Analysis on the Cessation Effect on Smoking among Patients with MDD with the Combination Treatment of Behavioral Activation and Varenicline

William Qian

November 2024

Introduction

Major Depressive Disorder (MDD) and smoking demonstrate a significant bidirectional relationship, with MDD patients showing higher probability of heavy smoking patterns, elevated nicotine dependence levels, and intensified withdrawal symptoms during cessation attempts. While previous studies have documented varenicline's effectiveness in smoking cessation treatment, and Behavioral Activation (BA) has emerged as a promising intervention for depression-related symptoms, we still lack comprehensive understanding of their combined therapeutic effects in treating smokers with MDD. In a recent randomized, placebo-controlled, 2×2 factorial trial comparing BA for smoking cessation (BASC) and standard behavioral treatment (ST), with and without varenicline administration, researchers found no significant advantage of BASC over ST in adult smokers with current or past MDD diagnosis.

In this report, we aim to extend the previous findings by analyzing the potential moderating effects of baseline variables on treatment outcomes and examining the predictive factors for smoking abstinence, while controlling for both behavioral intervention approaches and pharmacotherapy. We investigate multiple potential moderators, including depression-related variables (e.g., anhedonia severity, current MDD status, depression intensity), smoking-related characteristics (e.g., nicotine dependence level, daily cigarette consumption, cessation readiness), and demographic factors. Through this analysis, we expect to identify specific patient characteristics that may predict better responses to particular treatment approaches, ultimately contributing to the development of more individualized and effective smoking cessation strategies for this high-risk population.

Data

The dataset used in this analysis contains 300 observations with 25 variables (including 1 identification column) collected from a randomized, placebo-controlled, 2×2 factorial trial involving adult smokers diagnosed with current or past major depressive disorder (MDD). In order to ensure data quality and reliability of our analysis, we firstly conducted data preprocessing to handle missing values. After removing observations with incomplete records, our final analytical sample consisted of 241 participants.

The data documented participants' demographic information, depression-related variables, and smoking-related characteristics measured at baseline, as well as their treatment assignment and outcomes. The participants were randomized to receive either Behavioral Activation for smoking cessation (BASC) or standard behavioral treatment (ST), combined with either varenicline or placebo, forming a 2×2 factorial design. The completion of treatment and follow-up assessments were carefully monitored throughout the trial period.

Variable	Description
abst	Smoking Abstinence
Var	Pharmacotherapy (Varenicline)
BA	Psychotherapy (Behavioral Activation)
age_ps	Age at phone interview
sex_ps	Sex at phone interview
NHW	Non-Hispanic White indicator
Black	Black indicator
Hisp	Hispanic indicator
inc	Income (ordinal categorical, low to high)
edu	Education (ordinal categorical, low to high)
ftcd_score	FTCD score at baseline
ftcd.5.mins	Smoking within 5 mins of waking up
bdi_score_pq1	BDI score at baseline (a measure of depression)
cpd_ps	Cigarettes per day at baseline phone survey
crv_total_pq1	Cigarette reward value at baseline
hedonsum_n_pq1	Pleasurable Events Scale at baseline – substitute reinforcers
hedonsum_y_pq1	Pleasurable Events Scale at baseline – complementary reinforcers
shaps_score_pq1	Anhedonia
otherdiag	Other lifetime DSM-5 diagnosis
antidepmed	Taking antidepressant medication at baseline
mde_curr	Current vs past MDD
NMR	Nicotine Metabolism Ratio
Only.Menthol	Exclusive Mentholated Cigarette User
readiness	Baseline readiness to quit smoking

Analysis of the baseline characteristics by treatment groups reveals several distinctive patterns that need attention. When comparing the varenicline and placebo groups, we observe a substantial disparity in abstinence rates, with the varenicline group achieving 28% success compared to 8.4% in the placebo group ($p < 0.001$), indicating varenicline's potential therapeutic effectiveness. However, other baseline characteristics maintained balanced distributions between these pharmacotherapy groups, suggesting successful randomization. Similarly, when comparing behavioral activation (BA) and standard treatment groups, most characteristics were comparably distributed, with one notable exception: the proportion of participants using antidepressant medication differed significantly between groups (35% in BA group versus 19% in standard treatment group, $p = 0.004$).

Table 2: Data Summary Table

	Grouped by Pharmacotherapy Status			**Grouped by Psychotherapy Status**		
Characteristic	**No Varenicline**	**Varenicline**	**p-value**	**No Psychotherapy**	**Psychotherapy**	**p-value**
abst			<0.001			0.7
0	98 (92%)	96 (72%)		93 (79%)	101 (81%)	
1	9 (8.4%)	38 (28%)		24 (21%)	23 (19%)	
BA			0.6			
0	54 (50%)	63 (47%)				
1	53 (50%)	71 (53%)				
age_ps	52 (45, 59)	52 (40, 59)	0.4	51 (43, 58)	54 (41, 60)	0.2
sex_ps			0.4			>0.9
1	45 (42%)	63 (47%)		52 (44%)	56 (45%)	
2	62 (58%)	71 (53%)		65 (56%)	68 (55%)	
NHW			0.7			0.5
0	70 (65%)	85 (63%)		78 (67%)	77 (62%)	
1	37 (35%)	49 (37%)		39 (33%)	47 (38%)	
Black			0.2			0.6

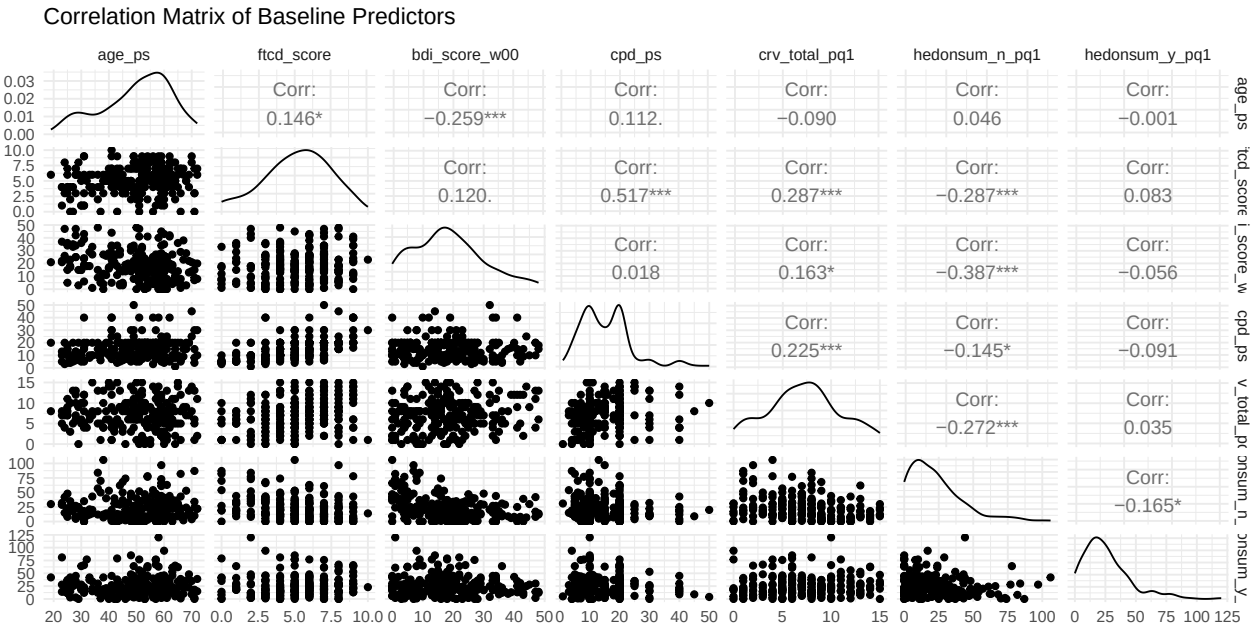
Table 2: Data Summary Table *(continued)*

	Grouped by Pharmacotherapy Status			Grouped by Psychotherapy Status		
Characteristic	**No Varenicline**	**Varenicline**	**p-value**	**No Psychotherapy**	**Psychotherapy**	**p-value**
0	47 (44%)	70 (52%)		55 (47%)	62 (50%)	
1	60 (56%)	64 (48%)		62 (53%)	62 (50%)	
Hisp			0.7			0.5
0	100 (93%)	127 (95%)		109 (93%)	118 (95%)	
1	7 (6.5%)	7 (5.2%)		8 (6.8%)	6 (4.8%)	
inc			0.6			0.3
1	38 (36%)	47 (35%)		40 (34%)	45 (36%)	
2	25 (23%)	31 (23%)		28 (24%)	28 (23%)	
3	18 (17%)	20 (15%)		21 (18%)	17 (14%)	
4	16 (15%)	15 (11%)		10 (8.5%)	21 (17%)	
5	10 (9.3%)	21 (16%)		18 (15%)	13 (10%)	
edu			0.2			0.6
1	1 (0.9%)	0 (0%)		0 (0%)	1 (0.8%)	
2	3 (2.8%)	9 (6.7%)		4 (3.4%)	8 (6.5%)	
3	29 (27%)	34 (25%)		30 (26%)	33 (27%)	
4	46 (43%)	45 (34%)		48 (41%)	43 (35%)	
5	28 (26%)	46 (34%)		35 (30%)	39 (31%)	
ftcd_score	6 (4, 7)	5 (4, 7)	0.3	6 (4, 7)	5 (4, 7)	0.4
ftcd.5.mins	51 (48%)	57 (43%)	0.4	57 (49%)	51 (41%)	0.2
bdi_score_w00	18 (9, 27)	18 (9, 24)	0.9	18 (12, 25)	18 (9, 26)	0.6
cpd_ps	15 (10, 20)	15 (10, 20)	0.5	15 (10, 20)	15 (10, 20)	>0.9
crv_total_pq1	7 (5, 9)	7 (5, 10)	0.8	7 (5, 9)	8 (5, 10)	0.4
hedonsum_n_pq1	17 (9, 30)	20 (9, 35)	0.3	18 (9, 32)	20 (9, 31)	>0.9
hedonsum_y_pq1	24 (12, 36)	19 (13, 31)	0.2	21 (13, 35)	22 (12, 33)	>0.9
shaps_score_pq1	1 (0, 4)	1 (0, 3)	>0.9	1 (0, 4)	1 (0, 3)	0.5
otherdiag			0.6			0.7
0	58 (54%)	77 (57%)		64 (55%)	71 (57%)	
1	49 (46%)	57 (43%)		53 (45%)	53 (43%)	
antidepmed			0.3			0.004
0	74 (69%)	101 (75%)		95 (81%)	80 (65%)	
1	33 (31%)	33 (25%)		22 (19%)	44 (35%)	
mde_curr			0.6			0.8
0	59 (55%)	69 (51%)		61 (52%)	67 (54%)	
1	48 (45%)	65 (49%)		56 (48%)	57 (46%)	
NMR	0.32 (0.20, 0.45)	0.31 (0.21, 0.50)	0.8	0.32 (0.20, 0.46)	0.32 (0.22, 0.48)	0.7
Only.Menthol			0.7			>0.9
0	41 (38%)	55 (41%)		47 (40%)	49 (40%)	
1	66 (62%)	79 (59%)		70 (60%)	75 (60%)	
readiness	7.00 (6.00, 8.00)	7.00 (6.00, 8.00)	0.2	7.00 (6.00, 8.00)	7.00 (6.00, 8.00)	0.9
Var						0.6
0				54 (46%)	53 (43%)	
1				63 (54%)	71 (57%)	

