Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature: _____ _____

Christina Chapman Date

Closing the Gap in Alzheimer's for African Americans: Analysis of the Forget Me Not Project Questionnaire

By

Christina Chapman
Master of Science in Public Health
Biostatistics

_____

Renee' Moore, Ph.D.
Thesis Advisor

_____

Glenda Wrenn Gordon, MD., MSHP
Thesis Collaborator and Reader

Closing the Gap in Alzheimer's for African Americans: Analysis of the Forget Me Not Project Questionnaire

By

Christina Chapman
B.A.
Washington University in St. Louis
2014

Thesis Committee Chair: Renee' Moore, Ph.D.

An abstract of
A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
In partial fulfillment of the requirements for the degree of
Master of Science in Public Health
in Biostatistics
2016

Abstract

Closing the Gap in Alzheimer's for African Americans:
Analysis of the Forget Me Not Project Questionnaire

By Christina Chapman

Older African Americans are about twice as likely to have Alzheimer's disease as older whites. Through the collaborative efforts of the African American Network Against Alzheimer's (AANAA) and playwright Garrett Davis, the Forget Me Not Project works to spread awareness of the growing burden of Alzheimer's in African American communities and to motivate members of minority groups to be more involved in Alzheimer's research and advocacy opportunities. Using the responses to the questionnaire distributed at five 2015 Forget Me Not Project events, this study has three main objectives: (1) we seek to summarize how questions in the survey were answered and to summarize the available demographic information; (2) we aim to determine whether questions considered to be indicators of program success are associated with the demographic information in the survey population; (3) we seek to compare methods treating the Likert-scale outcome variables as ordinal, nominal, or binary to determine the best fit for this dataset. To accomplish these objectives, we report descriptive statistics, conduct univariate analyses, and utilize multivariable logistic model selection procedures. The vast majority of program attendees were female and African American. The participants almost unanimously reported that they enjoyed the program, but often responses to other questionnaire items varied by survey site. Within the study population, the extent to which an individual agreed that they learned why older African Americans need to be involved in research was found to be associated with race and the survey location. Also, the extent to which an individual agrees that they would like to learn more about clinical trials was found to be associated with participants' age, survey location, and the participants' willingness to advocate for Alzheimer's research. To find these associations, models treating Likert-scale outcome variables as binary were found to be the most appropriate. In future studies, we suggest implementing changes to the questionnaire, addressing the limitations of the study, and expanding the program to a wider net of communities.

Closing the Gap in Alzheimer's for African Americans: Analysis of the Forget Me Not Project Questionnaire

By
Christina Chapman
B.A.
Washington University in St. Louis
2014

Thesis Committee Chair: Renee' Moore, Ph.D.

A thesis submitted to the faculty of the
Rollins School of Public Health of Emory University
In partial fulfillment of the requirements for the degree of
Master of Science in Public Health
in Biostatistics
2016

Acknowledgements

I would like to thank my thesis advisor, Dr. Renee' Moore, who so generously provided her guidance and support in the completion of this thesis.

I would also like to thank Dr. Glenda Wrenn Gordon for accepting me into this project and for giving me the opportunity to study the inspiring Forget Me Not program.

Lastly, I thank my grandfather, Marion E. Tuteral, for being a light in my life and for being the motivation behind my work on this project. I will never forget your love and kindness. The world is a better place having had you.

Contents

List of Tables

1 Introduction

1.1 Alzheimer's in the United States

Alzheimer's disease, a degenerative brain disease, is the sixth leading cause of death in the United States. It is the most common cause of dementia - accounting for an estimated 60-80 percent of dementia cases (Alzheimer's Association, 2015). "Dementia is a general term for the loss of memory and other intellectual abilities serious enough to interfere with daily life" (Basics, n.d.).

To summarize the current knowledge of Alzheimer's disease, the Alzheimer's Association produced the report *2015 Alzheimer's Disease Facts and Figures.* Based on this report, an estimated 5.4 million Americans have Alzheimer's disease. It also reports that one in nine people age 65 and older (11 percent) suffer with the disease. Finally, it is reported that an estimated 700,000 people in the United States age 65 and older died with Alzheimer's disease in the year 2015 (Alzheimer's Association, 2015).

While demonstrating the high prevalence and mortality in the United States, these statistics alone do not fully capture the burden of Alzheimer's disease. A person may experience several years of morbidity as the disease progresses. Worsening over time, the disease advances through three stages of illness: Mild, Moderate, and Severe Alzheimer's. While in the early-stage of the disease, an individual may be able to function independently, but may experience memory-lapses that are noticeable to others. Typically individuals will require a greater level of care as they advance into the later stages. In the middle-stage, a person may confuse their words, have bursts of anger, or act in unusual ways. These symptoms are typically coupled with difficulty expressing thoughts and performing routine tasks. Lastly, "in the final stage of the disease, individuals lose the ability to respond to their environment, to carry on a conversation and, eventually, to control movement" (Basics, n.d.). These stages are meant to provide

general guidelines of how abilities change after symptoms appear. However, the disease affects people in different ways, and the rate of progression varies (Basics, n.d.).

While the exact cause of Alzheimer's in unknown, scientists agree that the disease's mechanism involves the prevention of proper nerve cell function. One theory hinges on the fact that individuals with the disease tend to develop far more plaques (deposits of a protein fragment, beta-amyloid, in the spaces between nerve cells) and tangles (twisted fibers of the tau protein inside cells) in the brain than individuals without the illness. Other theories claim that genetic mutation may be the cause of nerve cell degradation (Alzheimer's Association, 2015). Regardless of the pathway leading to damage, any destruction or death of nerve cells can lead to "memory failure, personality changes, problems in carrying out daily activities and other symptoms of Alzheimer's disease" (Basics, n.d.).

Just as the precise cause of Alzheimer's has yet to be identified, there is currently no cure for the disease and no method of preventing further deterioration of brain cells. In fact, "Alzheimer's is the only disease in the top ten (causes of death in the United States) with no disease modifying treatment of cure" (Alzheimer's Disease, n.d.). Working to further Alzheimer's research and to discover these unknowns, scientists are constantly conducting new studies. By testing new theories in clinical studies, the origin of the disease and its remedy may be found (Basics, n.d.). Also while there is no cure, both pharmacologic and non-pharmacologic treatments may alleviate cognitive and behavioral symptoms of the disease (Alzheimer's Association, 2015).

As a slow progressing, incurable illness, a person with Alzheimer's can live for years with symptoms. "On average, a person with Alzheimer's lives four to eight years after diagnosis, but can live as long as 20 years, depending on other factors" (Basics, n.d.). For this reason, and because individuals progressing within the later stages of the disease require extensive care to survive, supporting an Alzheimer's patient can be both financially and emotionally straining. It costs approximately $20,000 a year to care for a patient with Alzheimer's, and the cost of Alzheimer's and other dementias in the entire United States is as much as $215 billion annually (Alzheimer's Disease, n.d.).

1.2 Alzheimer's Disproportionate Effect on African Americans

While the exact cause of Alzheimer's disease and its cure remain undiscovered, researchers have determined risk factors that are associated with the development of the disease. There are many such risk factors, the greatest of them being age. Alzheimer's is not a normal part of aging, and age alone is not sufficient to cause the disease. Other risk factors include family history, cardiovascular disease risk factors, education, and race (Alzheimer's Association, 2015).

Race is a particularly influential risk factor, and the unequal effect of Alzheimer's on the African American community is the motivation behind this initiative. Minorities, particularly African Americans, have a disproportionately high prevalence of Alzheimer's in their communities. "Age-specific prevalence of dementia has been found to be 14% to 100% higher in African Americans" (African-Americans and Alzheimer's disease, n.d.). In other words, older African Americans are about twice as likely to have Alzheimer's and other dementias as older whites. Additionally, several studies indicate that "African Americans with Alzheimer's disease or other dementias are less likely than whites to have been diagnosed with the condition" (Alzheimer's Discriminates, n.d.) (Alzheimer's Disease, n.d.).

