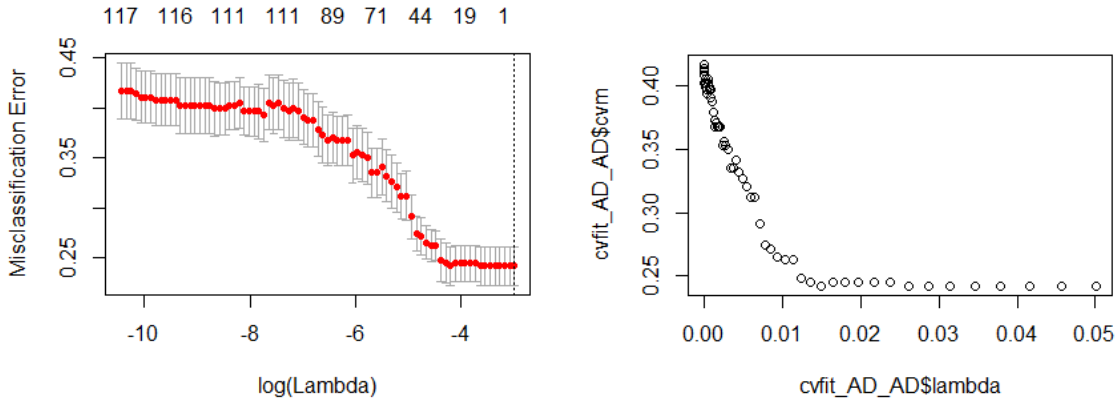


Figure 10: Cross-Validation Error (& Mean Error – Left) Curve for Multinomial LASSO – AD vs MCI vs. Control (λ_{min} = 0.01899294, λ_{1se} = 0.05800169, coefficients entering the model listed on Table 1)

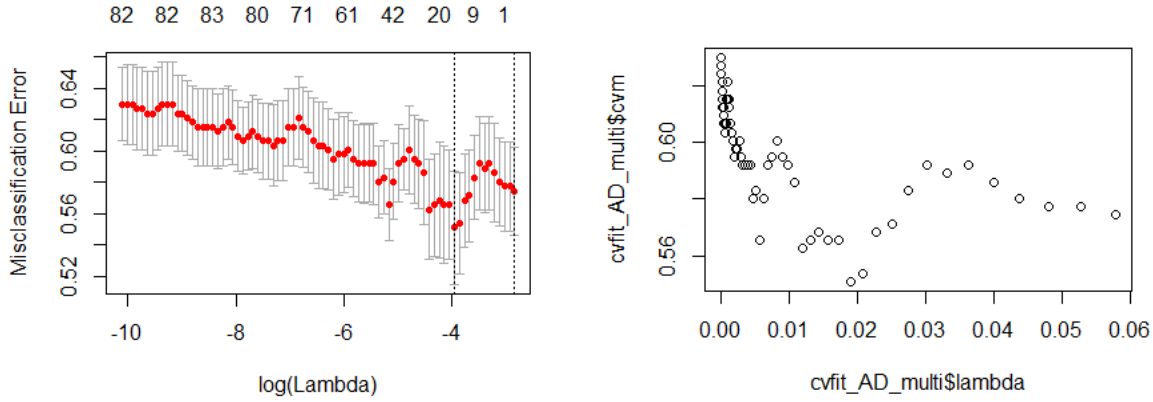


Figure 11: Cross-Validation Error (& Mean Error – Left) Curve for Grouped Multinomial LASSO – AD vs MCI vs. Control ($\lambda_{min} = \lambda_{1se}$ = 0.0715801 with no coefficients)