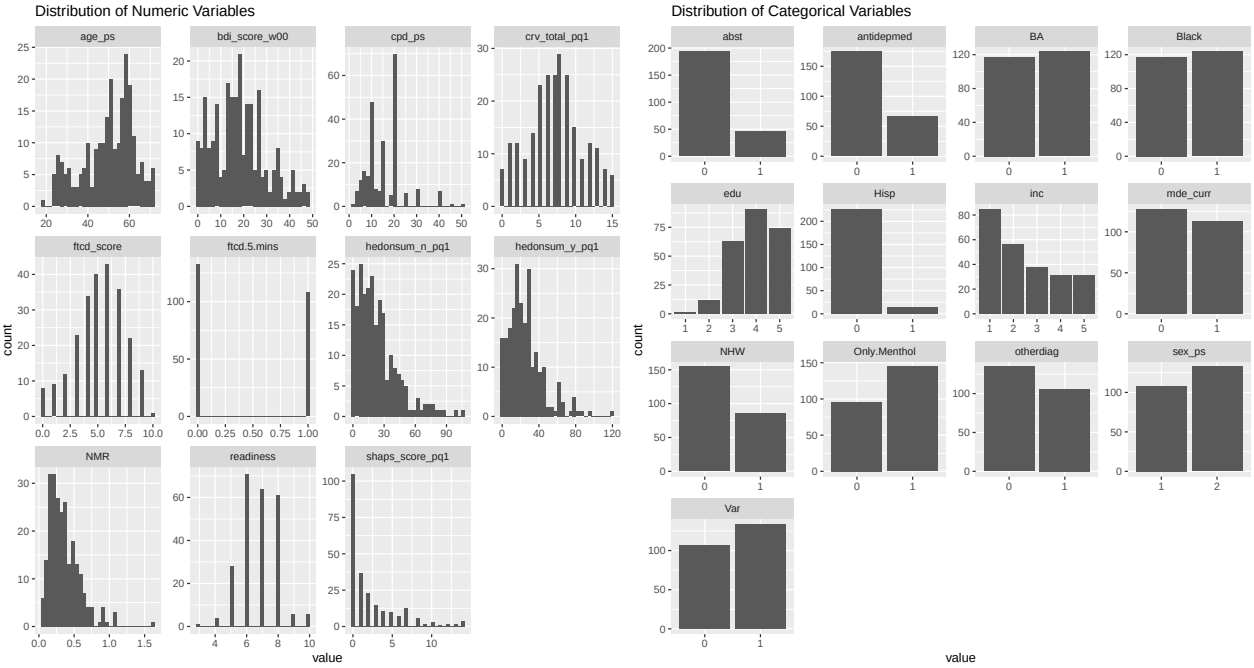
[1] n (%); Median (IQR)

[2] Pearson's Chi-squared test; Wilcoxon rank sum test; Fisher's exact test

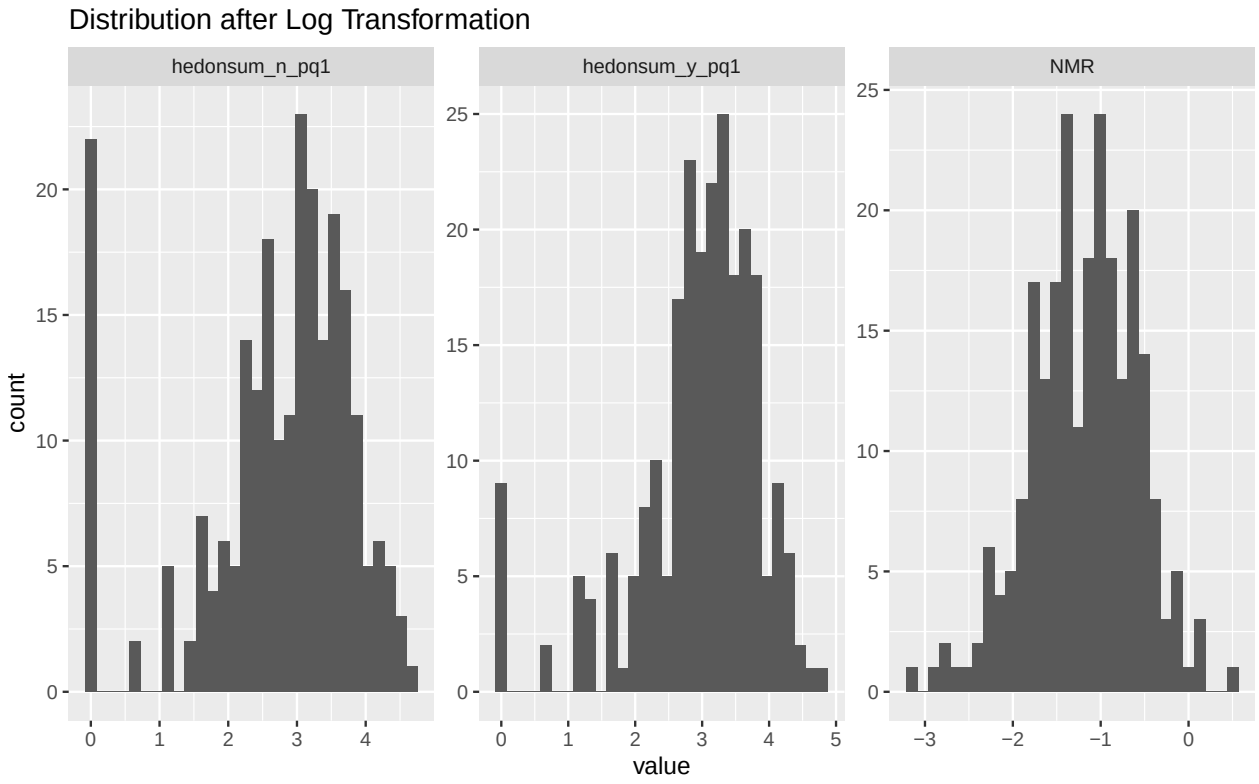
In our examination of the relationships among continuous variables in the dataset, we identified several significant correlations that deserve attention. The daily cigarette consumption (cpd_ps) exhibited a moderate positive correlation with the Fagerström Test for Cigarette Dependence score (FTCD score) ($r = 0.517$, $p < 0.001$). This correlation aligns with theoretical expectations, as both measures are designed to assess different aspects of smoking dependency. Furthermore, we observed that the Beck Depression Inventory score (bdi_score_w00) was negatively correlated with both measures of hedonic capacity (hedonsum_n_pq1 and hedonsum_y_pq1), a finding that supports the established relationship between increased depression severity and diminished ability to experience pleasure.



Our distribution analysis of numeric variables in the dataset identified notable right-skewed patterns in several key measures, particularly in both hedonic capacity indices (hedonsum_n_pq1, hedonsum_y_pq1) and the nicotine metabolite ratio (NMR). To address these violations of normality assumptions, we implemented log transformations for the affected variables. Post-transformation examination revealed substantially improved symmetry and reduced skewness in these distributions, ensuring their suitability for subsequent statistical analyses. We found that the remaining numeric variables, including participant age, Fagerström Test for Cigarette Dependence score, and depression-related measures, demonstrated approximately normal distributions and were therefore retained in their original scale for analysis.



This is the distribution of hedonsum_n_pq1, hedonsum_y_pq1, and NMR after log transformation.



Methods

In this study, we employed three regularized regression approaches - LASSO (L1), Ridge (L2), and Elastic Net regularization - to analyze the moderators and predictors of smoking cessation outcomes.

Given our binary outcome of smoking abstinence (y), we used logistic regression with different regularization terms. The basic form of our logistic regression model is:

$$P(y = 1|X) = \frac{1}{1 + e^{-X\beta}}$$

where X represents our predictor variables and β represents the coefficients.

LASSO (L1) Regularization

LASSO (Least Absolute Shrinkage and Selection Operator) adds an L1 penalty term to the log-likelihood function:

$$\hat{\beta}_{LASSO} = \underset{\beta}{\text{argmin}}\{-l(\beta) + \lambda \sum_{j=1}^{p} |\beta_j|\}$$

where $l(\beta)$ is the log-likelihood and λ is the regularization parameter. The L1 penalty encourages sparsity by shrinking some coefficients exactly to zero, effectively performing variable selection.

Ridge (L2) Regularization

Ridge regression uses an L2 penalty term:

$$\hat{\beta}_{Ridge} = \underset{\beta}{\operatorname{argmin}}\{-l(\beta) + \lambda \sum_{j=1}^{p} \beta_j^2\}$$

The L2 penalty shrinks coefficients toward zero but rarely sets them exactly to zero. This approach is particularly useful when dealing with correlated predictors.

