Function Draft

Yingxi Kong

Abstract

Background:

Methods:

Results:

Conclusion:

Introduction

Methods

The data in this analysis is a collaboration with Dr. George Papandonatos from a 2 x 2 factorial, randomized, placebo-controlled study examining the efficacy and safety of behavioral activation for smoking cessation (BASC) and varenicline in treating tobacco dependence among adults with current or past major depressive disorder (MDD). Our sample population consists of 300 adults smokers with or previously with MDD. Patients were randomly assigned to either behavioral activation for smoking cessation (BASC) or standard behavioral treatment (ST) and either varenicline or placebo groups. Randomization was stratified by clinical site, sex, and level of depressive symptoms to ensure balanced representation across these factors. The data also records patients' smoking cessation outcomes and relevant baseline characteristics. Key variables include smoking abstinence status, demographic characteristics (sex, age, income, education), their smoking behaviors (cigarettes per day, time to first cigarette after getting up, nicotine dependence score), and their psychiatric measures (MDD status, anhedonia score, other diagnoses, and antidepressant usage). Using this data, our analysis aims to identify baseline variables as moderators of the treatment effects on end-of-treatment (EOT) abstinence and as predictors of smoking cessation, controlling for behavioral treatment and pharmacotherapy.

Data Preprocessing

To prepare the data for analysis, we firstly convert all categorical variables to factor and for socioeconomic factors (income and education) with ordinal levels, we recoded levels in order to improve readability and interpretability. In addition, we generate a new treatment variables to capture the four distinct intervention groups formed by the 2x2 factorial design, including `ST + placebo`, `ST + varenicline`, `BASC + placebo`, and `BASC + varenicline.` This new treatment variable was set to reference `ST + placebo` for comparison among groups. Additionally, we combined race and ethnicity indicators into a single race variable with categories including "Black," "Hispanic," "Non-Hispanic White," "Mixed Race," and "Unknown."

The data also contains various levels of missingness across several variables presented in `Table 1.` Nicotine Metabolism Ratio (`NMR`) has the highest missingness rate, with 7% of observations missing. The FTCD score at baseline (`ftcd_score`) has the lowest missing rate, 0.33%, with only one patient missing information on this variable. Given the limited sample size of this data, we prefer to maintain as many observations as possible in our analysis. Thus, to address the missingness, we applied a multiple imputation approach using the `mice()` function from the mice package in R which provides plausible values for all missing entries across five imputed datasets.

Table 1: Summary of Missing Data Patterns Across Variables

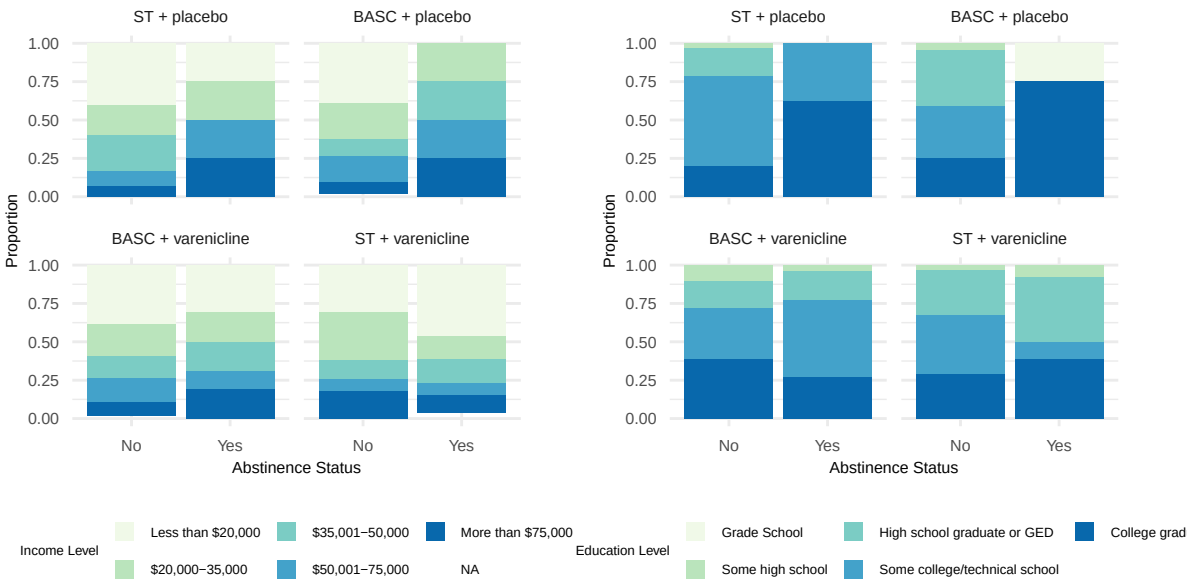
Variable	Missing Count	Missing Percentage
NMR	21	7 %
crv_total_pq1	18	6 %
readiness	17	5.67 %
inc	3	1 %
shaps_score_pq1	3	1 %
Only.Menthol	2	0.67 %
ftcd_score	1	0.33 %

Data Exploration and Transformation

Before conducting the primary analysis, we performed exploratory data analysis (EDA) to examine baseline characteristics, assess data distributions, and identify potential relationships within the dataset.