The causes of this disproportional influence of Alzheimer's disease on the African American population are uncertain. "While there is some evidence of racial differences in the genetic risk factors, these factors do not appear to account for the large prevalence differences across racial groups" (Alzheimer's discriminates, n.d.). A high prevalence of heart disease, diabetes, and stroke in older African Americans - all of which are known risk factors for Alzheimer's - likely accounts for a portion of this disparity.

Considering these disparities and the aging US population, researchers should focus their efforts especially to serve those that are most affected – African Americans. As the population ages, the

proportion of older individuals at high risk of developing Alzheimer's will increase. "The group at the highest risk – those 85 and older – is the fastest growing population in the US" (Alzheimer's discriminates, n.d.). Alzheimer's will then become a significant challenge to the United States health care system. Consequently, it is now a crucial time for advancing Alzheimer's research and increasing the number participating in clinical trials. Because of the higher prevalence of dementia in their communities, it is especially important for African Americans to be represented in these studies (Alzheimer's discriminates, n.d.).

1.3 African Americans Network Against Alzheimer's (AANAA) and the Forget Me Not Project

The grievous impact Alzheimer's has on African Americans motivated the foundation of the African American Network Against Alzheimer's (AANAA). An affiliate of USAgainstAlzheimer's, AANAA aspires to (1) unify the African American community on the issue of Alzheimer's and other dementias; (2) arm African Americans and others that are committed to their mission with information and connections needed for advocacy; and finally, (3) "engage and connect a network of individuals, businesses, and nonprofit organizations to mobilize the energy of African Americans in advocacy efforts that advance [their] national commitment to ending Alzheimer's" (AANAA, 2012). On the AANAA website, their founders listed the following five ways they planned to meet these aims:

- *"Inform and educate members of the African American community about our dramatically higher risk of developing Alzheimer's disease;"*
- *"Identify and engage political, policy and community leaders about the need for increased resources for Alzheimer's research, drug discovery, care, and services;"*

- *"Mobilize African Americans to become aware of and to participate in studies focused on finding a cure for the disease;"*
- *"Inform business, scientific, and industry leaders, especially leaders of the academic and pharmaceutical communities, about the need for increased research on the race-based differences in the incidence and treatment of Alzheimer's disease; and"*
- *"Speak with employers about the importance of job-friendly eldercare practices for families affected by the disease" (AANAA, 2012).*

The seventeen founders of AANAA are led by four distinguished Honorary Co-Chairs including the Reverend Al Sharpton, Director of the National Action Network; Dr. David Satcher, Former Surgeon General of the United States; The Honorable Kay Coles James, Former Director of the Office of Personnel Management; and Ms. Melody Barnes, Former head of President Obama's Domestic Policy Council (AANAA, 2012).

Under the leadership of Director Stephanie Monroe, the African American Network Against Alzheimer's forged a partnership with award-winning playwright Garrett Davis to support The Forget Me Not Project. The Forget Me Not Project is a program designed to engage and inform all peoples, especially within African American communities, about Alzheimer's disease. "It is the mission of The Forget Me Not Project to raise the awareness of Alzheimer's disease in all communities; to enhance the capacity of and support caregivers while educating the general public with special emphasis on underserved communities where health disparities exist" (Official, n.d.).

In 2015, the Forget Me Not Project held events in five American cities: Atlanta, Chicago, Houston, Indianapolis and Tampa. At each Forget Me Not event, the attendees were introduced to the featured play "Forget Me Not." Written and directed by Garrett Davis, "Forget Me Not" presented a humorous yet touching portrayal of Alzheimer's disease in hopes of entertaining and engaging the audience. During the intermission, an expert panel discussion was held with local doctors, researchers and

former/current African American research participants. This discussion was meant to educate attendees about Alzheimer's and the importance of clinical trial research and research advocacy. The attendees were then allowed time to ask the panel questions. Finally, at the end of each event, attendees were asked to fill out a questionnaire regarding their thoughts on the event and on Alzheimer's disease.

1.4 Purpose of This Study

Our study addresses how attendees responded to the questionnaire at the five 2015 Forget Me Not Project events. Using the responses from the five visited cities (Atlanta, Chicago, Houston, Indianapolis, and Tampa), we question how programs like the Forget Me Not Project can better promote Alzheimer's awareness and generate greater advocacy for Alzheimer's research – especially among African Americans.

Using the responses to the questionnaire, this study has three main objectives. First, we seek to summarize how questions in the survey are answered and to summarize the available demographic information of the survey population. Secondly, we aim to determine whether outcomes of interest are associated with the demographic information. Outcomes of interest include responses to questions regarding willingness to engage in Alzheimer's research, willingness to advocate for this research, and whether participants felt they learned anything from the program. Questions that meet this criteria will be considered indicators of the project's success. The third aim of the study is to determine the best method of analyzing outcome variables that are in a Likert scale. When responding to a Likert-scale questionnaire item, participants report their level of agreement or disagreement on a symmetric agree-disagree scale. In this last aim, we question whether analyzing these variables as ordinal, nominal, or binary is most appropriate.

This study was conducted under the advisement of Renee' Moore, PhD. and in collaboration with Glenda Wrenn Gordon, MD., MSHP, - and Stephanie Monroe. Acting as my thesis advisor, Dr. Moore is a Research Associate Professor in the Department of Biostatistics and Bioinformatics at the Rollins School of Public Health. Having served as a panel member at the Atlanta Forget Me Not event and as a member of the Diversity Council of the National Alzheimer's & Dementia Patient & Caregiver-Powered Research Network, Dr. Glenda Wrenn Gordon is the Director of the Division of Behavioral Health at the Satcher Health Leadership Institute and an Assistant Professor in the Department of Psychiatry and Behavioral Sciences at Morehouse School of Medicine. Having directed the partnership of Forget Me Not and AANAA, Stephanie Monroe is the current Director of African Americans Network Against Alzheimer's.

2 Methods

2.1 Data Collection

The data used in this study is derived from the questionnaires collected at the five Forget Me Not Project event locations. Of 3699 registrants, 1248 questionnaires were collected. The questionnaires comprise of 10 multiple choice questions, two short answer questions, and asks for basic demographic and contact information – the participants name, phone number, email address, race, gender and date of birth. The age of each participant is calculated by subtracting the date of birth from the appropriate event date. For reference, the questionnaire distributed in Indianapolis can be found in its entirety in **Appendix A**.

A complication arose while compiling the questionnaire data. When asked for their date of birth, the participants were to list their month, day, and year of birth. Often those who were born 1900-1915 or 2000-2015 listed their year of birth as ’00-’15. Because it was not clear in many cases to which century they referred, these individuals were removed from the study. It was then decided that the participants should be at least 18 years of age in the study. Therefore, the study was restricted to those between 18-99 years of age. This restriction left 1050 individual questionnaires in the study – 348 questionnaires were collected from Atlanta; 386 from Chicago, 128 from Houston, 88 from Indianapolis, and 100 from Tampa.

While the short answer questions are not used in this study, all 10 multiple questions are used. The 10 multiple choice questions used are listed along with the possible responses in **Table 1**.