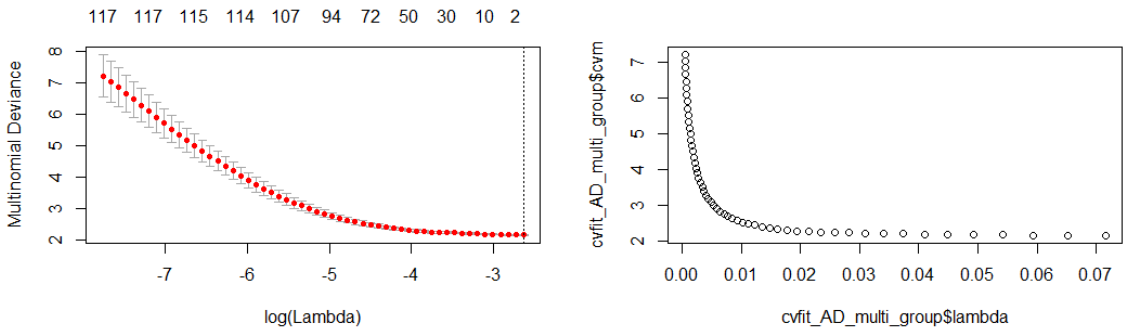
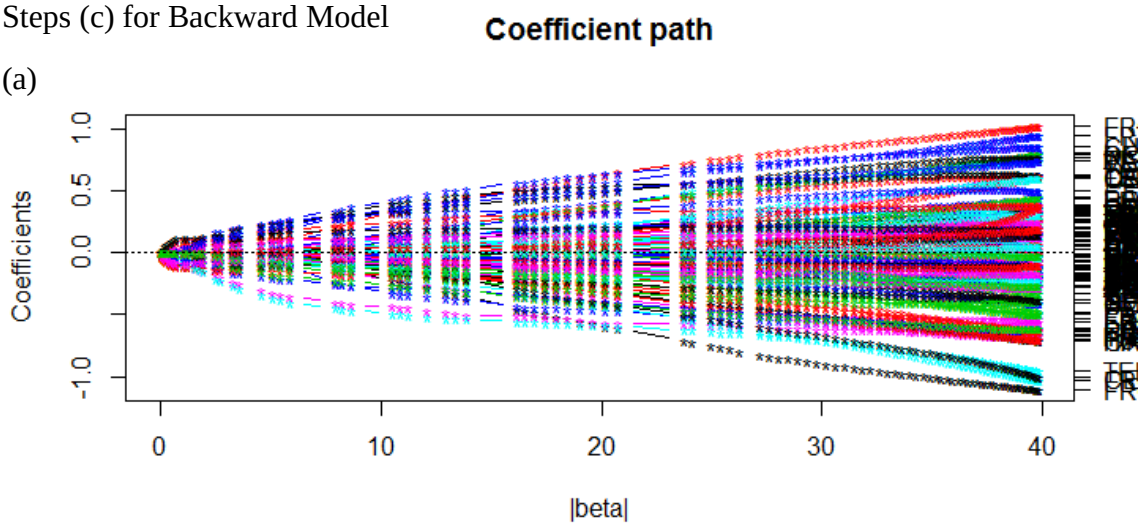
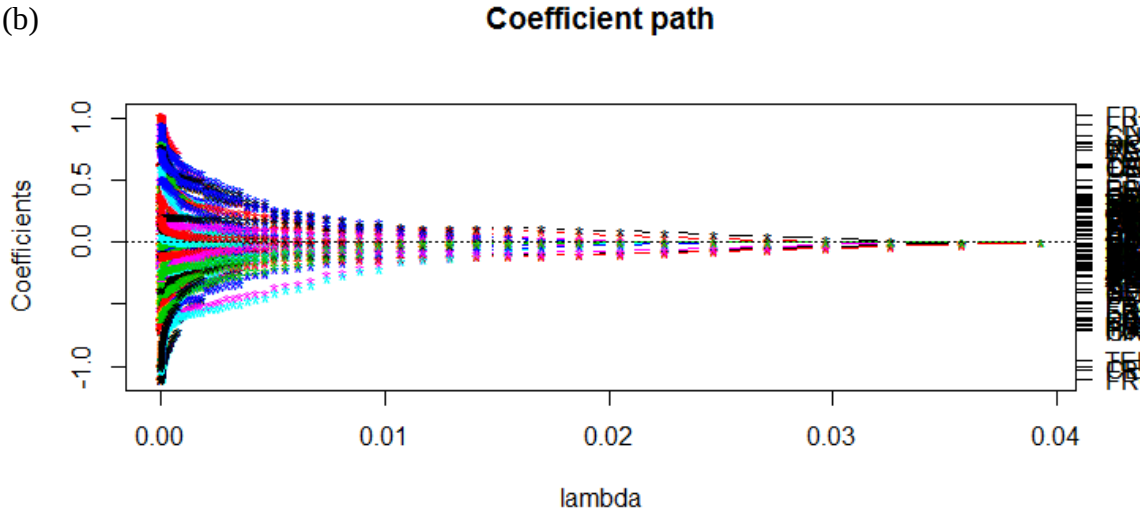