For categorical variables, we plot bar charts to show patterns across treatment groups and abstinence groups shown in `Figure 1` and `Figure 2.` Income in `Figure 1` exhibits differences among groups and abstinence outcomes. For example, participants with income less than $20,000 are less likely to stop smoking in both the `ST + placebo` and `BASC + varenicline` groups that more people would still continue to smoke at the week 27 follow-up. Even in the `BASC + placebo` group, no patients stop smoking at week 27. However, this pattern reverses in the `ST + varenicline` group that lower-income individuals show a relatively higher likelihood of smoking abstinence. That is, the combination of standard treatment with varenicline may have a greater impact on smoking cessation for lower-income participants. This reversal pattern suggests that income level might be a potential moderator of the treatment effectiveness on the EOT abstinence among people with MDD.

Figure 1: Baseline Characteristics by Abstinence Status and Treatment Group (Categorical 1)



Similar for education level presented in `Figure 2,` people with lower education level (grade school, some high school, or high school graduate) shows less probability of smoking cessation, especially in the `ST + placebo,` `BASC + placebo,` and `BASC + varenicline` group, suggesting the potential association between education level and smoking cessation after treatment.However, participants in the `ST + varenicline` groups show

relatively higher likelihood of abstinence, suggesting potential interacting relationship between education level and treatment assignment on the abstinence. In addition, college graduated participants in the two placebo groups show higher probability of abstinence while those participants in the two varenicline groups show reversed pattern as well, further suggesting that education level could be a potential moderator of the treatment effects on the EOT abstinence among people with MDD.

Race and the indicator of exclusive mentholated cigarette users (`Only.Menthol`) also exhibit difference distribution across treatment groups and outcome values shown in `Figure 2.` For instance, black people generally face greater challenges of smoking cessation as they continuously exhibit lower abstinence rate across different treatment groups, particularly in the `ST + placebo` and `BASC + placebo` groups. However, in the `BASC + varenicline` group, the difference between proportion of black participants who continue or stop smoking becomes less pronounced. Even in the `ST + varenicline` group, black participants exhibit higher abstinence rates, suggesting that varenicline may be particularly effective to mitigate challenges in cessation among black individuals. Also for `Only.Menthol`, among the first three treatment groups, mentholated cigarette users show lower likelihood to achieve smoking cessation and non-mentholated cigarette users exhibit higher probability to stop smoking. This pattern is reversed in the `ST + varenicline` group again. These findings suggest that race and the indicator of exclusive mentholated cigarette users could be potential predictors or moderators of the treatment effects on the EOT abstinence for people with MDD.

Figure 2: Baseline Characteristics by Abstinence Status and Treatment Group (Categorical 2)