Elastic Net

Elastic Net combines both L1 and L2 penalties:

$$\hat{\beta}_{ElasticNet} = \underset{\beta}{\operatorname{argmin}}\{-l(\beta) + \lambda(\alpha \sum_{j=1}^{p} |\beta_j| + (1-\alpha) \sum_{j=1}^{p} \beta_j^2)\}$$

where α controls the mix between L1 and L2 penalties ($\alpha = 1$ gives LASSO, $\alpha = 0$ gives Ridge). This combination allows for both variable selection and handling of correlated predictors.

Model Implementation

We implemented these models using the glmnet package in R, with cross-validation to select the optimal λ value. For the Elastic Net model, we set $\alpha = 0.5$ to give equal weight to L1 and L2 penalties. The models were evaluated using AUC and accuracy metrics on both training and test sets, with the data split in a 50:50 ratio to ensure robust validation of our findings.

Regresssion Analysis

Basic Models

We first implemented two fundamental regularized regression models (Ridge and LASSO regression without interaction terms) to examine the relationship between smoking abstinence and our predictors.

Lasso Regression

The LASSO regression with $\lambda = 0.033$ can be expressed as:

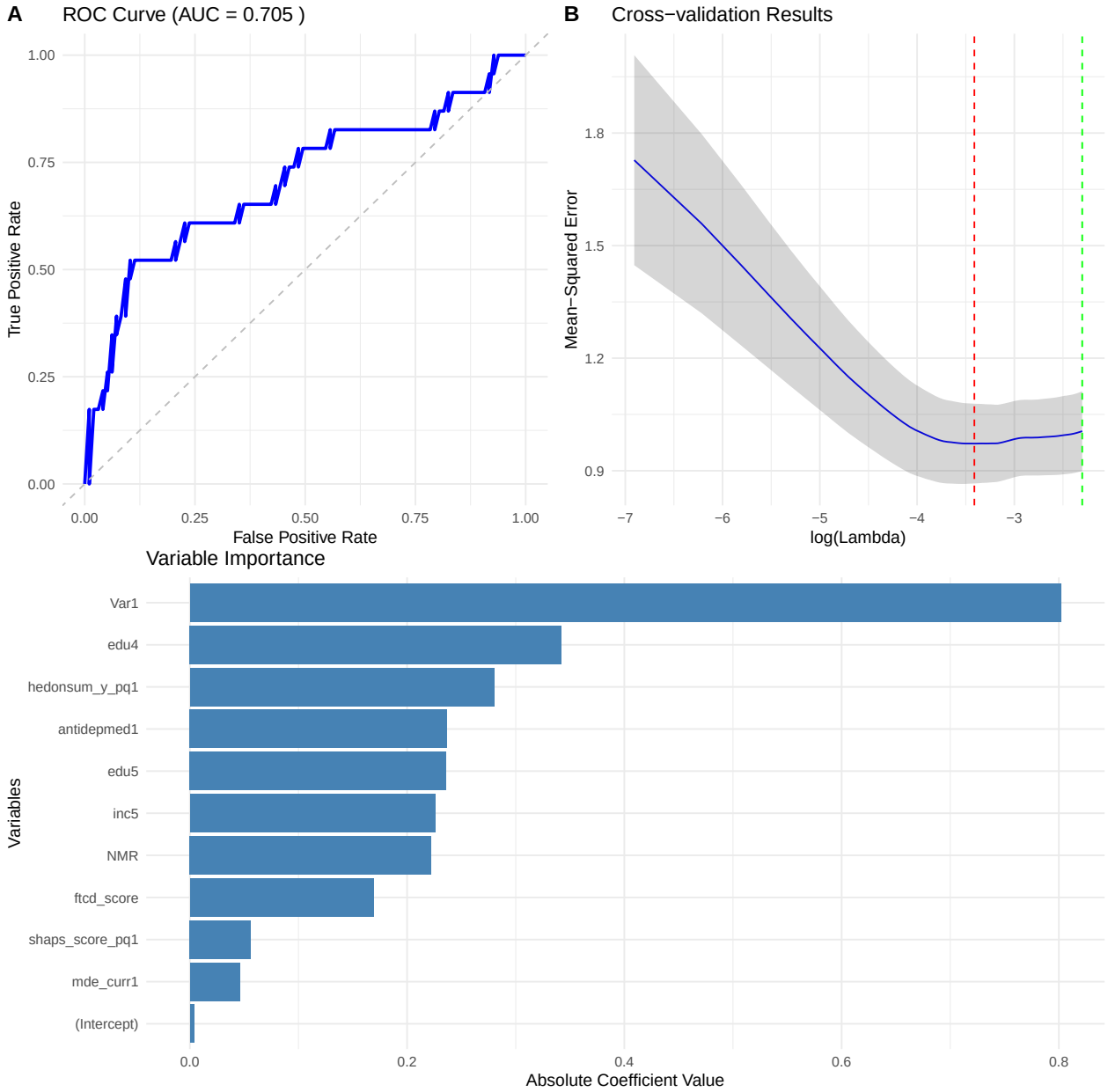
$$\begin{aligned} logit(P(abst = 1)) &= \beta_0 + \beta_1 \times Var + \beta_2 \times inc5 + \beta_3 \times edu4 + \beta_4 \times edu5 \\ &\quad + \beta_5 \times ftcd\_score + \beta_6 \times hedonsum\_y\_pq1 + \beta_7 \times shaps\_score\_pq1 \\ &\quad + \beta_8 \times mde\_curr1 + \beta_9 \times antidepmed1 + \beta_{10} \times log(NMR) \\ &= 0.004 + 0.802 \times Var + 0.226 \times inc5 - 0.342 \times edu4 + 0.236 \times edu5 \\ &\quad - 0.169 \times ftcd\_score - 0.280 \times hedonsum\_y\_pq1 - 0.056 \times shaps\_score\_pq1 \\ &\quad - 0.046 \times mde\_curr1 + 0.237 \times antidepmed1 + 0.222 \times log(NMR) \end{aligned}$$

Our LASSO regression model demonstrated strong predictive performance, achieving an Area Under the Curve (AUC) of 0.841 on the training dataset and 0.705 on the test dataset. Through its inherent variable selection mechanism, the LASSO model identified 10 key predictors of smoking cessation success.

Among these predictors, varenicline treatment emerged as the most influential factor, showing the strongest positive effect (coefficient = 0.802) on abstinence probability. This coefficient indicates that varenicline administration increases the log-odds of abstinence by 0.802, which translates to an odds ratio of exp(0.802) = 2.23. In the social domain, we found that both high income level (inc5: coefficient = 0.226, odds ratio =

exp(0.226) = 1.25) and antidepressant medication use (coefficient = 0.237, odds ratio = exp(0.237) = 1.27) were associated with increased likelihood of abstinence.

The impact of education levels showed heterogeneous effects in our analysis. While edu4 demonstrated a negative association with abstinence (coefficient = -0.342, reducing odds by 29% as exp(-0.342) = 0.71), edu5 showed a positive relationship (coefficient = 0.236, increasing odds by 27% as exp(0.236) = 1.27). Regarding clinical measures, we observed moderate negative associations with abstinence: each unit increase in Fagerström Test score (ftcd_score) was associated with a 16% reduction in abstinence odds (coefficient = -0.169, odds ratio = exp(-0.169) = 0.84), and similarly, each unit increase in positive hedonic capacity (hedonsum_y_pq1) corresponded to a 24% decrease in abstinence odds (coefficient = -0.280, odds ratio = exp(-0.280) = 0.76).



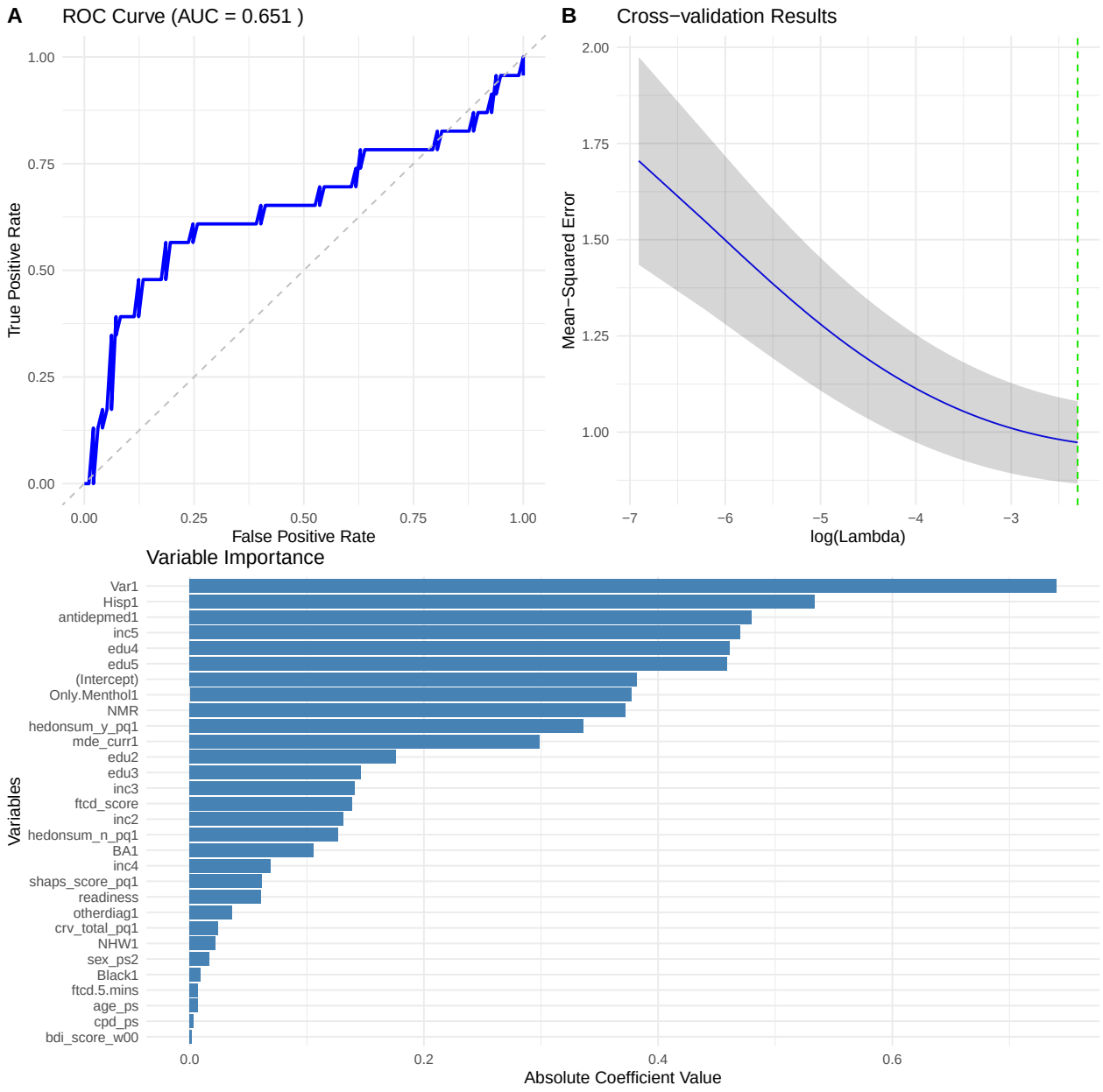
Ridge Regression

The Ridge regression with $\lambda = 0.100$ can be expressed as:

$$
\begin{aligned}
logit(P(abst = 1)) &= \beta_0 + \beta_1 \times Var + \beta_2 \times BA + \beta_3 \times age\_ps + \beta_4 \times sex\_ps2 \\
&+ \beta_5 \times NHW1 + \beta_6 \times Black1 + \beta_7 \times Hisp1 \\
&+ \beta_8 \times inc2 + \beta_9 \times inc3 + \beta_{10} \times inc4 + \beta_{11} \times inc5 \\
&+ \beta_{12} \times edu2 + \beta_{13} \times edu3 + \beta_{14} \times edu4 + \beta_{15} \times edu5 \\
&+ \beta_{16} \times ftcd\_score + \beta_{17} \times ftcd.5.mins + \beta_{18} \times bdi\_score\_w00 \\
&+ \beta_{19} \times cpd\_ps + \beta_{20} \times crv\_total\_pq1 + \beta_{21} \times log(hedonsum\_n\_pq1) \\
&+ \beta_{22} \times log(hedonsum\_y\_pq1) + \beta_{23} \times shaps\_score\_pq1 + \beta_{24} \times otherdiag1 \\
&+ \beta_{25} \times antidepmed1 + \beta_{26} \times mde\_curr1 + \beta_{27} \times log(NMR) \\
&+ \beta_{28} \times Only.Menthol1 + \beta_{29} \times readiness \\
&= -0.382 + 0.740 \times Var - 0.106 \times BA + 0.007 \times age\_ps + 0.017 \times sex\_ps2 \\
&+ 0.022 \times NHW1 + 0.009 \times Black1 + 0.534 \times Hisp1 \\
&+ 0.131 \times inc2 + 0.141 \times inc3 - 0.069 \times inc4 + 0.470 \times inc5 \\
&- 0.176 \times edu2 + 0.146 \times edu3 - 0.461 \times edu4 + 0.459 \times edu5 \\
&- 0.139 \times ftcd\_score + 0.007 \times ftcd.5.mins + 0.002 \times bdi\_score\_w00 \\
&+ 0.003 \times cpd\_ps - 0.024 \times crv\_total\_pq1 + 0.127 \times log(hedonsum\_n\_pq1) \\
&- 0.336 \times log(hedonsum\_y\_pq1) - 0.062 \times shaps\_score\_pq1 + 0.036 \times otherdiag1 \\
&+ 0.480 \times antidepmed1 - 0.299 \times mde\_curr1 + 0.372 \times log(NMR) \\
&+ 0.377 \times Only.Menthol1 - 0.061 \times readiness
\end{aligned}
$$

The Ridge regression model demonstrated stronger performance on the training set with an AUC of 0.878, though its test set performance (AUC = 0.651) was slightly lower than the LASSO model. While Ridge regression retains all variables in the model, the variable importance analysis revealed patterns similar to our LASSO findings, with varenicline treatment emerging as one of the strongest positive predictors (coefficient = 0.740, odds ratio = exp(0.740) = 2.10). Notably, both antidepressant medication use (coefficient = 0.480, odds ratio = exp(0.480) = 1.62) and Hispanic ethnicity (coefficient = 0.534, odds ratio = exp(0.534) = 1.71) showed substantial positive associations with abstinence outcomes.

Among the clinical predictors, we observed moderate positive associations for both nicotine metabolism ratio (NMR: coefficient = 0.372, increasing odds by 45% as exp(0.372) = 1.45) and menthol cigarette use (coefficient = 0.377, increasing odds by 46% as exp(0.377) = 1.46). In contrast, depression-related measures demonstrated consistent negative associations with abstinence probability. Specifically, current MDD status decreased the odds of abstinence by 26% (coefficient = -0.299, odds ratio = exp(-0.299) = 0.74), and higher scores on the positive hedonic capacity measure (hedonsum_y_pq1) were associated with reduced odds of abstinence by 28% (coefficient = -0.336, odds ratio = exp(-0.336) = 0.72).



The comparison of the Ridge model and the Lasso model is shown below. The table of coefficients are also included.

Table 3: Comparison of Basic Ridge and Lasso Regression Models

Variable	Ridge Regression Model	Lasso Regression Model
(Intercept)	-0.382	0.004
Var1	0.740	0.802
BA1	-0.106	Eliminated
age_ps	0.007	Eliminated
sex_ps2	0.017	Eliminated
NHW1	0.022	Eliminated
Black1	0.009	Eliminated
Hisp1	0.534	Eliminated
inc2	0.131	Eliminated
inc3	0.141	Eliminated

Table 3: Comparison of Basic Ridge and Lasso Regression Models *(continued)*

Variable	Ridge Regression Model	Lasso Regression Model
inc4	-0.069	Eliminated
inc5	0.470	0.226
edu2	-0.176	Eliminated
edu3	0.146	Eliminated
edu4	-0.461	-0.342
edu5	0.459	0.236
ftcd_score	-0.139	-0.169
ftcd.5.mins	0.007	Eliminated
bdi_score_w00	0.002	Eliminated
cpd_ps	0.003	Eliminated
crv_total_pq1	-0.024	Eliminated
hedonsum_n_pq1	0.127	Eliminated
hedonsum_y_pq1	-0.336	-0.28
shaps_score_pq1	-0.062	-0.056
otherdiag1	0.036	Eliminated
antidepmed1	0.480	0.237
mde_curr1	-0.299	-0.046
NMR	0.372	0.222
Only.Menthol1	0.377	Eliminated
readiness	-0.061	Eliminated

Table 4: Results of Basic Ridge and Lasso Regression Models

Metric	Ridge Regression Model	Lasso Regression Model
AUC (Train)	0.878	0.841
AUC (Test)	0.651	0.705
Accuracy (Train)	0.835	0.818
Accuracy (Test)	0.800	0.808
Lambda Min	0.100	0.033

Model with Interaction Terms

Our investigation of interaction effects progressed through two stages, each guided by specific theoretical considerations.

Basic Interaction Model

In constructing our basic interaction model, we prioritized the examination of treatment-specific interactions that aligned with our primary research objectives regarding treatment effectiveness. For Behavioral Activation (BA), we specifically focused on its interactions with depression-related variables, including current MDD status (mde_curr), Beck Depression Inventory score (bdi_score_w00), and Snaith-Hamilton Pleasure Scale score (shaps_score_pq1). This selection was theoretically driven, as BA was explicitly designed to target depression symptoms, with particular emphasis on addressing anhedonic symptoms. Additionally, we incorporated the interaction between BA and current MDD status to evaluate whether BA's therapeutic effectiveness varies between patients with current versus past MDD diagnoses.

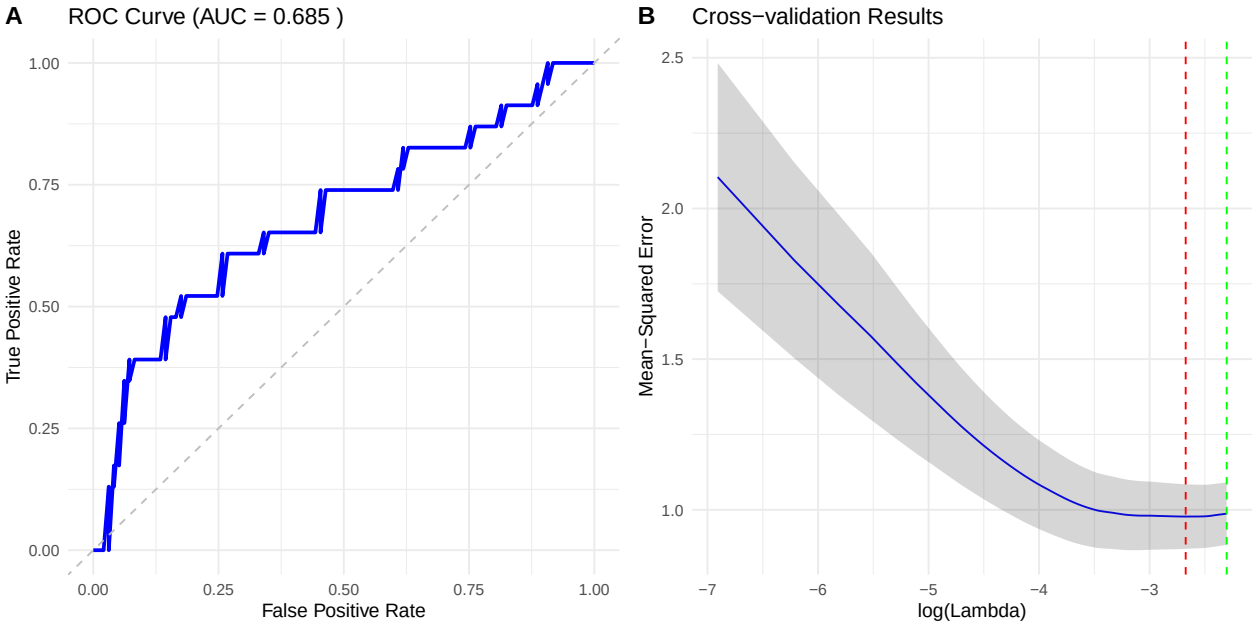
Meanwhile, our analysis of varenicline interactions concentrated on smoking-related characteristics, specifically the Fagerström Test score (ftcd_score), cigarettes per day (cpd_ps), and readiness to quit measures. This selection was informed by previous research findings suggesting that pharmacotherapy outcomes may be moderated by baseline smoking intensity and motivational factors. These interaction terms were essential for understanding the differential effects of varenicline across varying levels of nicotine dependence and quit motivation.