Figure 12: Ordinal LASSO Logistic Regression Coefficient Path by Betas (a) Lambda (b), and Steps (c) for Backward Model

(a)



(b)



(c)

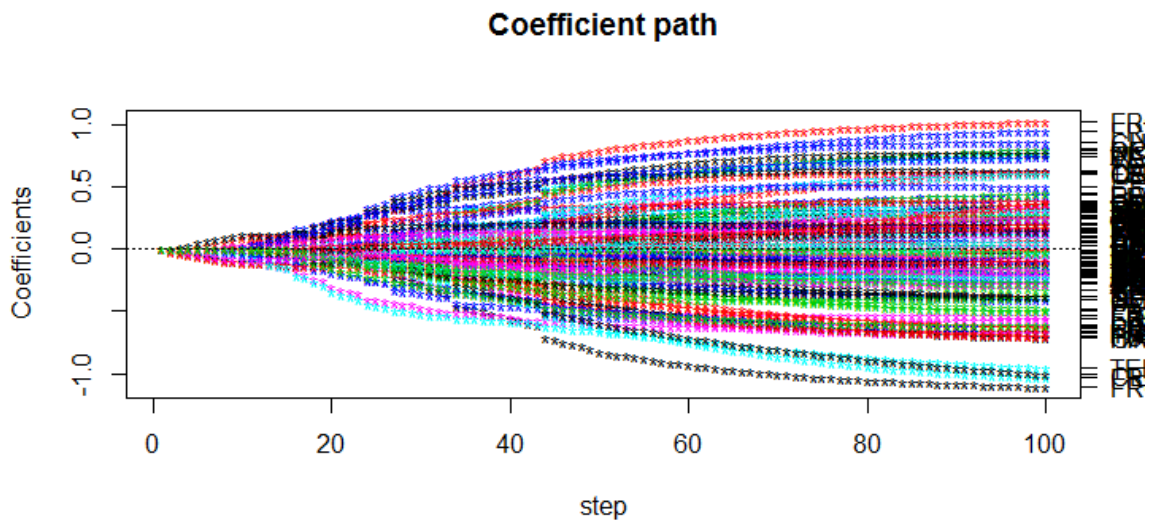
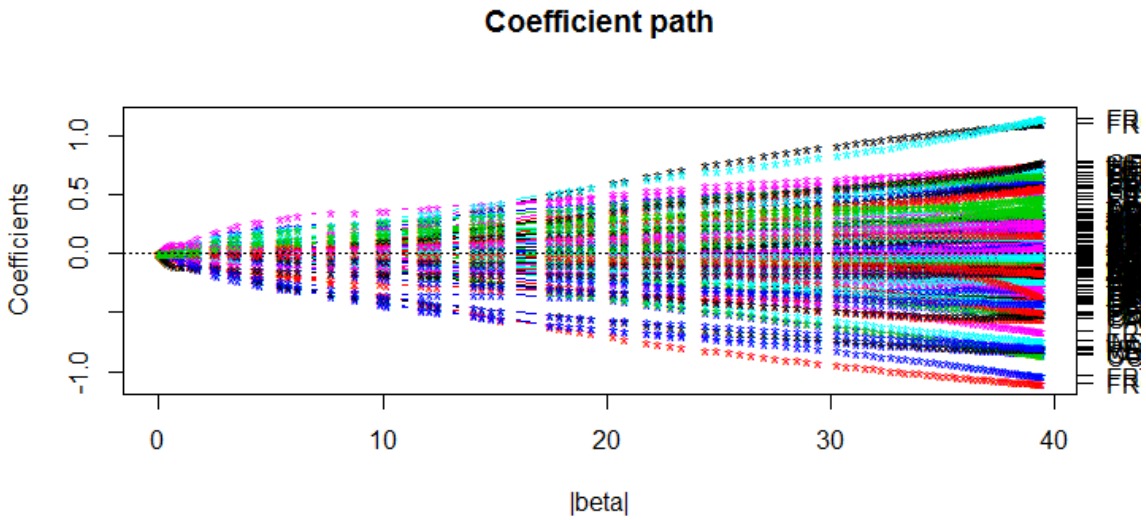
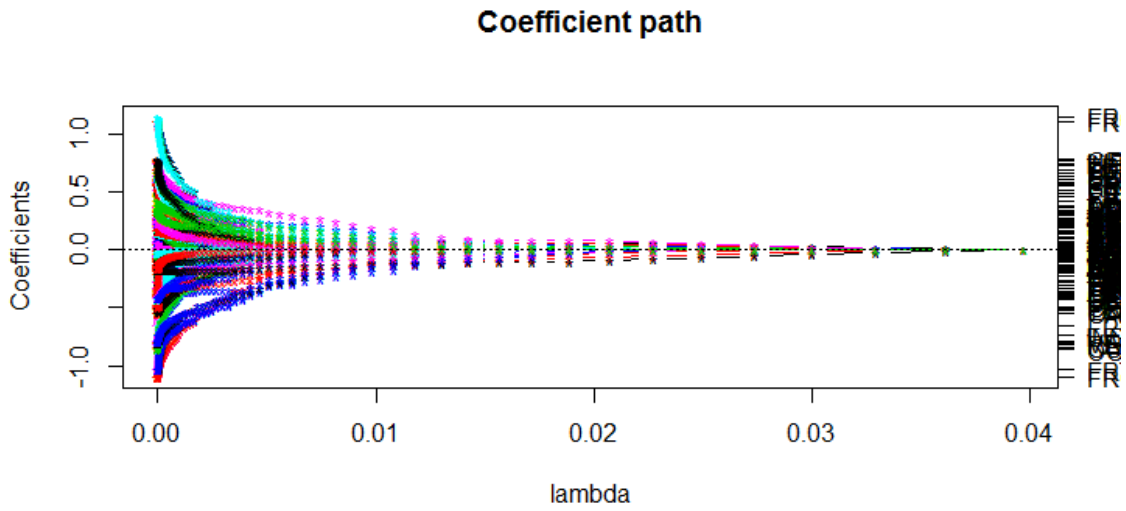