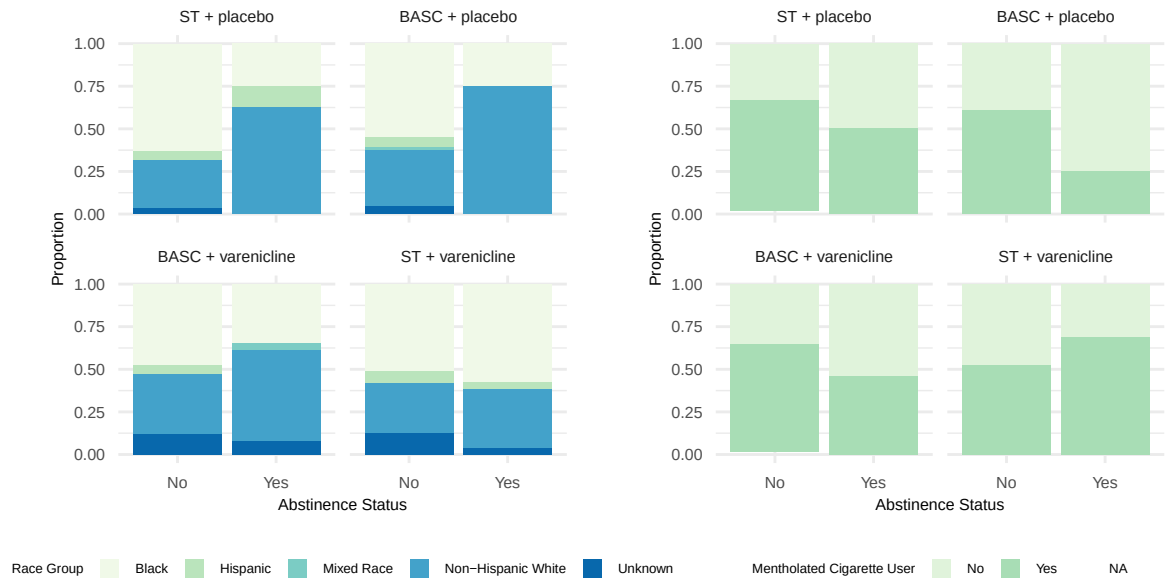


Figure 3: Baseline Characteristics by Abstinence Status and Treatment Group (Continuous 1)

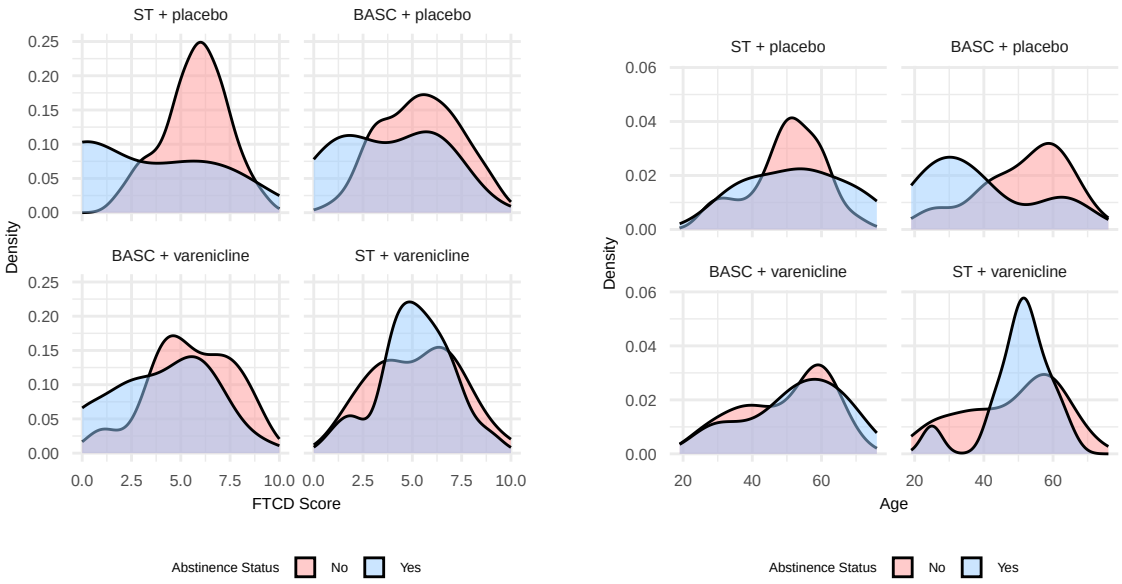


Figure 4: Baseline Characteristics by Abstinence Status and Treatment Group (Continuous 2)