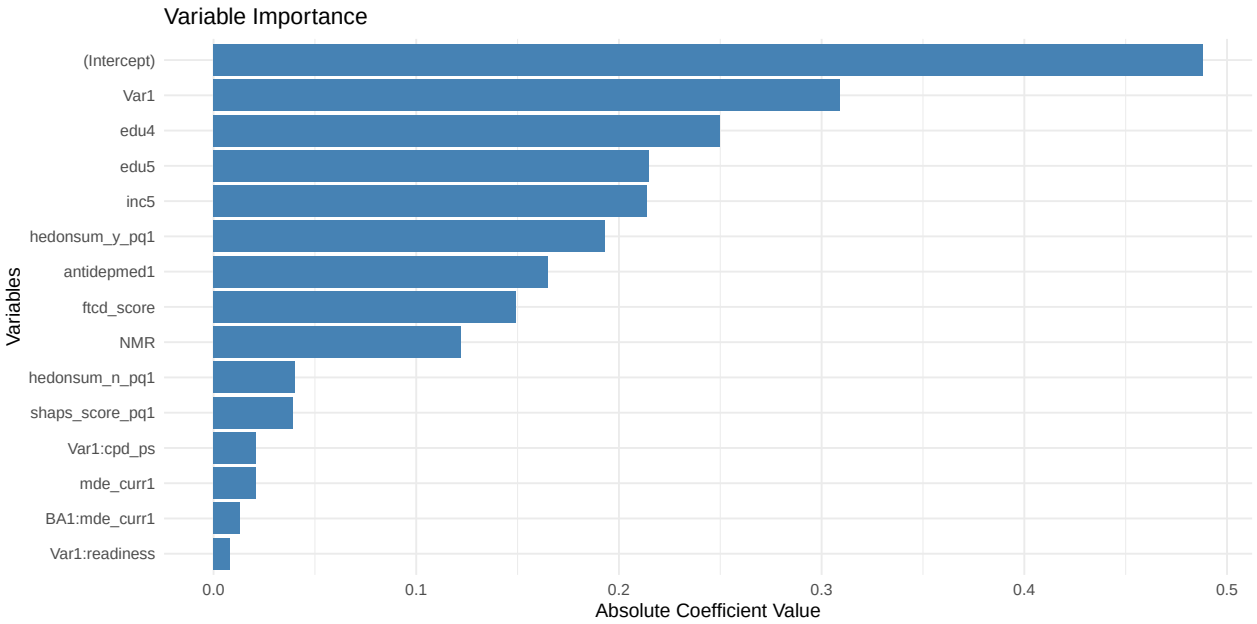
The elastic net model with $\alpha = 0.5$ and $\lambda = 0.069$ can be expressed as:

$$\begin{aligned}
logit(P(abst = 1)) &= \beta_0 + \beta_1 \times Var + \beta_2 \times BA + \beta_3 \times ftcd_score \\
&+ \beta_4 \times antidepmed1 + \beta_5 \times hedonsum_n_pq1 + \beta_6 \times hedonsum_y_pq1 \\
&+ \beta_7 \times shaps_score_pq1 + \beta_8 \times mde_curr1 + \beta_9 \times NMR \\
&+ \beta_{10} \times (BA \times mde_curr1) + \beta_{11} \times (Var \times cpd_ps) + \beta_{12} \times (Var \times readiness) \\
&= -0.488 + 0.309 \times Var + 0.215 \times edu5 - 0.250 \times edu4 \\
&+ 0.165 \times antidepmed1 - 0.149 \times ftcd_score + 0.040 \times hedonsum_n_pq1 \\
&- 0.193 \times hedonsum_y_pq1 - 0.039 \times shaps_score_pq1 - 0.021 \times mde_curr1 \\
&+ 0.122 \times NMR - 0.013 \times (BA \times mde_curr1) + 0.021 \times (Var \times cpd_ps) \\
&+ 0.008 \times (Var \times readiness)
\end{aligned}$$

Our initial interaction model demonstrated moderate predictive performance, achieving an AUC of 0.685 on the test dataset. The variable importance analysis revealed that varenicline maintained its position as a strong predictor, showing a substantial main effect (coefficient = 0.309) that increased the base odds of abstinence by exp(0.309) = 1.36 times. However, we found that varenicline's effectiveness exhibited variation across different baseline smoking characteristics. Particularly noteworthy was the positive interaction with daily cigarette consumption (coefficient = 0.021, odds ratio = exp(0.021) = 1.02), indicating an enhanced treatment effect with higher baseline smoking intensity. To illustrate, for participants reporting 20 cigarettes per day, varenicline's effectiveness showed an enhancement factor of exp(0.021 * 20) = 1.52 compared to its baseline effect.

In examining behavioral treatment interactions, we observed that the relationship between Behavioral Activation (BA) and current MDD status revealed a subtle but important pattern. The negative interaction coefficient (-0.013) suggests that for patients with current MDD, BA's effectiveness was marginally reduced (odds ratio = exp(-0.013) = 0.987, approximately 1.3% reduction in abstinence odds). This finding underscores the particular challenges in treating this specific patient subgroup. Additionally, our analysis revealed that educational level served as a significant moderator of treatment effectiveness, with contrasting effects observed at different levels: edu4 was associated with reduced odds of abstinence (coefficient = -0.250, odds ratio = exp(-0.250) = 0.779, representing a 22.1% reduction), while edu5 showed increased odds (coefficient = 0.215, odds ratio = exp(0.215) = 1.24, indicating a 24% increase).





Advanced Interaction Model

In developing our advanced interaction model, we expanded the scope of our analysis to incorporate a more comprehensive set of theoretically grounded interactions, focusing on three fundamental mechanisms. Firstly, we investigated the interaction between depression symptoms, specifically examining the relationship between Beck Depression Inventory score and Snaith-Hamilton Pleasure Scale score (bdi_score_w00 $\times$ shaps_score_pq1). This interaction was selected based on the hypothesis that concurrent presence of severe depression and pronounced anhedonia might create uniquely challenging conditions for achieving smoking cessation.

Secondly, we examined potential interactions between smoking behavior characteristics and quit motivation (ftcd_score $\times$ readiness). This investigation was driven by the theoretical premise that elevated levels of nicotine dependence, as measured by the Fagerström Test score, might attenuate the typically positive influence of quit motivation on cessation outcomes. Finally, we explored social interactions, particularly focusing on education level's interaction with quit readiness (edu $\times$ readiness) and income level's interaction with antidepressant medication use (inc $\times$ antidepmed). These interactions were chosen to investigate how educational background might moderate the translation of motivation into successful cessation, and how social factors might influence medication effectiveness through differential access to healthcare resources and support systems.

The advanced interaction model with $\alpha = 0.5$ and $\lambda = 0.071$ can be expressed as:

$$
\begin{aligned}
logit(P(abst = 1)) &= \beta_0 + \beta_1 \times Var + \beta_2 \times inc5 + \beta_3 \times edu4 + \beta_4 \times edu5 \\
&+ \beta_5 \times ftcd\_score + \beta_6 \times hedonsum\_n\_pq1 + \beta_7 \times hedonsum\_y\_pq1 \\
&+ \beta_8 \times shaps\_score\_pq1 + \beta_9 \times antidepmed1 + \beta_{10} \times NMR \\
&+ \beta_{11} \times (Var \times cpd\_ps) + \beta_{12} \times (Var \times readiness) \\
&+ \beta_{13} \times (bdi\_score\_w00 \times shaps\_score\_pq1) + \beta_{14} \times (ftcd\_score \times readiness) \\
&+ \beta_{15} \times (edu4 \times readiness) + \beta_{16} \times (inc3 \times antidepmed1) \\
&= -0.430 + 0.242 \times Var + 0.244 \times inc5 - 0.206 \times edu4 + 0.138 \times edu5 \\
&- 0.097 \times ftcd\_score + 0.046 \times hedonsum\_n\_pq1 - 0.156 \times hedonsum\_y\_pq1 \\
&- 0.033 \times shaps\_score\_pq1 + 0.015 \times antidepmed1 + 0.149 \times NMR \\
&+ 0.019 \times (Var \times cpd\_ps) + 0.017 \times (Var \times readiness) \\
&+ 0.000 \times (bdi\_score\_w00 \times shaps\_score\_pq1) - 0.011 \times (ftcd\_score \times readiness) \\
&- 0.014 \times (edu4 \times readiness) + 1.204 \times (inc3 \times antidepmed1)
\end{aligned}
$$

Our advanced interaction model demonstrated superior predictive performance, achieving the highest test set Area Under the Curve (AUC) of 0.711. This improvement in predictive accuracy suggests that the incorporation of comprehensive interaction terms successfully captured important complex relationships in the data. The variable importance analysis revealed several notable interaction patterns that warrant detailed examination.

The most prominent finding emerged from the social domain, where we observed a strong interaction between middle income level and antidepressant medication use (coefficient = 1.204, odds ratio = exp(1.204) = 3.33). This substantial effect indicates that participants in the middle income group experienced more than triple the effectiveness from antidepressant medication compared to other income groups. Treatment-related interactions also demonstrated meaningful patterns: varenicline's effectiveness showed positive interaction with both daily cigarette consumption (coefficient = 0.019, odds ratio per cigarette = exp(0.019) = 1.019) and readiness to quit (coefficient = 0.017, odds ratio per unit = exp(0.017) = 1.017). To illustrate, for participants smoking 20 cigarettes per day, the varenicline treatment effect was enhanced by a factor of exp(0.019 * 20) = 1.46.