Table 1 – Summary of Survey Questions		
Question Number	**Question Content**	**Available Answers**
Question 1	How did you hear about the event?	[ ] Church [ ] Event [ ] Friend [ ] Word of Mouth [ ] Flier [ ] EventBrite [ ] Postcard [ ] Other [ ] Website
Question 2	Did you enjoy the play?	[ ] Yes [ ] No
Question 3	Did you enjoy the panel discussion?	[ ] Yes [ ] No
Question 4	Was the event format and length appropriate?	[ ] Yes [ ] No
Question 5	Was the information helpful to you?	[ ] Yes [ ] No
Question 6	To what extent do you agree or disagree with the following statements? "As a result of attending this event, I have a better understanding of why older African Americans need to get involved in Alzheimer's research."	[ ] Strongly Disagree [ ] Disagree [ ] Neither Disagree or Agree [ ] Agree [ ] Strongly Agree
Question 7	To what extent do you agree or disagree with the following statements? "I would like to learn more about clinical trials and other opportunities that may be available to me."	[ ] Strongly Disagree [ ] Disagree [ ] Neither Disagree or Agree [ ] Agree [ ] Strongly Agree
Question 8	Would you like to receive information about actions you can take to persuade Congress to provide more funding for Alzheimer's?	[ ] Yes [ ] No
Question 9	Do you want to be contacted about other events?	[ ] Yes [ ] No
Question 10	Every 3-6 months, would you be willing to go to the Brain Health Registry website, play simple online brain games and provide basic health and lifestyle information to help researchers understand the brain as it develops and ages? We will send you information on how if you check 'yes'.	[ ] Yes [ ] No

2.2 Outcomes of Interest

As stated previously, the primary objective of this study is to determine how attendees responded to questions regarding willingness to engage in Alzheimer's research, willingness to advocate for this research, and whether participants felt they learned anything from the program.

Out of the 10 multiple choice questions in the survey, four questions meet this criteria and are therefore treated as indicators of the project's success – Question 6, 7, 8 and 10. These questions will be treated as outcomes of interest for the remainder of the study. Question 6 addresses whether participants felt they learned anything from attending this event. Question 7 and Question 10 address whether they would like to learn about ways to get involved in clinical trials. Question 8 addresses whether a participant would like to learn about ways to advocate for Alzheimer's research funding.

While Question 8 and Question 10 have binary (yes or no) responses, responses to Question 6 and Question 7 are in a five-member Likert scale (strongly agree, agree, neither disagree or agree, agree, or strongly agree). In this study, the analysis is conducted using both the original, Likert-scale form of these questions and a new binary form. To transform these variables to binary, responses of "Strongly Disagree", "Disagree", and "Neither Disagree or Agree" are collapsed into one "Do Not Agree" category. Responses of "Agree" and "Strongly Agree" are collapsed into one "Agree" category. A summary of questions treated as outcomes of interest is presented in **Table 2**.

Table 2 – Summary of Outcomes of Interest			
Question Number	**Question Content**	**Possible Responses**	
		Likert-Scale Form	**Binary Form**
Question 6	To what extent do you agree or disagree with the following statement? "As a result of attending this event, I have a better understanding of why older African Americans need to get involved in Alzheimer's research."	[ ] Strongly Disagree [ ] Disagree [ ] Neither Disagree or Agree [ ] Agree [ ] Strongly Agree	[ ] Do Not Agree [ ] Agree
Question 7	To what extent do you agree or disagree with the following statement? "I would like to learn more about clinical trials and other opportunities that may be available to me."	[ ] Strongly Disagree [ ] Disagree [ ] Neither Disagree or Agree [ ] Agree [ ] Strongly Agree	[ ] Do Not Agree [ ] Agree
Question 8	Would you like to receive information about actions you can take to persuade Congress to provide more funding for Alzheimer's?	NA	[ ] Yes [ ] No
Question 10	Every 3-6 months, would you be willing to go to the Brain Health Registry website, play simple online brain games and provide basic health and lifestyle information to help researchers understand the brain as it develops and ages? We will send you information on how if you check 'yes'.	NA	[ ] Yes [ ] No

2.3 Statistical Methods

To summarize how all questions and demographic information in the survey are answered – which is the first aim of this study, a descriptive analysis is conducted. This analysis is conducted for the entire study population and then again for each program location.

To then address the second aim, univariate analyses and multivariable binary logistic regression analyses are conducted using the binary forms of the outcomes of interest seen in **Table 2**. It is the purpose of these analyses to determine if the binary responses of Questions 6, 7, 8, and 10 are associated

with demographic information collected in the study. Also of interest, we can determine whether the response to Questions 6, 7, and 10 are associated with Question 8 responses – representing an individual's motivation to advocate for Alzheimer's research after attending the program. The univariate analyses include Chi-Square tests, Fisher's exact tests, Student t-tests, and univariate logistic regression.

To further meet the second aim of this study, a model selection procedure is conducted to find the best multivariable binary logistic regression model for each of the four binary outcomes of interest. All multivariable binary logistic models consider the available demographic information and, except for the model containing Question 8 as an outcome, the responses to Question 8 as a potential covariate.

Finally, the last portion of the analysis – meeting the third and final aim of the study - is a comparison of three multivariable logistic regression models for both Question 6 and Question 7: a binary logistic model, an ordered logistic model, and a nominal logistic model. Using the Likert forms of Questions 6 and 7, model selection procedures are conducted to find the best multivariable ordered and multivariable nominal logistic models for these two outcomes. The ordered logistic regression model treats the Likert scale responses as ordinal variables. Nominal logistic regression models treat the Likert scale responses as nominal, unordered variables. Having found the best multivariable ordered and generalized logistic models for both Question 6 and Question7, the associations found in these models and each model's criteria measuring goodness of fit are compared to those found in the multivariable binary logistic models.

2.3.1 Descriptive Analysis

The purpose of this first step of the statistical analysis is to understand how the program participants respond to all questionnaire items. These items include all available demographic information – participants' race, gender, age, and program (survey) location – and the responses to Question 1-Question 10. The frequency of each available response and the percent of people who selected a given

response are listed for the entire study population. This procedure is then repeated within each survey location. Additionally, the distribution of the continuous age variable is described via characteristics such as skewness and kurtosis. Age is summarized via measures of center (mean, median) and measures of spread (standard deviation, interquartile range). These descriptive statistics are summarized in **Table 3** of the Results section.

2.3.2 Univariate Analyses

In this second step of the analysis, the aim is to determine whether our outcomes of interest (Questions 6, 7, 8, and 10) are associated with other questionnaire items. Specifically, we seek to determine whether the participants' gender, age, and survey location are significantly associated with the outcomes. In addition, we are interested in whether the participants' motivation to advocate for the program's cause (represented by the responses to Question 8) are associated with Question 6, 7, and 10.

For this portion of the analysis, only the binary forms of Questions 6, 7, 8, and 10 are used. Questions 8 and 10 are originally binary. Questions 6 and 7, on the other hand, are transformed as such– if the respondents answer these questions in a way that conveys agreement with responses "Agree" or "Strongly Agree", this suggests program success. If they did not agree with responses "Strongly Disagree", "Disagree", or "Neither Disagree or Agree", than the program was not successful. A binary split of Likert scale variables may eliminate challenges when interpreting results.

As stated, the independent variables being examined are participants' race, gender, age, survey site, and response to Question 8. All of these variables are categorical except for age which is continuous. For the remainder of the study, a binary version of the race variable is used; instead of having three categories (African American, Caucasian, and Other), we collapse this variable into two categories (African American, Not African American). This is done because only a small proportion of study participants claim to be "Caucasian" (5.2%) or "Other" (2.3%). Also, within the survey site variable,

because Atlanta was the location of the first Forget Me Not program, those surveyed in Atlanta are treated as the reference group.

For all categorical variables, either a Chi-Square test or Fisher's exact test is conducted between each outcome of interest (Question 6, 7, 8, and 10) and each independent variable. Both Chi-Square and Fisher's exact tests are used when testing whether the proportion of participants falling into each category of a nominal variable differ when comparing these values across another nominal variable. Within a contingency table, a chi-square test can only be used if more than 80% of the contingency cells have expected values greater than five. If this assumption is not met, the Fisher's exact test is more appropriate and is then used.