Figure 13: Ordinal LASSO Logistic Regression Coefficient Path by Betas (a) Lambda (b), and Steps (c) for Forward Model

(a)



(b)



(c)

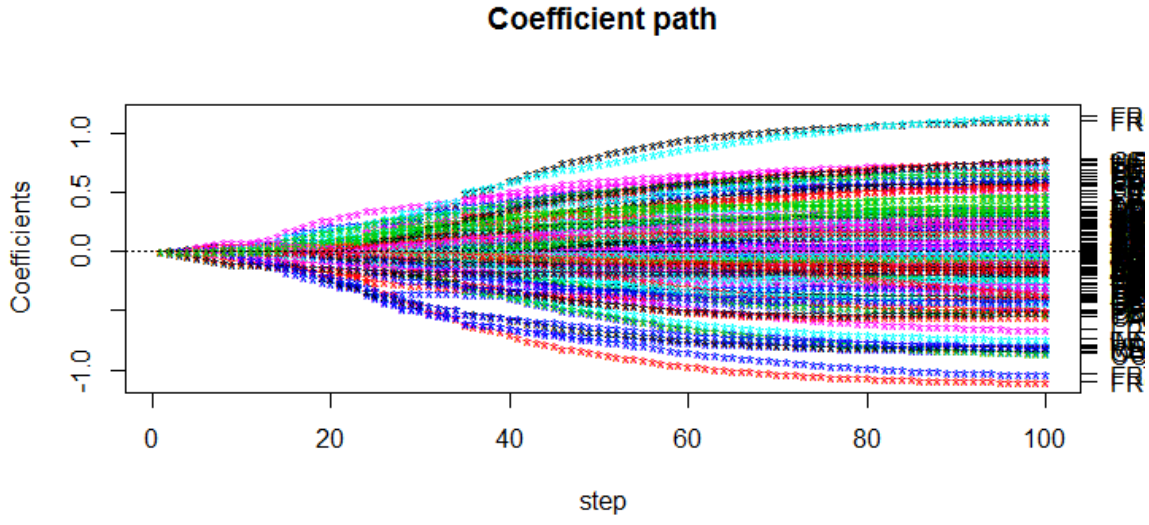
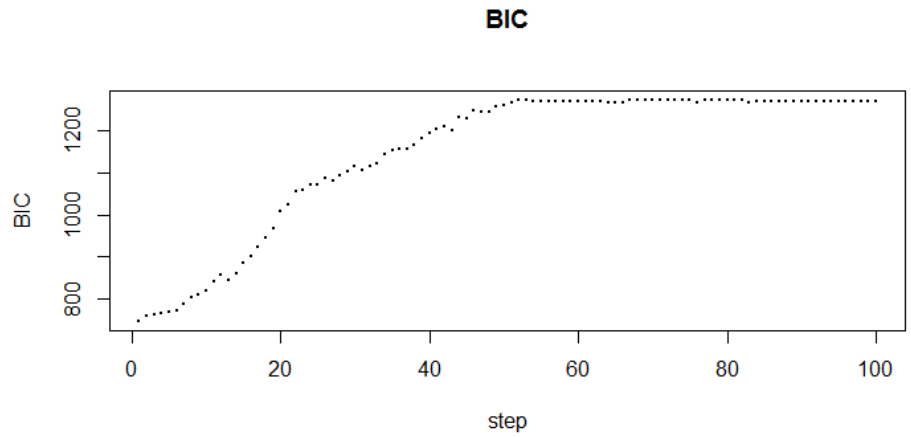