The model additionally captured significant behavioral interactions that provide insights into treatment mechanisms. We found that increased nicotine dependence slightly attenuated the benefit of quit readiness (coefficient = -0.011, odds ratio = exp(-0.011) = 0.989 per unit increase in dependence score). Similarly, educational background emerged as a modifier of quit readiness effects, with edu4 level showing a modest negative interaction (coefficient = -0.014, odds ratio = exp(-0.014) = 0.986 per unit increase in readiness score). These findings suggest that both nicotine dependence and educational background influence how effectively quit motivation translates into successful abstinence outcomes.

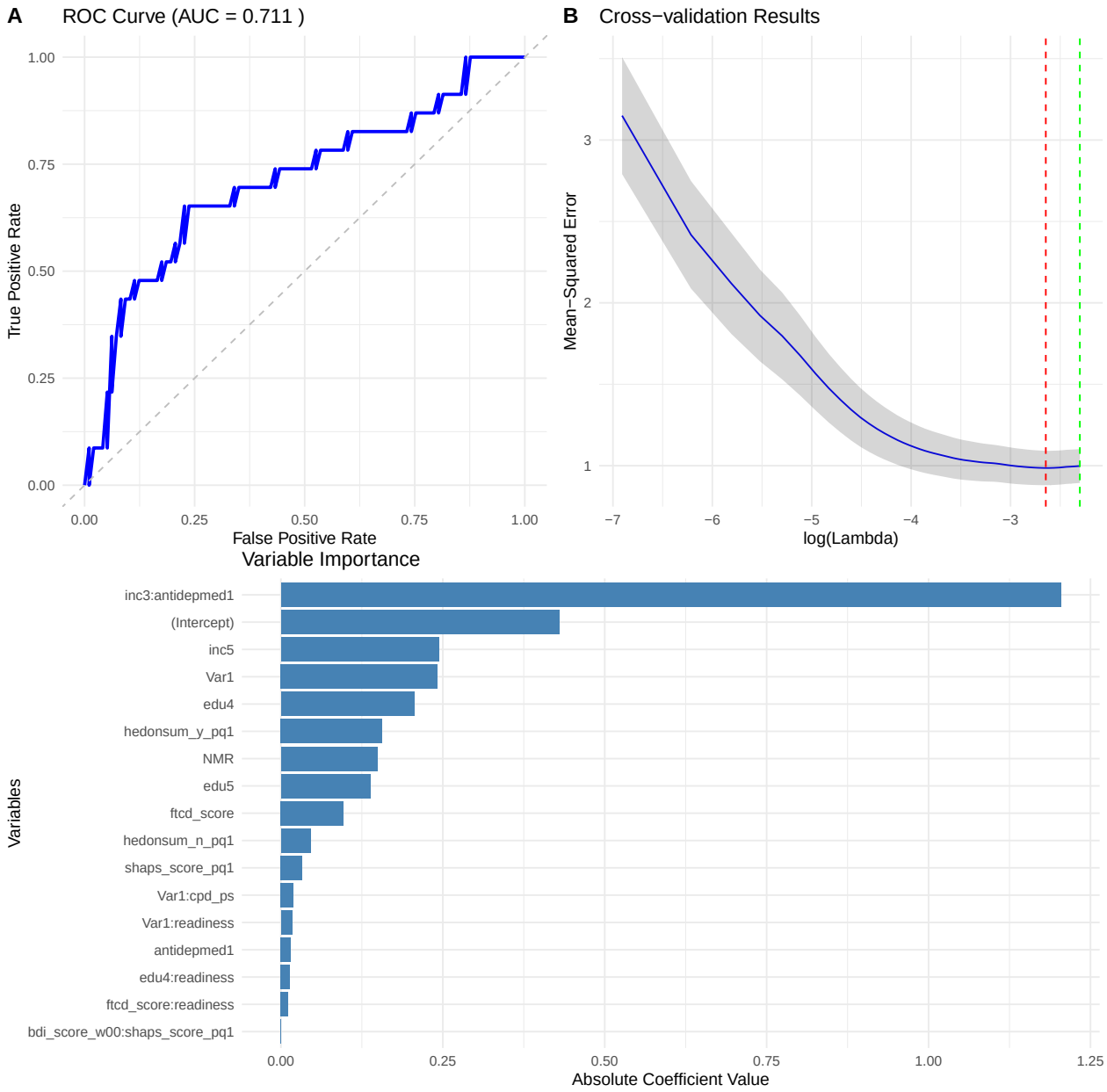


Table 5: Comparison of Models with Interaction Terms

Variable	Model with Basic Interaction Term	Model with Advanced Interaction Terms
(Intercept)	-0.488	-0.43
BA1	Eliminated	Eliminated
Var1	0.309	0.242
age_ps	Eliminated	Eliminated
sex_ps2	Eliminated	Eliminated
NHW1	Eliminated	Eliminated
Black1	Eliminated	Eliminated
Hisp1	Eliminated	Eliminated
inc2	Eliminated	Eliminated
inc3	Eliminated	Eliminated
inc4	Eliminated	Eliminated
inc5	0.214	0.244
edu2	Eliminated	Eliminated

Table 5: Comparison of Models with Interaction Terms *(continued)*

Variable	Model with Basic Interaction Term	Model with Advanced Interaction Terms
edu3	Eliminated	Eliminated
edu4	-0.25	-0.206
edu5	0.215	0.138
bdi_score_w00	Eliminated	Eliminated
cpd_ps	Eliminated	Eliminated
ftcd_score	-0.149	-0.097
ftcd.5.mins	Eliminated	Eliminated
shaps_score_pq1	-0.039	-0.033
mde_curr1	-0.021	Eliminated
crv_total_pq1	Eliminated	Eliminated
hedonsum_n_pq1	0.04	0.046
hedonsum_y_pq1	-0.193	-0.156
otherdiag1	Eliminated	Eliminated
antidepmed1	0.165	0.015
NMR	0.122	0.149
Only.Menthol1	Eliminated	Eliminated
readiness	Eliminated	Eliminated
BA1:shaps_score_pq1	Eliminated	Eliminated
BA1:mde_curr1	-0.013	Eliminated
BA1:bdi_score_w00	Eliminated	Eliminated
Var1:ftcd_score	Eliminated	Eliminated
Var1:cpd_ps	0.021	0.019
Var1:readiness	0.008	0.017
BA1:Black1	NA	Eliminated
BA1:NHW1	NA	Eliminated
Var1:Black1	NA	Eliminated
Var1:NHW1	NA	Eliminated
bdi_score_w00:shaps_score_pq1	NA	0
shaps_score_pq1:mde_curr1	NA	Eliminated
cpd_ps:ftcd_score	NA	Eliminated
ftcd_score:readiness	NA	-0.011
inc2:antidepmed1	NA	Eliminated
inc3:antidepmed1	NA	1.204
inc4:antidepmed1	NA	Eliminated
inc5:antidepmed1	NA	Eliminated
edu2:readiness	NA	Eliminated
edu3:readiness	NA	Eliminated
edu4:readiness	NA	-0.014
edu5:readiness	NA	Eliminated
ftcd_score:Only.Menthol1	NA	Eliminated
NMR:Only.Menthol1	NA	Eliminated

Table 6: Results of Models with Interaction Terms

Metric	Model with Basic Interaction Term	Model with Advanced Interaction Terms
AUC (Train)	0.845	0.854
AUC (Test)	0.685	0.711
Accuracy (Train)	0.802	0.802
Accuracy (Test)	0.800	0.808
Lambda Min	0.069	0.071

Discussion

In this study, we employed a series of regularized regression models to investigate the effectiveness of behavioral activation and varenicline for smoking cessation among individuals with current or past MDD. Through progressively complex modeling approaches, from basic LASSO and Ridge regression to interaction models, we gained several key insights about treatment effects and their moderators.

Our analytical models demonstrated consistently robust predictive performance, with test AUC values ranging from 0.651 to 0.711. The advanced interaction model achieved the highest predictive accuracy, highlighting the importance of considering treatment effect heterogeneity in our analysis. Notably, the relatively small gap between training and test performance (AUC difference < 0.15) suggests our models successfully avoided severe overfitting, which we attribute to the effective implementation of regularization strategies.

Regarding treatment effectiveness, varenicline emerged as a consistently strong predictor of smoking cessation success across all model specifications. The basic LASSO model identified a substantial positive effect (coefficient = 0.802), which maintained its robustness even after accounting for various interaction terms. In contrast, behavioral activation demonstrated more modest and variable effects, with the lack of a strong main effect aligning with previous trial findings where BASC showed no significant advantage over standard behavioral treatment.

Our interaction models revealed significant heterogeneity in treatment effects across different patient subgroups. For varenicline, we observed enhanced effectiveness among heavier smokers (Var $\times$ cpd_ps: 0.019) and participants with higher readiness to quit (Var $\times$ readiness: 0.017), suggesting particular benefit for more severely dependent smokers who demonstrate strong quit motivation. For behavioral activation, the negative interaction with current MDD status (BA $\times$ mde_curr: -0.013) indicated potentially reduced effectiveness among currently depressed patients, underscoring the challenges in treating this specific subgroup.

The advanced interaction model uncovered important social dimensions of treatment response. We identified a strong positive interaction between middle income level and antidepressant medication use (inc3 $\times$ antidepmed: 1.204), suggesting significant social moderation of depression treatment effectiveness. Additionally, the negative interaction between education level and quit readiness (edu4 $\times$ readiness: -0.014) indicated that the translation of quit motivation into successful cessation varies across educational backgrounds.

Several limitations should be noted in our analysis. First, our relatively small sample size (n=241 after removing missing values) may have limited our ability to detect smaller interaction effects. Second, our binary outcome measure (7-day point prevalence abstinence) may not fully capture the complexity of smoking cessation trajectories. Third, our models did not account for potential time-varying treatment effects or the dynamic nature of depression symptoms during quit attempts.