For the continuous age variable, if found to be approximately normally distributed, a two-sample independent Student t-test is used to examine the association between age and the outcomes of interest. Further, if normally distributed and the Equality of Variance test is not significant, then we use the Pooled Student-t test. If normally distributed and the Equality of Variance test is significant, then we use the Satterthwaite Student-t test. If the age variable is not found to be normally distributed, a nonparametric alternative, the two-sample Wilcoxon Rank-Sum test will be used. Both the Student t-test and the Wilcoxon test have a null hypothesis that the group means are equal; the Wilcoxon test utilizes the sample medians for comparison.

As the last step in the univariate analysis, we fit univariate binary logistic regression models to predict the binary outcomes of interest using each of the possible predictor variables. The univariate binary logistic model is used to estimate the probability of a binary response based on one independent predictor variable. These models generate odds ratios. An odds ratio is the odds of an event occurring in one group of the predictor variable divided by the odds of an event occurring in another group.

2.3.3 Multivariable Binary Logistic Regression Model Selection

For the next step of the analysis, we conduct a multivariable binary logistic regression model selection procedure to find the best possible model for the binary versions of each outcome of interest (Question 6, 7, 8, and 10). The purpose of this analysis is to determine whether there is a statistically significant association between the independent variables and Questions 6, 7, 8, and 10, all while controlling for other significant variables.

To execute the model selection procedure, we first use the univariate logistic regression models to find the most significant predictor of the binary version of each outcome of interest. All possible two-variable models that contain the most significant predictor are then fit, and the most significant of the two-variable models is determined. All possible three-variable models that contain these two variables are fit. The most significant variables are added one-by-one into the model until all of the variables have been examined. If during this procedure, there are no additional significant variables with an alpha level of 0.05, then the selection procedure ceases, and the last model with all significant predictors is the final, best model.

After conducting this step-by-step selection procedure, all final models are compared to SAS's automated forward, backward, and stepwise selection procedures.

2.3.4 Comparison of Three Multivariable Logistic Regression Models

Finally, both multivariable ordered and multivariable nominal logistic regression model selection procedures are conducted to find the best possible multivariable logistic models for the Likert-scale versions of Questions 6 and 7. The purpose of this step of the analysis is to compare the results of these models with those of the multivariable binary logistic regression. We seek to determine the best method

of analyzing the Likert-scale variables in the questionnaire – whether that is by transforming these responses into binary variables as in previous steps of the analysis, treating them as ordinal variables, or treating them as nominal variables.

An ordered logistic regression model, also known as the proportional odds model, is a model for dependent, ordinal variables, as opposed to a binary logistic regression models that model dependent, binary variables. While this model can make fuller use of ordinal information, the ordered logic regression model can only be used if the proportional odds assumption is met. The proportional odds assumption states that the odds ratios assessing the effect of an exposure variable over a one-level increase in the ordinal variable is the same regardless of what two levels are compared. This assumption is examined with a Score test whose null hypothesis is that the proportional odds assumption is met (Kleinbaum, 2010).

A nominal logistic regression model, also known as a generalized model, is a model for dependent, nominal variables with no natural ordering. Although this does not incorporate the ordinal information found in the Likert-scale variables, nominal logit models may, in many cases, be superior given that they are less restrictive than ordered logistic models. Unlike the ordered logistic model, the nominal logistic model does not have to follow the proportional odds assumption. Therefore, the odds ratios assessing the effect of an exposure variable over a one-level increase in the ordinal variable are free to differ across each comparison (Kleinbaum, 2010).

Both the multivariable ordered logistic and the multivariable nominal logistic model selection procedures are conducted similarly to the multivariable binary logistic regression model. Just as before, only the most significant variables with an alpha level of 0.05 remain in the model. The most significant variables are added one-by-one into the model until we go through all of the variables. If during this procedure, there are no additional significant variables with an alpha level of 0.05, then the selection procedure ceases, and the last model with all significant predictors is the final, best model. All final models are compared to SAS's automated forward, backward, and stepwise selection procedures.

After fitting a model for each of the binary, ordinal, and nominal versions of Questions 6 and 7, the three models for each question are compared. First, a comparison of variables found to be significant in each model is conducted – these comparisons are summarized in **Table 11**. Then, a comparison of criteria describing each model's goodness of fit is conducted. Criteria describing goodness of fit include model AIC, the Likelihood Ratio (LR) test statistic, corresponding LR p-value, and the model Schwarz Criterion (SC). Model AIC, or the Akaike Information Criterion, offers a relative estimate of information lost when a given model in used. In this way, AIC estimates the quality of each model – the lower the AIC, the better the model fits the data. The Likelihood Ratio test statistic is a Chi-Squared statistic that measures how well the independent variables affect the outcome or dependent variable. The larger the LR test statistic, the larger the affect the independent variables affect the outcome, and the better the model fits the data. The LR test statistic corresponds to a p-value. If this p-value is below 0.05, then there is evidence that at least one of the independent variables contributes to the prediction of the dependent, outcome variable. Lastly, the Schwarz Criterion (SC), or the Bayesian Information Criterion (BIC), is similar to the AIC in that it offers a relative estimate of information lost, but the penalty for overfitting a model is larger when calculating SC than AIC. The lower the SC, the better the model fits the data.

2.4 Statistical Analysis Software (SAS)

All data are analyzed in SAS version 9.4 statistical software. All SAS code and output used in this methodology can be found in the **Appendix B**.

3 Results

3.1 Results of Descriptive Analysis

For the first portion of the analysis, the aim is to summarize how questions in the survey are answered and to summarize the available demographic information of the survey population. The results of the descriptive analysis of all items in the Forget Me Not Project questionnaire are summarized in **Table 3**.

In **Table 3**, the variables listed along the left side of the table represent items answered in the questionnaire – listed directly below each variable are the possible responses to that item. The first column lists the overall count and the overall percentage of people who gave a particular response to each item. The next four columns list the count and the percentage of people who gave a particular response at each program location (Atlanta, Chicago, Houston, Indianapolis, and Tampa). At the bottom of the table, the descriptive statistics of the only continuous variable in the questionnaire, age, are listed. The mean, standard deviation, median, the 25th and 75th percentiles, and the range of age are listed for the overall study population and by program location.

Table 3 – Descriptive Statistics

	Overall N (%)	**Site**				
		Atlanta N (%)	**Chicago N (%)**	**Houston N (%)**	**Indianapolis N (%)**	**Tampa N (%)**
Gender						
Female	903 (87.8)	285 (83.1)	347 (91.8)	110 (90.2)	76 (88.4)	85 (85.0)
Male	126 (12.2)	58 (16.9)	31 (8.20)	12 (9.84)	10 (11.6)	15 (15.0)
Question 1						
Church	215 (21.3)	80 (24.1)	102 (27.7)	9 (7.14)	22 (25.3)	2 (2.02)
Friend	356 (35.2)	76 (22.3)	142 (38.6)	56 (44.4)	33 (37.9)	49 (49.5)
Flier	37 (3.66)	12 (3.61)	20 (5.43)	1 (0.79)	3 (3.45)	1 (1.01)
Postcard	8 (0.79)	5 (1.51)	3 (0.82)	0 (0.0)	0 (0.0)	0 (0.0)
Website	21 (2.08)	8 (2.41)	5 (1.36)	1 (0.79)	2 (2.30)	5 (5.05)
Event	55 (5.43)	43 (13.0)	7 (1.90)	2 (1.59)	2 (2.30)	1 (1.01)
Word of Mouth	41 (4.05)	17 (5.12)	7 (1.90)	10 (7.94)	3 (3.45)	4 (4.04)
EventBrite	37 (3.66)	16 (4.82)	9 (2.45)	2 (1.59)	2 (2.30)	8 (8.08)
Other	242 (23.9)	75 (22.6)	73 (19.8)	45 (35.7)	20 (23.0)	29 (29.3)
Question 2						
Yes	1026 (99.7)	340 (99.7)	381 (100)	126 (100)	88 (100)	91 (97.9)
No	3 (0.29)	1 (0.29)	0 (0.0)	0 (0.0)	0 (0.0)	2 (2.15)