Figure 14: Ordinal LASSO Backward Model's BIC (a), and AIC (a) by Steps.

(a)



(b)

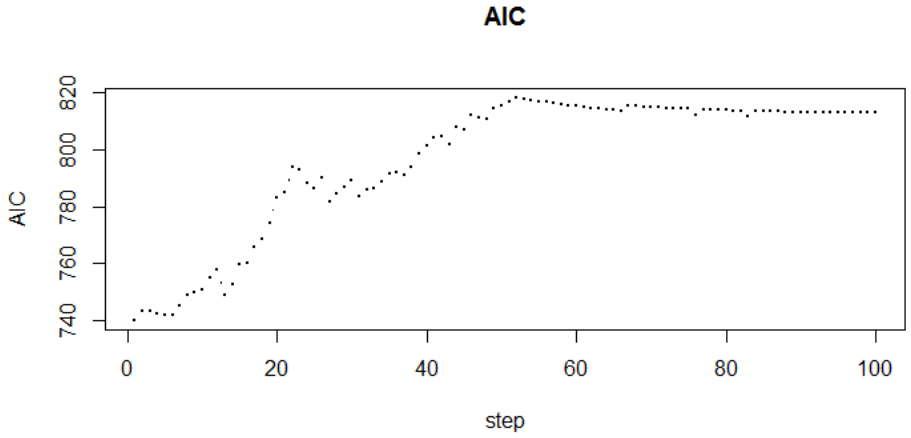
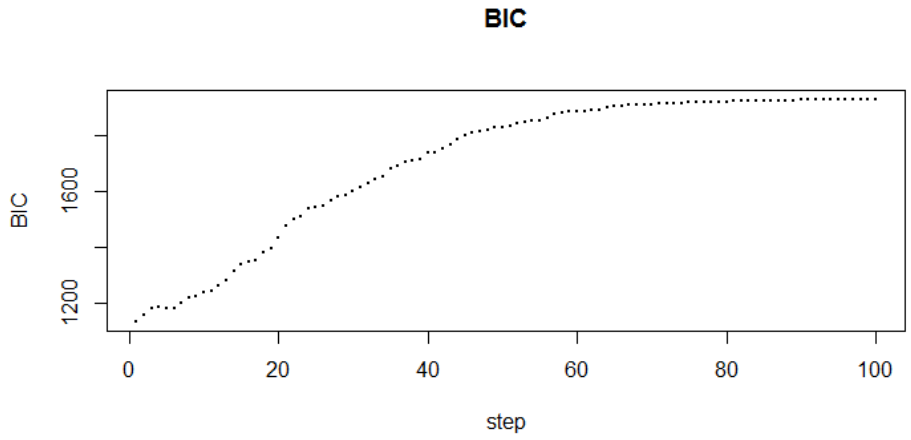


Figure 15: Ordinal LASSO Forward Model's BIC (a), and AIC (a) by Steps.