Future research directions should address several key areas. Larger sample sizes would enable more precise estimation of interaction effects and potentially reveal additional treatment effect moderators. The incorporation of longitudinal measures for both smoking behavior and depression symptoms could provide valuable insights into temporal dynamics of treatment effects. Additionally, more sophisticated modeling approaches, such as Bayesian methods or advanced machine learning techniques, might better capture the complex relationships between patient characteristics and treatment outcomes.

Despite these limitations, our findings yield important clinical implications. The consistent effectiveness of varenicline, particularly among heavier smokers, supports its continued use as a first-line treatment. However, the variable effects of behavioral activation and the significant role of social factors suggest the need for more personalized treatment approaches. Clinicians should carefully consider both clinical and social factors when developing smoking cessation interventions for patients with MDD.

References

1. Hitsman, Brian, George D. Papandonatos, Jacqueline K. Gollan, Mark D. Huffman, Raymond Niaura, David C. Mohr, Anna K. Veluz-Wilkins, et al. "Efficacy and Safety of Combination Behavioral Activation for Smoking Cessation and Varenicline for Treating Tobacco Dependence among Individuals with Current or Past Major Depressive Disorder: A 2 × 2 Factorial, Randomized, Placebo-controlled Trial." Addiction 118, no. 9 (September 2023): 1710–25. https://doi.org/10.1111/add.16209.