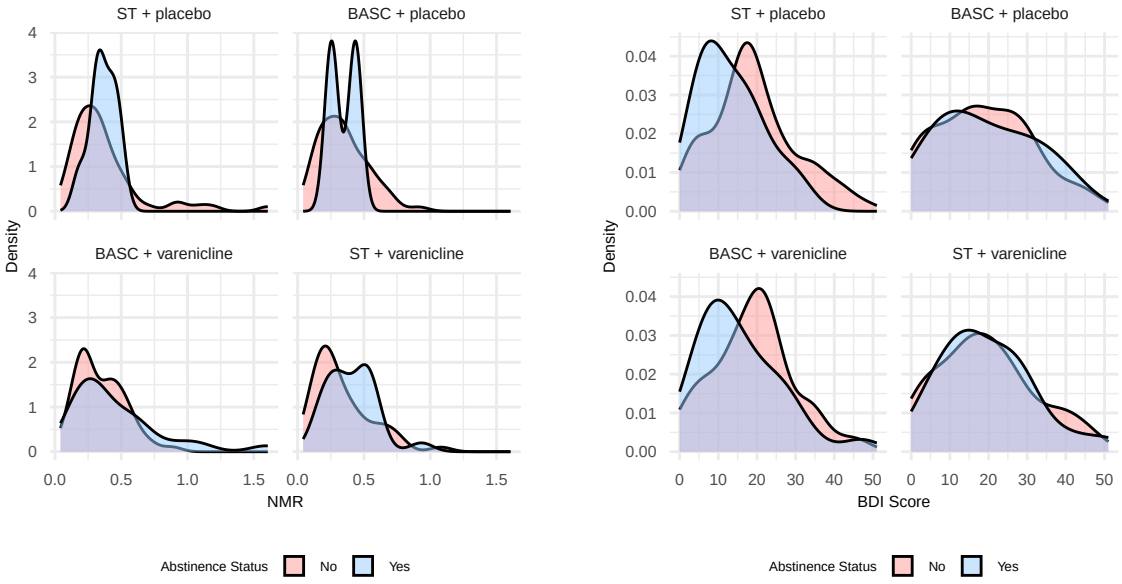


Table 2: Variable Transformation on Skewness

Variable	Transformation	Skewness before Transformation	Skewness after Transformation
hedonsum_n_pq1	Square Root Transformation	1.338843	-0.0591728

Table 2: Variable Transformation on Skewness *(continued)*

Variable	Transformation	Skewness before Transformation	Skewness after Transformation
hedonsum_y_pq1	Square Root Transformation	1.391398	0.0620129
shaps_score_pq1	Inverse Hyperbolic Sine Transformation	1.707230	0.5217093
NMR	Log Transformation	1.915358	-0.2241582

Data Preprocessing

Results

`Table 1` presents an overall summary statistics of patients' baseline characteristics by their behavioral and pharmacological treatment assignment. Since our study is a 2×2, factorial, randomized, placebo-controlled trial, patients are randomly assigned to either behavioral activation for smoking cessation group (BASC) or standard behavioral treatment group (ST) and either varenicline or placebo blister pack. Patients can be categorized into four treatment arm groups: BASC + placebo, BASC + varenicline, ST + placebo, and ST + varenicline. Seeing from `Table 1`, the two placebo groups both have 68 obervations while the two varenicline groups both have 83 observations.

Most variables are evenly distributed across the four treatment arms, which reflects successful randomization in this factorial trial. However, a few key factors, such as socioeconomic indicators (income and education) and specific mental health variables (MDD status, DSM-5 diagnoses), exhibit slight variations that may influence outcomes. Notably, treatment arms with varenicline show higher abstinence rates than placebo groups, suggesting the potential efficacy of this pharmacotherapy in combination with behavioral interventions. While many baseline characteristics are evenly distributed across groups, some may still function as moderators, potentially interacting with treatment assignment to affect abstinence success.

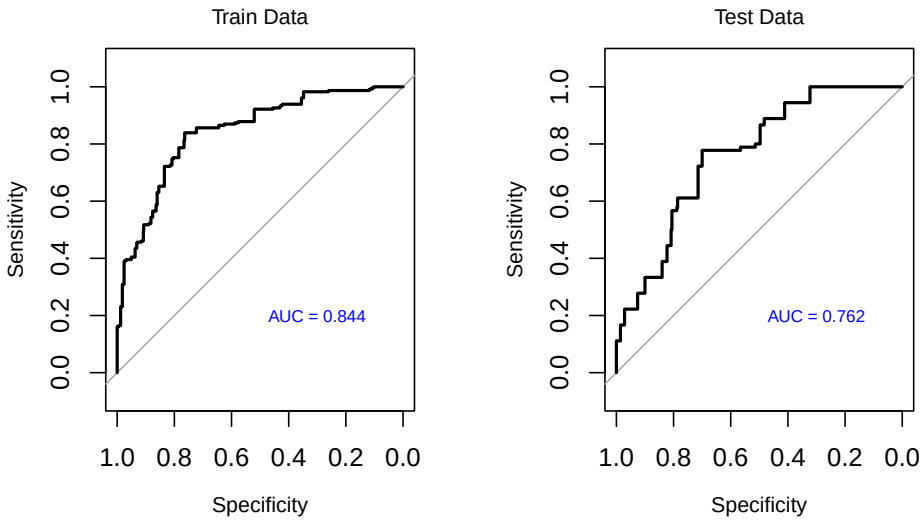
Table 3: Participant Characteristics by Treatment Arm