Question 3						
Yes	841 (99.6)	261 (100)	312 (99.7)	103 (99.0)	81 (100)	84 (98.8)
No	3 (0.36)	0 (0.0)	1 (0.32)	1 (0.96)	0 (0.0)	1 (1.18)
Question 4						
Yes	903 (96.9)	306 (98.4)	328 (97.3)	113 (96.6)	76 (95.0)	80 (92.0)
No	29 (3.11)	5 (1.61)	4 (3.42)	4 (3.42)	4 (5.00)	7 (8.05)
Question 5						
Yes	958 (100)	319 (100)	343 (100)	123 (100)	81 (100)	92 (100)
No	0 (0)	0 (0.0)	0 (0.0)	0 (0.0)	(0.0)	0 (0.0)
Question 6						
Strongly Disagree	85 (8.73)	67 (20.7)	0 (0.0)	1 (0.81)	0 (0.0)	17 (17.9)
Disagree	2 (0.21)	2 (0.62)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Neither Disagree or Agree	21 (2.16)	11 (3.41)	4 (1.13)	2 (1.63)	3 (3.80)	1 (1.05)
Agree	244 (25.1)	91 (28.2)	77 (21.8)	20 (16.3)	26 (32.9)	30 (31.6)
Strongly Agree	622 (63.7)	152 (47.1)	273 (77.1)	100 (81.3)	50 (63.3)	47 (49.5)
Question 7						
Strongly Disagree	63 (7.07)	46 (15.0)	4 (1.28)	2 (1.72)	0 (0.0)	11 (12.4)
Disagree	15 (1.68)	6 (1.95)	4 (1.28)	1 (0.86)	0 (0.0)	4 (4.49)
Neither Disagree or Agree	89 (9.99)	27 (8.79)	26 (8.31)	16 (13.8)	6 (9.09)	14 (15.7)
Agree	310 (34.8)	111 (36.2)	102 (32.6)	38 (32.8)	27 (40.9)	32 (36.0)
Strongly Agree	414 (46.5)	117 (38.1)	177 (56.6)	59 (50.9)	33 (50.0)	28 (31.5)
Question 8						
Yes	920 (91.7)	316 (92.7)	333 (91.2)	114 (92.7)	76 (97.4)	81 (84.4)
No	83 (8.28)	25 (7.33)	32 (8.77)	9 (7.32)	2 (2.56)	15 (15.6)
Question 9						
Yes	816 (84.2)	285 (84.8)	269 (79.6)	108 (90.0)	73 (90.1)	81 (86.2)
No	153 (15.8)	51 (15.2)	69 (20.4)	12 (10.0)	8 (9.88)	13 (13.8)
Question 10						
Yes	770 (81.7)	256 (80.8)	278 (81.8)	94 (82.5)	61 (80.3)	81 (85.3)
No	172 (18.3)	61 (19.2)	62 (18.2)	20 (17.5)	15 (19.7)	14 (14.7)
Race						
African American	949 (92.5)	336 (98.3)	355 (95.4)	112 (88.2)	67 (77.0)	79 (80.6)
Caucasian	53 (5.17)	6 (1.75)	8 (2.15)	8 (6.30)	17 (19.5)	14 (14.3)
Other	24 (2.34)	0 (0.0)	9 (2.42)	7 (5.51)	3 (3.45)	5 (5.10)
Survey Site						
Atlanta	348 (33.1)	----	----	----	----	----
Chicago	386 (36.8)	----	----	----	----	----
Houston	128 (12.2)	----	----	----	----	----
Indianapolis	88 (8.38)	----	----	----	----	----
Tampa	100 (9.52)	----	----	----	----	----

	Overall	**Site**				
		Atlanta	**Chicago**	**Houston**	**Indianapolis**	**Tampa**
Age						
Mean (SD)	63.0 (14.3)	59.9 (13.4)	69.5 (11.7)	59.4 (14.0)	60.9 (16.9)	55.9 (15.7)
Median	65.0	62.5	71.0	62.0	64.0	57.5
(Q1, Q3)	(55.0, 72.0)	(52.5, 69.0)	(64.0, 78.0)	(52.0, 68.0)	(53.0, 72.0)	(45.0, 67.0)
(Min, Max)	(18.0, 99.0)	(18.0, 88.0)	(22.0, 96.0)	(21.0, 85.0)	(18.0, 99.0)	(20.0, 94.0)

Seen in **Table 3**, 87.8% of participants are female, and this number is fairly consistent across program sites – ranging from 83.1% in Atlanta to 91.8% in Chicago. The responses to Question 1, representing how participants claimed they heard about the program, varied drastically across the program

locations. For example, 49.5% of Tampa participants said they heard about the program from a friend while only 22.3% of Atlanta participants said the same. Almost unanimously, participants answered "yes" for Questions 2-5 – meaning that they enjoyed the play, enjoyed the panel discussion, thought that the event length and format were appropriate, and thought the information presented to them was helpful. Over 60% of participants strongly agreed that after the program they understood why older African Americans should be involved in Alzheimer's research (ranging from 47.1% in Atlanta to 81.3% in Houston). The number of participants who said they would like to learn more about opportunities to get involved in clinical trials varied from 56.6% in Chicago to 31.5% in Tampa. Over 75% of participants answered "yes" they would like to learn more about future events, ways to advocate for Alzheimer's funding, and about the Brain Health Registry. 92.5% participants were African American (ranging from 77.0% of Indianapolis participants to 98.3% of Atlanta participants), and the mean age was 63.0 years old.

3.2 Results of Univariate Analysis

The purpose of the univariate analysis is to determine whether Questions 6, 7, 8, and 10 are associated with gender, race, and survey site. We also aim to determine whether responses to Question 8 are significantly associated with Questions 6, 7, and 10.

Table 4A displays the univariate analyses for Question 6; **Table 4B** displays the univariate analyses for Question 7; **Table 4C** displays the univariate analyses for Question 8; and **Table 4D** displays the univariate analyses for Question 10. In all four tables, the possible covariates are listed in the leftmost column. The second column represents the number and percent of individuals answering positively for a given question. The third column represents the odds ratio (OR) of responding positively to the question given a one year increase in age or as compared to reference group for a level of a categorical variable.

The 95% confidence interval of the OR is also included. The fourth column lists the univariate test performed for each covariate, and the last column lists the resulting p-value of this univariate test.

Table 4A – Univariate Analysis for Binary Question 6:
6. To what extend do you agree or disagree with the following statements?
"As a result of attending this event, I have a better understanding of why older African Americans need to get involved in Alzheimer's research."
[ ] Do Not Agree [ ] Agree

	Agree N (%)	**Logistic Regression OR (95% CI)**	**Test**	**P-value**
Age				
18-99 Years Old	NA	1.01 (1.00, 1.03)	Pooled T-Test	0.0931
Gender				
Female	748 (88.9)	1.05 (0.56,1.94)	Chi-Square	0.8873
Male (reference)	100 (88.5)			
Question 8				
Yes	767 (89.4)	1.82 (0.98, 3.37)	Chi-Square	0.0550
No (reference)	65 (82.3)			
Race**				
African American	785 (89.1)	1.48 (0.75, 2.90)	Chi-Square	0.2576
Not African American (reference)	61 (84.7)			
Survey Site				
Atlanta (reference)	243 (75.2)	1.00 (reference)	Fisher's Exact	<.0001
Chicago	350 (98.9)	28.8 (10.4, 79.7)		
Houston	120 (97.6)	13.2 (4.07, 42.6)		
Indianapolis	76 (96.2)	8.34 (2.56, 27.2)		
Tampa	77 (81.1)	1.41 (0.80, 2.50)		

** P-values for Race compare African American to non-African American.

Based on the results found in **Table 4A**, age is not significantly associated with the binary version of Question 6 (p-value = 0.0931) using a pooled t-test. The odds ratio of responding positively to the question given a one year increase in age is 1.01 (95% CI: 1.00-1.03) indicating no effect. Gender and race are not significantly associated with Question 6 through the use of Chi-Square tests. The responses to Question 8 are not significantly associated with Question 6 (p-value=0.0550) with a Chi-Square test. Those that claimed they are willing to learn how to advocate for Alzheimer's research (responding "Yes" to Question 8) have 1.82 (95% CI: 0.98-3.37) times higher odds of responding "Agree" to Question 6 than those that responded "No" to Question 8. The only significant predictor is the survey site with those surveyed in Atlanta as the reference group (p-value < .0001; Fisher's exact test). As an example, those

that were surveyed in Chicago have 28.8 (95% CI: 10.4-79.7) times higher odds of responding "Agree" to Question 6 as those surveyed in Atlanta.