(a)



(b)

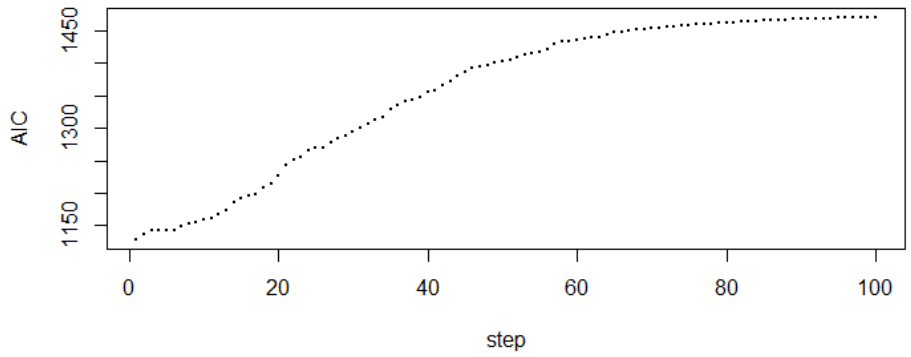
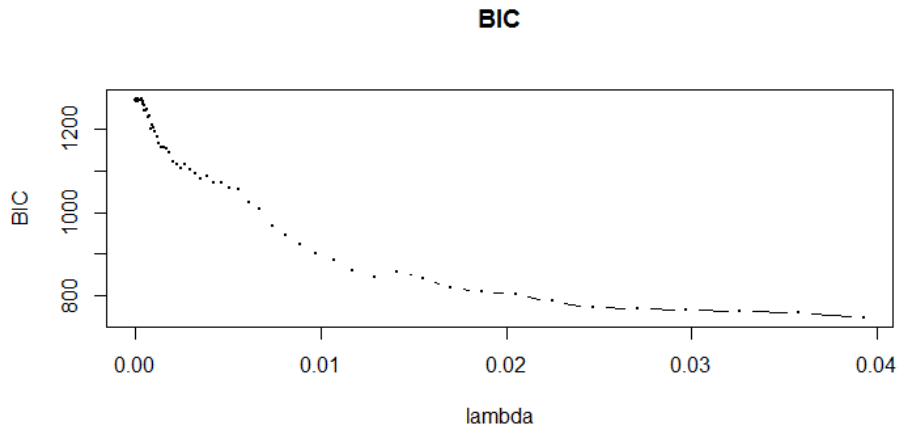


Figure 16: Ordinal LASSO Backward Model's BIC (a), and AIC (a) by Lambda Values.

(a)



(b)

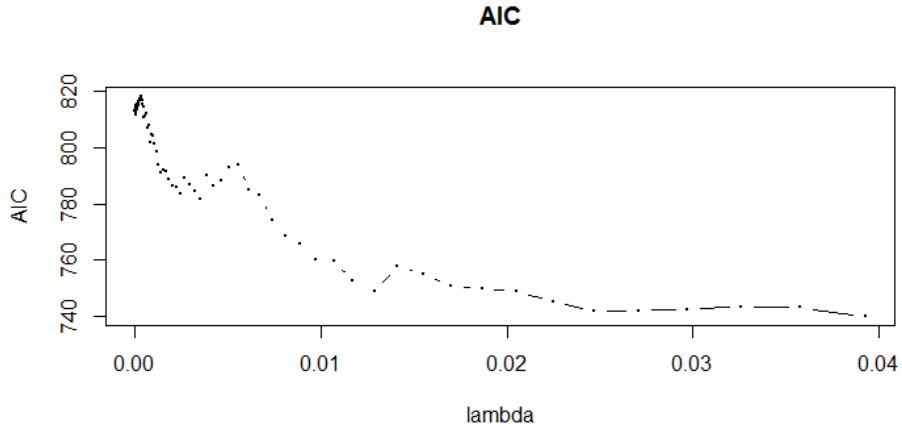
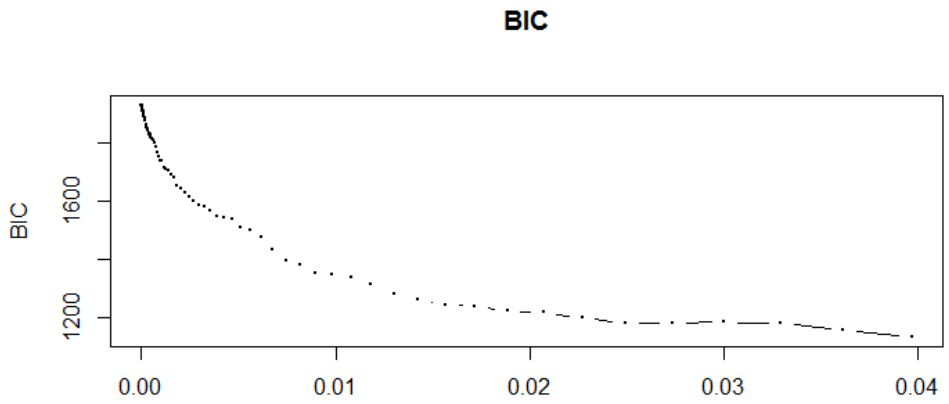