Characteristic	**Behavioral and Pharmacological Treatment Assignment**				
	ST + placebo, N = 68	**BASC + placebo, N = 68**	**BASC + varenicline, N = 83**	**ST + varenicline, N = 81**	**Overall, N = 300**
Smoking abstinence	8 (12%)	4 (5.9%)	26 (31%)	26 (32%)	64 (21%)
Age	50 (11)	51 (14)	50 (13)	49 (13)	50 (13)
Sex					
Male	29 (43%)	30 (44%)	39 (47%)	37 (46%)	135 (45%)
Female	39 (57%)	38 (56%)	44 (53%)	44 (54%)	165 (55%)
Income					
Less than $20,000	26 (38%)	25 (37%)	30 (37%)	29 (36%)	110 (37%)
$20,000-35,000	14 (21%)	16 (24%)	17 (21%)	21 (26%)	68 (23%)
$35,001-50,000	14 (21%)	8 (12%)	13 (16%)	11 (14%)	46 (15%)
$50,001-75,000	8 (12%)	12 (18%)	12 (15%)	6 (7.5%)	38 (13%)
More than $75,000	6 (8.8%)	6 (9.0%)	10 (12%)	13 (16%)	35 (12%)
Missing	0	1	1	1	3
Education					
Grade School	0 (0%)	1 (1.5%)	0 (0%)	0 (0%)	1 (0.3%)
Some high school	2 (2.9%)	3 (4.4%)	7 (8.4%)	4 (4.9%)	16 (5.3%)
High school graduate or GED	11 (16%)	23 (34%)	15 (18%)	27 (33%)	76 (25%)
Some college/technical school	38 (56%)	22 (32%)	32 (39%)	24 (30%)	116 (39%)
College graduate	17 (25%)	19 (28%)	29 (35%)	26 (32%)	91 (30%)
FTCD score	5 (2)	5 (2)	5 (2)	5 (2)	5 (2)
Missing	1	0	0	0	1
Smoking within 5 mins of waking up	35 (51%)	32 (47%)	33 (40%)	38 (47%)	138 (46%)
BDI score	18 (11)	19 (12)	18 (11)	20 (12)	19 (11)
Cigarettes smoked per day	15 (7)	16 (9)	16 (9)	14 (7)	15 (8)

Table 3: Participant Characteristics by Treatment Arm *(continued)*

Characteristic	Behavioral and Pharmacological Treatment Assignment				
	ST + placebo, N = 68	**BASC + placebo, N = 68**	**BASC + varenicline, N = 83**	**ST + varenicline, N = 81**	**Overall, N = 300**
Cigarette reward value	7 (4)	7 (4)	7 (4)	7 (3)	7 (4)
Missing	8	1	3	6	18
Pleasurable events (substitute reinforcers)	21 (20)	23 (20)	23 (19)	23 (19)	23 (20)
Pleasurable events (complementary reinforcers)	27 (20)	28 (22)	22 (17)	25 (19)	25 (19)
Anhedonia	3 (3)	2 (3)	2 (3)	2 (3)	2 (3)
Missing	1	2	0	0	3
Other lifetime DSM-5 diagnosis	28 (41%)	35 (51%)	30 (36%)	40 (49%)	133 (44%)
Taking antidepressant	15 (22%)	28 (41%)	24 (29%)	15 (19%)	82 (27%)
Current vs. past MDD					
Past MDD	37 (54%)	36 (53%)	43 (52%)	37 (46%)	153 (51%)
Current MDD	31 (46%)	32 (47%)	40 (48%)	44 (54%)	147 (49%)
Nicotine metabolism ratio	0.37 (0.27)	0.34 (0.18)	0.38 (0.25)	0.36 (0.21)	0.36 (0.23)
Missing	2	7	3	9	21
Exclusive mentholated cigarette user	43 (64%)	40 (59%)	48 (59%)	47 (58%)	178 (60%)
Missing	1	0	1	0	2
Readiness to quit smoking	7 (1)	7 (1)	7 (1)	7 (1)	7 (1)
Missing	4	4	5	4	17
Race					
Black	40 (59%)	36 (53%)	36 (43%)	43 (53%)	155 (52%)
Hispanic	4 (5.9%)	4 (5.9%)	3 (3.6%)	5 (6.2%)	16 (5.3%)
Mixed Race	0 (0%)	1 (1.5%)	1 (1.2%)	0 (0%)	2 (0.7%)
Non-Hispanic White	22 (32%)	24 (35%)	34 (41%)	25 (31%)	105 (35%)
Unknown	2 (2.9%)	3 (4.4%)	9 (11%)	8 (9.9%)	22 (7.3%)

[1] Mean (SD) for continuous; n (%) for categorical



Discussion