Table 4B – Univariate Analysis for Binary Question 7:
7. To what extent do you agree or disagree with the following statements?
"I would like to learn more about clinical trials and other opportunities that may be available to me."
[ ] Do Not Agree [ ] Agree

	Agree N (%)	**Logistic Regression OR (95% CI)**	**Test**	**P-value**
Age				
18-99 Years Old	NA	1.02 (1.01, 1.03)	Satterthwaite T-Test	0.0017
Gender				
Female	630 (81.5)	1.13 (0.68, 1.88)	Chi-Square	0.6443
Male (reference)	82 (79.6)			
Question 8				
Yes	674 (84.1)	6.59 (3.90, 11.1)	Chi-Square	<.0001
No (reference)	29 (44.6)			
Race**				
African American	665 (82.0)	1.82 (1.03, 3.24)	Chi-Square	0.0385
Not African American (reference)	45 (71.4)			
Survey Site				
Atlanta (reference)	228 (74.3)	1.00 (reference)	Chi-Square	<.0001
Chicago	279 (89.1)	2.84 (1.83, 4.41)		
Houston	97 (83.6)	1.77 (1.02, 3.08)		
Indianapolis	60 (90.9)	3.47 (1.44, 8.33)		
Tampa	60 (67.4)	0.72 (0.43, 1.20)		

** P-values for Race compare African American to non-African American.

As seen in **Table 4B**, a Satterthwaite t-test finds that age is a statistically significant predictor of the binary version of Question 7 (p-value = 0.0017), but with an estimated odds ratio of only 1.02 (95% CI: 1.01-1.03) times the odds of agreement with each one year increase in age. The responses to Question 8, race, and the program location are found to be associated with binary Question 7 (p-values <.0001, =0.0385, and <.0001, respectively). Those that responded "Yes" to Question 8 – representing that they want to learn how to advocate for Alzheimer's research, have 6.59 (95% CI: 3.90-11.1) times higher odds of responding "Agree" compared to those who responded "No" to Question 8. African Americans have 1.82 (95% CI: 1.03-3.24) times higher odds of responding "Agree" to binary Question 7 than those who are not African American. Compared to those surveyed in Atlanta, those surveyed in Indianapolis have

3.47 (95% CI: 1.44-8.33) times higher odds of responding "Agree"; those from Chicago have 2.84 (95% CI: 1.83-4.41) times higher odds; those from Houston have 1.77 (95% CI: 1.02-3.08) times the odds; and Tampa has 0.72 (95% CI: 0.43-1.20) times higher odds of responding "Agree". A Chi-Square test finds that gender is not a significantly associated with the binary responses to Question 7 (p-value = 0.6443).

Table 4C – Univariate Analysis for Question 8:
8. Would you like to receive information about actions you can take to persuade Congress to provide more funding for Alzheimer's?
[ ] Yes [ ] No

	Yes N (%)	**Logistic Regression OR (95% CI)**	**Test**	**P-value**
Age				
18-99 Years Old	NA	1.00 (0.99, 1.02)	Satterthwaite T-Test	0.6550
Gender				
Female	793 (91.8)	1.12 (0.57, 2.17)	Chi-Square	0.7447
Male (reference)	110 (90.9)			
Race**				
African American	837 (92.4)	1.93 (0.95, 3.92)	Chi-Square	0.0664
Not African American (reference)	63 (86.3)			
Survey Site				
Atlanta (reference)	316 (92.7)	1.00 (reference)	Fisher's Exact	0.0360
Chicago	333 (91.2)	0.82 (0.48, 1.42)		
Houston	114 (92.7)	1.00 (0.45, 2.21)		
Indianapolis	76 (97.4)	3.01 (0.70, 13.0)		
Tampa	81 (84.4)	0.43 (0.22, 0.85)		

** P-values for Race compare African American to non-African American.

Seen in **Table 4C**, through the use of a Satterthwaite t-test and a Chi-Square test respectively, both age (p-value = 0.6550) and gender (p-value=0.7447) are found to be insignificant predictors of responses to Question 8. Race is not significant associated with Question 8 through the use of a Chi-Square test (p-value = 0.0664), though a higher percentage of African Americans answered "Yes". African Americans have 1.93 (95% CI: 0.95-3.92) times higher odds of answering "Yes" to Question 8 than those who are not African American. Using a Fisher's exact test, the survey site is determined to have a significant association with Question 8 (p-value = 0.0360). Compared to those attending the Forget Me Not program in Atlanta, those in Chicago are estimated to have 0.82 (95% CI: 0.48-1.42) times lower odds of indicating "Yes"; those in Houston have 1.00 (95% CI: 0.45-2.21) times higher odds; those in

Indianapolis have 3.01 (95% CI: 0.70-13.0) times higher odds; and those in Tampa have 0.43 (95% CI: 0.22-0.85) times lower odds of responding "Yes" to Question 8.

Table 4D – Univariate Analysis for Question 10:

10. Every 3-6 months, would you be willing to go to the Brain Health Registry website, play simple online brain games and provide basic health and lifestyle information to help researchers understand the brain as it develops and ages? We will send you information on how if you check 'yes'.

[ ] Yes [ ] No

	Yes N (%)	**Logistic Regression OR (95% CI)**	**Test**	**P-value**
Age				
18-99 Years Old	NA	1.00 (0.99, 1.01)	Satterthwaite T-Test	0.9105
Gender				
Female	658 (81.5)	1.03 (0.63, 1.70)	Chi-Square	0.8965
Male (reference)	94 (81.0)			
Question 8				
Yes	723 (86.3)	10.1 (6.12, 16.5)	Chi-Square	<.0001
No (reference)	30 (38.5)			
Race**				
African American	704 (82.8)	1.64 (0.93, 2.88)	Chi-Square	0.0836
Not African American (reference)	53 (74.7)			
Survey Site				
Atlanta (reference)	256 (80.8)	1.00 (reference)	Chi-Square	0.8870
Chicago	278 (81.8)	1.07 (0.72, 1.58)		
Houston	94 (82.5)	1.12 (0.64, 1.96)		
Indianapolis	61 (80.3)	0.97 (0.52, 1.82)		
Tampa	81 (85.3)	1.38 (0.73, 2.59)		

** P-values for Race compare African American to non-African American.

Finally, based on the results found in **Table 4D**, the responses to Question 8 are found to be significantly associated with the responses to Question 10 using a Chi-Square test (p-value < .0001). Those that respond "Yes" to Question 8 (representing a willingness to learn how to advocate for Alzheimer's research) have 10.1 (95% CI: 6.12-16.5) times the odds of responding "Yes" to Question 10 compared to those that respond "No" to Question 8. Race is not significantly associated with the responses to Question 10 (p-value = 0.0836). However, African Americans have 1.64 (95% CI: 0.93-2.88) times higher odds of responding "Yes" to Question 10 as compared to those who are not African American. Age (p-value = 0.9105), gender (p-value = 0.8965), and the survey location (p-value = 0.8870) are all not significant predictors of responses to Question 10.