Figure 17: Ordinal LASSO Forward Model's BIC (a), and AIC (a) by Lambda Values.

(a)



(b)

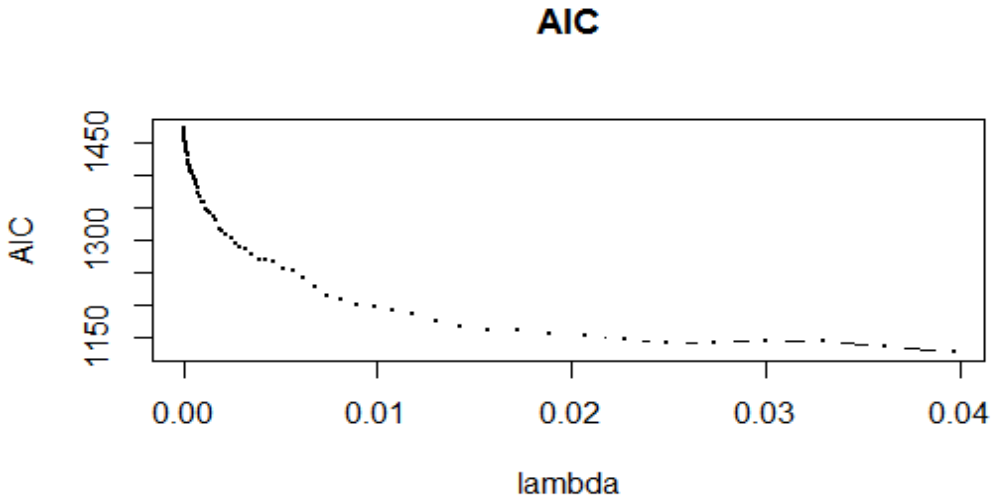
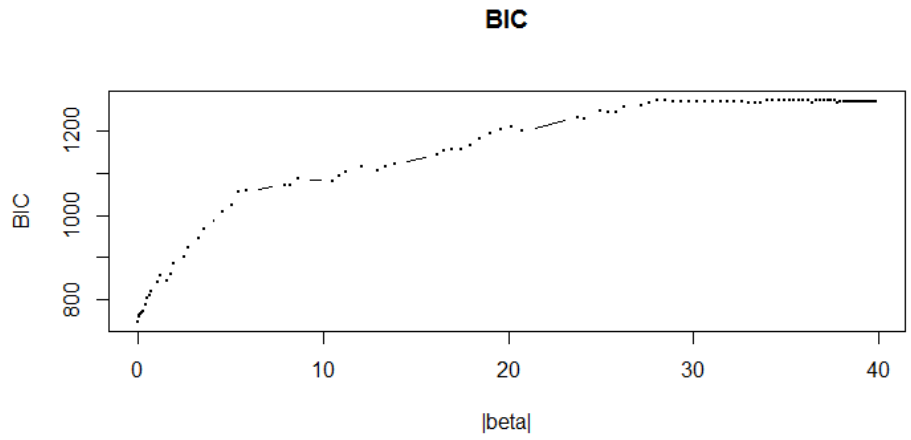


Figure 18: Ordinal LASSO Backward Model's BIC (a), and AIC (a) by Beta

(a)



(b)