Code Appendix

```
knitr::opts_chunk$set(echo = FALSE)
knitr::opts_chunk$set(message = FALSE)
knitr::opts_chunk$set(warning = FALSE)
library(mice, warn.conflicts = FALSE)
library(naniar)
library(ggplot2)
library(dplyr)
library(readr)
library(tidyr)
library(readxl)
library(ggpubr)
library(gtsummary)
library(GGally)
library(ggcorrplot)
library(knitr)
library(kableExtra)
library(lubridate)
library(patchwork)
library(introdataviz)
library(glmnet)
library(gt)
library(LOLearn)
library(gridExtra)
library(purrr)
library(pROC)
library(caret)

# Load the data
data <- read.csv("../Data/project2.csv")
data
# Transform data type
data$abst <- as.factor(data$abst)
data$Var <- as.factor(data$Var)
data$BA <- as.factor(data$BA)
data$sex_ps <- as.factor(data$sex_ps)
data$NHW <- as.factor(data$NHW)
data$Black <- as.factor(data$Black)
data$Hisp <- as.factor(data$Hisp)
data$inc <- as.factor(data$inc)
data$edu <- as.factor(data$edu)
data$otherdiag <- as.factor(data$otherdiag)
data$antidepmed <- as.factor(data$antidepmed)
data$mde_curr <- as.factor(data$mde_curr)
data$Only.Menthol <- as.factor(data$Only.Menthol)
data$readiness <- as.numeric(data$readiness)

dim(data)

summary(data)

data <- na.omit(data)
```

```r
summary(data)
dim(data)

summary_var1 <- tbl_summary(data %>% select(-id), by = Var, type = list(readiness ~ "continuous")) %>%
  add_p() %>%
  modify_header(stat_1 = "**No Varenicline**", stat_2 = "**Varenicline**")
summary_var2 <- tbl_summary(data  %>% select(-id), by = BA, type = list(readiness ~ "continuous")) %>%
  add_p() %>%
  modify_header(stat_1 = "**No Psychotherapy**", stat_2 = "**Psychotherapy**")

tbl_merge(
  tbls = list(summary_var1, summary_var2),
  tab_spanner = c("**Grouped by Pharmacotherapy Status**", "**Grouped by Psychotherapy Status**")
) %>%
  as_kable_extra(booktabs = TRUE,
caption = "Data Summary Table",
longtable = TRUE, linesep = "") %>%
kableExtra::kable_styling(font_size = 8, latex_options = c("repeat_header", "HOLD_position"))

continuous_vars <- data[, c("age_ps", "ftcd_score", "bdi_score_w00", "cpd_ps", "crv_total_pq1", "hedonsu
corr_plot <- ggpairs(continuous_vars, title = "Correlation Matrix of Baseline Predictors") +
  theme_minimal()

ggsave("../Figures/corr_plot.png", plot = corr_plot, dpi = 300)
corr_plot

dist_numeric <- data %>%
  select(-id) %>%
  keep(is.numeric) %>%
  gather() %>%
  ggplot(aes(value)) +
    facet_wrap(~ key, scales = "free") +
    geom_histogram() +
    ggtitle("Distribution of Numeric Variables")

dist_cat <- data %>%
  select(-id) %>%
  keep(is.factor) %>%
  gather(key = "key", value = "value") %>%
  ggplot(aes(x = value)) +
    facet_wrap(~ key, scales = "free") +
    geom_bar()+
    ggtitle("Distribution of Categorical Variables")

gridExtra::grid.arrange(dist_numeric, dist_cat, ncol = 2)


data$hedonsum_n_pq1 <- log(data$hedonsum_n_pq1 + 1)
data$hedonsum_y_pq1 <- log(data$hedonsum_y_pq1 + 1)
data$NMR <- log(data$NMR)

dist_after_log <- data %>%
  select(c(hedonsum_n_pq1, hedonsum_y_pq1, NMR)) %>%
```

```
  gather() %>%
  ggplot(aes(value)) +
    facet_wrap(~ key, scales = "free") +
    geom_histogram() +
    ggtitle("Distribution after Log Transformation")

ggsave("../Figures/dist_after_log.png", plot = dist_after_log, dpi = 300)
dist_after_log

# Split the data
set.seed(123)

train_index <- createDataPartition(data$abst, p = 0.5, list = FALSE)

train_data <- data[train_index, ] %>% select(-id)
test_data <- data[-train_index, ] %>% select(-id)

X_train <- model.matrix(abst ~ ., data = train_data)[, -1]
Y_train <- factor(train_data$abst)

X_test <- model.matrix(abst ~ ., data = test_data)[, -1]
Y_test <- factor(test_data$abst)
model_fit <- function(X_train, Y_train, X_test, Y_test, alpha, lambda_seq = seq(0.001, 0.1, length.out =
 set.seed(123)
 cv_model <- cv.glmnet(X_train, Y_train, alpha = alpha, family = "binomial", lambda = lambda_seq)
 best_lambda <- cv_model$lambda.min
 lasso_coef <- coef(cv_model, s = best_lambda)

 # Get predictions for both train and test sets
 train_pred <- predict(cv_model, newx = X_train, s = best_lambda, type = "response")
 test_pred <- predict(cv_model, newx = X_test, s = best_lambda, type = "response")

 # Convert predictions to binary using 0.5 threshold
 train_pred_binary <- ifelse(train_pred > 0.5, 1, 0)
 test_pred_binary <- ifelse(test_pred > 0.5, 1, 0)

 # ROC and AUC
 train_roc <- roc(Y_train, as.numeric(train_pred))
 test_roc <- roc(Y_test, as.numeric(test_pred))

 # Calculate metrics
 metrics <- data.frame(
   Metric = c(
     "AUC (Train)",
     "AUC (Test)",
     "Accuracy (Train)",
     "Accuracy (Test)",
     "Lambda Min"
   ),
   Value = c(
     auc(train_roc),
     auc(test_roc),
     mean(train_pred_binary == Y_train),
```

```
    mean(test_pred_binary == Y_test),
    cv_model$lambda.min
  )
) %>%
  mutate(Value = round(Value, 3))

# Create ROC curve data
roc_df <- data.frame(
  FPR = 1 - test_roc$specificities,
  TPR = test_roc$sensitivities
)

# Create ROC plot using ggplot2
roc_plot <- ggplot(roc_df, aes(x = FPR, y = TPR)) +
  geom_line(color = "blue", size = 1) +
  geom_abline(slope = 1, intercept = 0, linetype = "dashed", color = "gray") +
  labs(
    title = paste("ROC Curve (AUC =", round(auc(test_roc), 3), ")"),
    x = "False Positive Rate",
    y = "True Positive Rate"
  ) +
  theme_minimal()

# Create CV plot data
cv_df <- data.frame(
  lambda = log(cv_model$lambda),
  mean = cv_model$cvm,
  upper = cv_model$cvup,
  lower = cv_model$cvlo
)

# Create CV plot using ggplot2
cv_plot <- ggplot(cv_df, aes(x = lambda)) +
  geom_line(aes(y = mean), color = "blue") +
  geom_ribbon(aes(ymin = lower, ymax = upper), alpha = 0.2) +
  geom_vline(xintercept = log(cv_model$lambda.min), linetype = "dashed", color = "red") +
  geom_vline(xintercept = log(cv_model$lambda.1se), linetype = "dashed", color = "green") +
  labs(
    title = "Cross-validation Results",
    x = "log(Lambda)",
    y = "Mean-Squared Error"
  ) +
  theme_minimal()

# Combine plots
combined_plot <- ggarrange(roc_plot, cv_plot,
                           ncol = 2,
                           labels = c("A", "B"))

# Create coefficient dataframe
coef_df <- data.frame(
  Variable = rownames(lasso_coef),
  Coefficient = as.vector(lasso_coef)
```

```
 ) %>%
 mutate(Coefficient = ifelse(Coefficient == 0, "Eliminated", round(Coefficient, 3)))

 return(list(
   combined_plot = combined_plot,
   coef_df = coef_df,
   metrics = metrics
 ))
}

plot_variable_importance <- function(coef_df) {
 # Calculate variable importance by taking absolute values of non-eliminated coefficients
 var_importance <- coef_df %>%
   filter(Coefficient != "Eliminated") %>%
   mutate(
     Coefficient = as.numeric(Coefficient),
     Importance = abs(Coefficient)
   ) %>%
   arrange(desc(Importance))

 # Create bar plot of variable importance
 importance_plot <- ggplot(var_importance, aes(x = reorder(Variable, Importance), y = Importance)) +
   geom_bar(stat = "identity", fill = "steelblue") +
   coord_flip() + # Flip coordinates for horizontal bars
   theme_minimal() + # Use minimal theme for clean look
   labs(title = "Variable Importance",
        x = "Variables",
        y = "Absolute Coefficient Value")

 # Return both plot and processed data
 return(list(
   plot = importance_plot,
   importance_df = var_importance
 ))
}
lambda_seq <- seq(0.001, 0.1, length.out = 100)
lasso_model <- model_fit(X_train, Y_train, X_test, Y_test, alpha = 1, lambda_seq = lambda_seq)

print(lasso_model$combined_plot)
plot_variable_importance(lasso_model$coef_df)$plot

ggsave("../Figures/lasso_model.png", plot = lasso_model$combined_plot, dpi = 300)
ggsave("../Figures/lasso_variable_importance.png", plot = plot_variable_importance(lasso_model$coef_df)$

lambda_seq <- seq(0.001, 0.1, length.out = 100)
ridge_model <- model_fit(X_train, Y_train, X_test, Y_test, alpha = 0, lambda_seq = lambda_seq)

print(ridge_model$combined_plot)
plot_variable_importance(ridge_model$coef_df)$plot

ggsave("../Figures/ridge_model.png", plot = ridge_model$combined_plot, dpi = 300)
ggsave("../Figures/ridge_variable_importance.png", plot = plot_variable_importance(ridge_model$coef_df)$
```

```r
combined_table <- full_join(ridge_model$coef_df, lasso_model$coef_df, by = "Variable") %>%
  kable(
    col.names = c("Variable", "Ridge Regression Model", "Lasso Regression Model"),
    caption = "Comparison of Basic Ridge and Lasso Regression Models"
  ) %>%
  kable_styling(bootstrap_options = c("striped", "hover"),
                full_width = FALSE, font_size = 8, latex_options = c("repeat_header", "HOLD_position")) 
  column_spec(1, bold = TRUE)

combined_table

combined_result <- full_join(ridge_model$metrics, lasso_model$metrics, by = "Metric") %>%
  kable(
    col.names = c("Metric", "Ridge Regression Model", "Lasso Regression Model"),
    caption = "Results of Basic Ridge and Lasso Regression Models"
  ) %>%
  kable_styling(bootstrap_options = c("striped", "hover"),
                full_width = FALSE, font_size = 8, latex_options = c("repeat_header", "HOLD_position")) 
  column_spec(1, bold = TRUE)

combined_result
X_train_interactions <- model.matrix(
 ~ BA + Var +
   # Demographic variables
   age_ps + sex_ps + NHW + Black + Hisp + inc + edu +
   # Clinical and smoking-related variables
   bdi_score_w00 + cpd_ps + ftcd_score + ftcd.5.mins +
   shaps_score_pq1 + mde_curr + crv_total_pq1 +
   hedonsum_n_pq1 + hedonsum_y_pq1 +
   otherdiag + antidepmed + NMR + Only.Menthol + readiness +
   # BA interaction terms
   BA:shaps_score_pq1 +    # BA * anhedonia
   BA:mde_curr +           # BA * current depression status
   BA:bdi_score_w00 +      # BA * depression severity
   # Varenicline interaction terms
   Var:ftcd_score +        # Varenicline * nicotine dependence
   Var:cpd_ps +            # Varenicline * cigarettes per day
   Var:readiness,          # Varenicline * readiness to quit
 data = train_data %>% select(-abst))[, -1]

X_test_interactions <- model.matrix(
 ~ BA + Var +
   # Demographic variables
   age_ps + sex_ps + NHW + Black + Hisp + inc + edu +
   # Clinical and smoking-related variables
   bdi_score_w00 + cpd_ps + ftcd_score + ftcd.5.mins +
   shaps_score_pq1 + mde_curr + crv_total_pq1 +
   hedonsum_n_pq1 + hedonsum_y_pq1 +
   otherdiag + antidepmed + NMR + Only.Menthol + readiness +
   # BA interaction terms
   BA:shaps_score_pq1 +    # BA * anhedonia
   BA:mde_curr +           # BA * current depression status
   BA:bdi_score_w00 +      # BA * depression severity
```

```
   # Varenicline interaction terms
   Var:ftcd_score +       # Varenicline * nicotine dependence
   Var:cpd_ps +           # Varenicline * cigarettes per day
   Var:readiness,         # Varenicline * readiness to quit
 data = test_data %>% select(-abst))[, -1]
lambda_seq <- seq(0.001, 0.1, length.out = 100)
basic_interaction_model <- model_fit(X_train_interactions, Y_train, X_test_interactions, Y_test, alpha =

print(basic_interaction_model$combined_plot)
plot_variable_importance(basic_interaction_model$coef_df)$plot

ggsave("../Figures/basic_interaction_model.png", plot = basic_interaction_model$combined_plot, dpi = 300
ggsave("../Figures/basic_interaction_variable_importance.png", plot = plot_variable_importance(basic_int

# Create model matrix with extended interaction terms
X_train_interactions <- model.matrix(
 ~ BA + Var +
   # Demographic variables
   age_ps + sex_ps + NHW + Black + Hisp + inc + edu +
   # Clinical and smoking-related variables
   bdi_score_w00 + cpd_ps + ftcd_score + ftcd.5.mins +
   shaps_score_pq1 + mde_curr + crv_total_pq1 +
   hedonsum_n_pq1 + hedonsum_y_pq1 +
   otherdiag + antidepmed + NMR + Only.Menthol + readiness +
   # BA interaction terms
   BA:shaps_score_pq1 +   # BA * anhedonia
   BA:mde_curr +          # BA * current depression status
   BA:bdi_score_w00 +     # BA * depression severity
   BA:Black +             # BA * race (Black)
   BA:NHW +               # BA * race (Non-Hispanic White)
   # Varenicline interaction terms
   Var:ftcd_score +       # Varenicline * nicotine dependence
   Var:cpd_ps +           # Varenicline * cigarettes per day
   Var:readiness +        # Varenicline * readiness to quit
   Var:Black +            # Varenicline * race (Black)
   Var:NHW +              # Varenicline * race (Non-Hispanic White)
   # Depression symptom interactions
   bdi_score_w00:shaps_score_pq1 +  # Depression severity * anhedonia
   mde_curr:shaps_score_pq1 +       # Current depression status * anhedonia
   # Smoking behavior interactions
   ftcd_score:cpd_ps +          # Nicotine dependence * cigarettes per day
   ftcd_score:readiness +       # Nicotine dependence * readiness to quit
   # social status interactions
   antidepmed:inc +             # Antidepressant medication * income
   edu:readiness +              # Education * readiness to quit
   # Menthol cigarette interactions
   Only.Menthol:ftcd_score +    # Menthol use * nicotine dependence
   Only.Menthol:NMR,            # Menthol use * nicotine metabolism ratio
 data = train_data %>% select(-abst))[, -1]

X_test_interactions <- model.matrix(
 ~ BA + Var +
   # Demographic variables
```

```
    age_ps + sex_ps + NHW + Black + Hisp + inc + edu +
    # Clinical and smoking-related variables
    bdi_score_w00 + cpd_ps + ftcd_score + ftcd.5.mins +
    shaps_score_pq1 + mde_curr + crv_total_pq1 +
    hedonsum_n_pq1 + hedonsum_y_pq1 +
    otherdiag + antidepmed + NMR + Only.Menthol + readiness +
    # BA interaction terms
    BA:shaps_score_pq1 +    # BA * anhedonia
    BA:mde_curr +           # BA * current depression status
    BA:bdi_score_w00 +      # BA * depression severity
    BA:Black +              # BA * race (Black)
    BA:NHW +                # BA * race (Non-Hispanic White)
    # Varenicline interaction terms
    Var:ftcd_score +        # Varenicline * nicotine dependence
    Var:cpd_ps +            # Varenicline * cigarettes per day
    Var:readiness +         # Varenicline * readiness to quit
    Var:Black +             # Varenicline * race (Black)
    Var:NHW +               # Varenicline * race (Non-Hispanic White)
    # Depression symptom interactions
    bdi_score_w00:shaps_score_pq1 +  # Depression severity * anhedonia
    mde_curr:shaps_score_pq1 +       # Current depression status * anhedonia
    # Smoking behavior interactions
    ftcd_score:cpd_ps +          # Nicotine dependence * cigarettes per day
    ftcd_score:readiness +       # Nicotine dependence * readiness to quit
    # social status interactions
    antidepmed:inc +             # Antidepressant medication * income
    edu:readiness +              # Education * readiness to quit
    # Menthol cigarette interactions
    Only.Menthol:ftcd_score +    # Menthol use * nicotine dependence
    Only.Menthol:NMR,            # Menthol use * nicotine metabolism ratio
  data = test_data %>% select(-abst))[, -1]
lambda_seq <- seq(0.001, 0.1, length.out = 100)
advanced_interaction_model <- model_fit(X_train_interactions, Y_train, X_test_interactions, Y_test, alpl

print(advanced_interaction_model$combined_plot)
plot_variable_importance(advanced_interaction_model$coef_df)$plot

ggsave("../Figures/advanced_interaction_model.png", plot = advanced_interaction_model$combined_plot, dpi
ggsave("../Figures/advanced_interaction_variable_importance.png", plot = plot_variable_importance(advanc

combined_table <- full_join(basic_interaction_model$coef_df, advanced_interaction_model$coef_df, by = "V
  kable(
    col.names = c("Variable", "Model with Basic Interaction Term", "Model with Advanced Interaction Terr
    caption = "Comparison of Models with Interaction Terms"
  ) %>%
  kable_styling(bootstrap_options = c("striped", "hover"),
                full_width = FALSE, font_size = 8, latex_options = c("repeat_header", "HOLD_position"))
  column_spec(1, bold = TRUE) # Make Variable column bold

combined_table

combined_result <- full_join(basic_interaction_model$metrics, advanced_interaction_model$metrics, by =
  kable(
```

```
    col.names = c("Metric", "Model with Basic Interaction Term", "Model with Advanced Interaction Terms
    caption = "Results of Models with Interaction Terms"
  ) %>%
  kable_styling(bootstrap_options = c("striped", "hover"),
                full_width = FALSE, font_size = 8, latex_options = c("repeat_header", "HOLD_position"))
  column_spec(1, bold = TRUE) # Make Metric column bold

combined_result
```