3.3 Results of Binary Logistic Model Selection

The purpose of this portion of the analysis is to address the second aim of this study – to determine whether are outcomes of interest (Questions 6, 7, 8, and 10) are associated with the demographic information. The multivariable binary logistic model selection procedure determines the best possible model predicting the binary versions of all four outcomes of interest. Each outcome's model will determine which potential covariates are significantly associated with the outcome while adjusting for other significant variables.

Tables 5A-5D depict the selection procedures for binary Questions 6, 7, 8, 9 and 10, respectively. In **Tables 5A-5D**, all possible covariates are listed along the leftmost column. The remaining columns summarize the results of each round of the model selection process. For the first round of selection, or "Round 1", only the intercept and each given row's covariate are incorporated in the model. If there is a covariate that results in a significant association – indicated by a Wald Chi-Square p-value of 0.05 or less, the most significant variable is held in the model for the next round. The next round, "Round 2", incorporates the intercept, the most significant covariate in Round 1 and the given row's covariate into the model. Again, if there is a significant covariate, it is retained in the model and the procedure moves on to the next round. This process continues until either there are no significant covariates in a round or until there are no more covariates to incorporate into the model. For each round of selection, the Model AIC and the Wald Chi-Square p-value are listed. If a variable is selected in a round, the variable's p-value is marked with an asterisk and will not be listed again for any subsequent round.

Table 5A – Multivariable Binary Logistic Model Selection Procedure for Question 6:
6. To what extend do you agree or disagree with the following statements?
"As a result of attending this event, I have a better understanding of why older African Americans need to get involved in Alzheimer's research."
[ ] Do Not Agree [ ] Agree

	Round 1 Containing Intercept and Each Given Variable		**Round 2** Containing Intercept, Survey Site, and Each Given Variable		**Round 3** Containing Intercept, Survey Site, Race, and Each Given Variable	
	Model AIC	**Wald P-value**	**Model AIC**	**Wald P-value**	**Model AIC**	**Wald P-value**
Age **18-99**	679.859	0.0935	562.338	0.3103	553.534	0.3571
Gender **Female** **Male**	669.552	0.8873	554.535	0.2400	544.355	0.1106
Question 8 **Yes** **No**	658.158	0.0582	541.670	0.0584	533.662	0.0654
Race** **AA** **Not AA**	672.305	0.2602	552.395	0.0247*	-----	-----
Survey Site **Atlanta** **Chicago** **Houston** **Indianapolis** **Tampa**	561.384	<.0001*	-----	-----	-----	-----

* Represents the significant Wald Chi-Square p-values resulting in a variable being selected.
** P-values for Race compare African American to non-African American.

As seen in **Table 5A**, the final multivariable binary logistic model for Question 6 contains the survey site and race. In the first round of selection, when the survey site is the only variable in the model, survey site is significantly associated with binary Question 6 (p-value < 0.0001). No other covariates are seen to have a significant association in their respective models, and the survey site variable is retained in all "Round 2" models. During the second round, race is the only variable significantly associated with binary Question 6 (p-value = 0.0247), and it is also retained in the model. The third round produces no additional significant variables, so the final model contains only the survey site and race variables determined to be significant in the first two rounds.

Table 5B – Multivariable Binary Logistic Model Selection Procedure for Question 7:

7. To what extent do you agree or disagree with the following statements?

"I would like to learn more about clinical trials and other opportunities that may be available to me."

[ ] Do Not Agree [ ] Agree

	Round 1 Containing Intercept and Each Given Variable		**Round 2** Containing Intercept, Question 8, and Each Given Variable		**Round 3** Containing Intercept, Question 8, Survey Site, and Each Given Variable		**Round 4** Containing Intercept, Question 8, Survey Site, Age, and Each Given Variable	
	Model AIC	**Wald P-value**	**Model AIC**	**Wald P-value**	**Model AIC**	**Wald P-value**	**Model AIC**	**Wald P-value**
Age **18-99**	852.182	0.0006	783.585	0.0004	762.852	0.0415*	-----	-----
Gender **Female** **Male**	848.533	0.6445	780.970	0.8841	752.802	0.4619	750.697	0.4159
Question 8 **Yes** **No**	793.845	<.0001*	------	-----	-----	-----	-----	-----
Race** **AA** **Not AA**	844.043	0.0409	781.386	0.1197	752.139	0.0901	750.680	0.0928
Survey Site **Atlanta** **Chicago** **Houston** **Indianapolis** **Tampa**	831.260	<.0001	764.981	<.0001*	-----	-----	-----	-----

* Represents the significant Wald Chi-Square p-values resulting in a variable being selected.

** P-values for Race compare African American to non-African American.

Based on the results summarized in **Table 5B**, the final multivariable binary logistic model for binary Question 7 contains the Question 8, survey site, and age variables. In the first round of selection, when the Question 8 is the only variable in the model, responses to Question 8 are significantly associated with binary Question 7 (p-value < 0.0001). While age (p-value = 0.0006), race (p-value = 0.0409), and survey site (p-value < 0.0001) are all statistically significant, Questions 8 is the most significant. Therefore the Question 8 variable is retained in all "Round 2" models. During the second round, survey site is the most significant variable – out of the four remaining variables - associated with binary Question 7 (p-value < 0.0001), and it is also retained in the model. In the third round, age is the only significant (p-value = 0.0415) of the remaining three variables, and it is retained. The fourth round produces no

additional significant variables, so the final model contains only the variables selected in the previous rounds: Question 8, survey site, and age.

Table 5C – Multivariable Binary Logistic Model Selection Procedure for Question 8:
8. Would you like to receive information about actions you can take to persuade Congress to provide more funding for Alzheimer's?
[ ] Yes [ ] No

	Round 1 Containing Intercept and Each Given Variable		**Round 2** Containing Intercept, Survey Site, and Each Given Variable	
	Model AIC	**Wald P-value**	**Model AIC**	**Wald P-value**
Age **18-99**	576.316	0.6020	573.444	0.5140
Gender **Female** **Male**	568.564	0.7448	565.776	0.7740
Race** **AA** **Not AA**	550.267	0.0709	548.906	0.0832
Survey Site **Atlanta** **Chicago** **Houston** **Indianapolis** **Tampa**	571.875	0.0419*	-----	-----

* Represents the significant Wald Chi-Square p-values resulting in a variable being selected.
** P-values for Race compare African American to non-African American.

Seen in **Table 5C** and **Table 5D**, both Question 8 and Question 10 only go through two rounds of selection. During the first round of selection for Question 8, only survey site is significantly associated with the outcome (p-value = 0.0419). It is retained in the model. Because there are no additional significant variables seen in "Round 2", survey site is the only variable in then binary logistic model for Question 8. Similarly for Question 10, only Question 8 is significantly associated with Question 10 (p-value) in the first round. With no additional significant variables seen in the second round, the Question 8 variable is the only variable in the binary logistic model for Question 10.

Table 5D – Multivariable Binary Logistic Model Selection Procedure for Question 10:
10. Every 3-6 months, would you be willing to go to the Brain Health Registry website, play simple online brain games and provide basic health and lifestyle information to help researchers understand the brain as it develops and ages? We will send you information on how if you check 'yes'.
[ ] Yes [ ] No

	Round 1 Containing Intercept and Each Given Variable		**Round 2** Containing Intercept, Survey Site, and Each Given Variable	
	Model AIC	**Wald P-value**	**Model AIC**	**Wald P-value**
Age **18-99**	899.446	0.8986	779.779	0.5249
Gender **Female** **Male**	888.742	0.8956	769.528	0.9278
Question 8 **Yes** **No**	778.182	<.0001*	-----	-----
Race** **AA** **Not AA**	864.141	0.0863	754.129	0.1197
Survey Site **Atlanta** **Chicago** **Houston** **Indianapolis** **Tampa**	904.277	0.8880	782.770	0.5173

* Represents the significant Wald Chi-Square p-values resulting in a variable being selected.
** P-values for Race compare African American to non-African American.