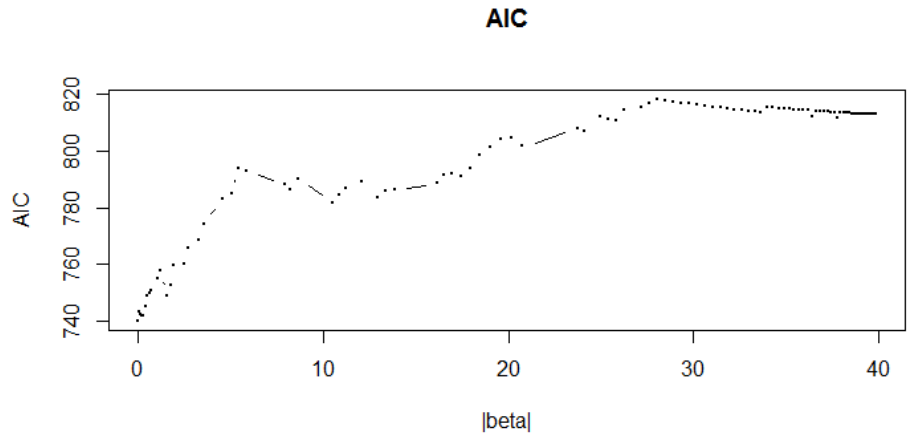
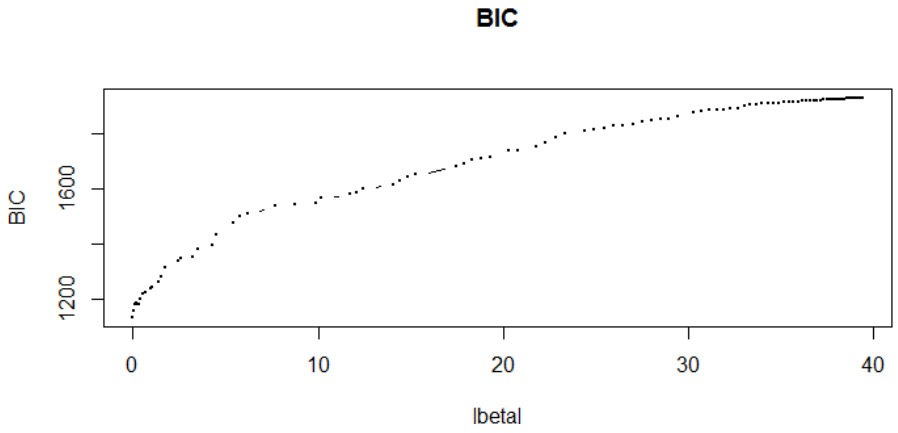


Figure 19: Ordinal LASSO Forward Model's BIC (a), and AIC (a) by Beta

(a)



(b)

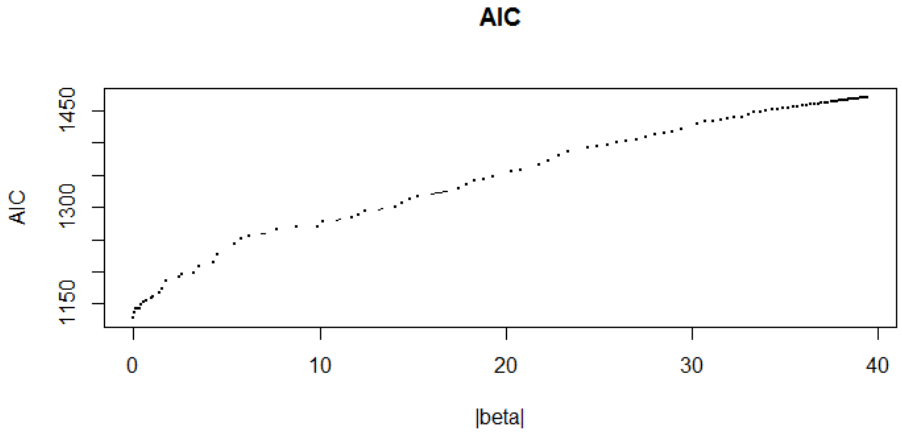


Figure 20: Sensitivity, Specificity, and Accuracy Rates for Backward and Forward Ordinal Models by Steps.