Table 6A displays the resulting best binary logistic model for binary Question 6; **Table 6B** displays the resulting best model for binary Question 7; **Table 6C** displays the resulting best model for Question 8; and **Table 6D** displays the resulting best model for Question 10. In all four tables, the covariates remaining in the final model are listed down the leftmost column. The second column represents the number and percent of those responding "Agree" (for **Table 6A** and **Table 6B**) or "Yes" (for **Table 6C** and **Table 6D**) to the outcome question. The third column represents the estimated beta coefficient and standard error for each covariate. The fourth column represents the odds ratio (OR) with 95% confidence intervals of responding positively. The last column represents the p-value of each covariate's Wald Chi-Square test from the binary logistic regression model.

Table 6A – Final Multivariable Binary Logistic Model for Question 6:
6. To what extend do you agree or disagree with the following statements?
"As a result of attending this event, I have a better understanding of why older African Americans need to get involved in Alzheimer's research."
[ ] Do Not Agree [ ] Agree

	Agree N (%)	**β (SE)**	**OR (95% CI)**	**P-value**
Race**				
African American	785 (89.1)	0.925 (0.412)	2.523 (1.125, 5.657)	0.0247
Not African American (reference)	61 (84.7)			
Survey Site				
Atlanta (reference)	243 (75.2)	---	---	---
Chicago	350 (98.9)	3.376 (0.521)	29.272 (10.55, 81.24)	<.0001
Houston	120 (97.6)	2.710 (0.606)	15.023 (4.579, 49.29)	<.0001
Indianapolis	76 (96.2)	2.384 (0.626)	10.843 (3.181, 36.96)	0.0001
Tampa	77 (81.1)	0.519 (0.314)	1.681 (0.908, 3.112)	0.0986

AIC = 552.395, Likelihood Ratio = 129.0869, p-value < .0001, SC = 581.553
** P-values for Race compare African American to non-African American.

As displayed in **Table 6A**, after the model selection procedure for binary Question 6, the survey site and race are the covariates remaining in the final model – this indicates that these covariates are together significantly associated with binary Question 6. All odds ratio estimates take into account the other covariates in the model. African Americans have 2.523 (95% CI: 1.124-5.657) times higher odds of responding "Agree" compared to those who are not African American. Compared to those surveyed in Atlanta, those surveyed in Chicago have 29.272 (95% CI: 10.55-81.24) times higher odds of responding "Agree"; those surveyed in Houston have 15.023 (4.579-49.29) times higher odds of responding "Agree"; those surveyed in Indianapolis have 10.843 (95% CI: 3.181-36.96) times higher odds of responding "Agree"; and those surveyed in Tampa have 1.681 (95% CI: 0.908-3.112) times higher odds of responding "Agree".

Table 6B – Final Multivariable Binary Logistic Model for Question 7:
7. To what extent do you agree or disagree with the following statements?
"I would like to learn more about clinical trials and other opportunities that may be available to me."
[ ] Do Not Agree [ ] Agree

	Agree N (%)	**β (SE)**	**OR (95% CI)**	**P-value**
Age				
18-99 Years Old	NA	0.013 (0.007)	1.013 (1.001, 1.026)	0.0415
Question 8				
Yes	674 (84.1)	1.954 (0.284)	7.056 (4.041, 12.32)	<.0001
No (reference)	29 (44.6)			
Survey Site				
Atlanta (reference)	228 (74.3)	---	---	---
Chicago	279 (89.1)	1.030 (0.244)	2.800 (1.735, 4.518)	<.0001
Houston	97 (83.6)	0.765 (0.307)	2.149 (1.179, 3.919)	0.0125
Indianapolis	60 (90.9)	1.046 (0.454)	2.847 (1.170, 6.928)	0.0211
Tampa	60 (67.4)	-0.183 (0.279)	0.833 (0.482, 1.440)	0.5130

AIC = 762.852, Likelihood Ratio = 88.8022, p-value < .0001, SC = 796.199
** P-values for Race compare African American to non-African American.

Based on the results found in **Table 6B**, the responses to Question 8, age, and survey site remain in the final binary logistic model predicting the responses to binary Question 7 and are together significantly associated with this outcome. All odds ratio estimates take into account the other covariates in the model. Every one year increase in age corresponds to a 1.01 (1.00-1.03) times increase in odds of responding "Agree" to the binary version of Question 7. Those that respond "Yes" to Question 8 (indicating willingness to learn how to advocate for Alzheimer's research) have 7.06 (95% CI: 4.04-12.32) times higher odds of responding "Agree" to binary Question 7 compared to those who respond "No" to Question 8. Compared to those surveyed in Atlanta, those surveyed in Chicago have 2.80 (95% CI: 1.74-4.52) times higher odds of responding "Agree"; those surveyed in Houston have 2.15 (95% CI: 1.18-3.92) times higher odds of responding "Agree"; those surveyed in Indianapolis have 2.85 (95% CI: 1.17-6.93) times higher odds of responding "Agree"; and those surveyed in Tampa have 0.83 (95% CI: 0.48-1.44) times lower odds of responding "Agree".

Table 6C – Final Multivariable Binary Logistic Model for Question 8:
8. Would you like to receive information about actions you can take to persuade Congress to provide more funding for Alzheimer's?
[ ] Yes [ ] No

	Yes N (%)	**β (SE)**	**OR (95% CI)**	**P-value**
Survey Site				
Atlanta (reference)	316 (92.7)	---	---	---
Chicago	333 (91.2)	-0.195 (0.278)	0.823 (0.477, 1.420)	0.485
Houston	114 (92.7)	0.002 (0.404)	1.002 (0.454, 2.211)	0.996
Indianapolis	76 (97.4)	1.101 (0.746)	3.006 (0.697, 12.97)	0.140
Tampa	81 (84.4)	-0.851 (0.350)	0.427 (0.215, 0.848)	0.015
AIC = 571.875, Likelihood Ratio = 10.7158, p-value = 0.0300, SC = 596.429				

** P-values for Race compare African American to non-African American.

The result of Question 8's multivariable binary logistic model selection procedure is summarized in **Table 5C**. Based on these results, survey site is the only covariate significantly associated with Question 8, and only survey site remains in the final model. Compared to those surveyed in Atlanta, those surveyed in Chicago have 0.82 (95% CI: 0.48-1.42) times lower odds of responding "Yes"; those surveyed in Houston have 1.00 (95% CI: 0.45-2.21) times the odds of responding "Yes"; those surveyed in Indianapolis have 3.01 (95% CI: 0.70-12.97) times higher odds of responding "Yes"; and those surveyed in Tampa have 0.43 (95% CI: 0.22-0.85) times lower odds of responding "Yes". Of these, only the comparison between Atlanta and Tampa is statistically significantly different.

Table 6D – Final Multivariable Binary Logistic Model for Question 10:
10. Every 3-6 months, would you be willing to go to the Brain Health Registry website, play simple online brain games and provide basic health and lifestyle information to help researchers understand the brain as it develops and ages? We will send you information on how if you check 'yes'.
[ ] Yes [ ] No

	Yes N (%)	**β (SE)**	**OR (95% CI)**	**P-value**
Question 8				
Yes	723 (86.3)	2.253 (0.260)	9.513 (5.714, 15.84)	<.0001
No (reference)	30 (38.5)			
AIC = 748.477, Likelihood Ratio = 74.7747, p-value < .0001, SC = 787.822				

Finally, the result of Question 10's model selection procedure is summarized in **Table 6D**. Only the responses to Question 8 remain in the final model, and they therefore are the only variable significantly associated with this outcome. Those that respond "Yes" to Question 8 (indicating a willingness to learn how they can advocate for Alzheimer's research) have 9.51 (95% CI: 5.71-15.84) times higher odds of responding "Yes" to Question 10 compared to those who respond "No" to Question 8.

3.4 Results of Ordinal Logistic Model Selection

The purpose of the ordinal logistic model selection portion of the analysis is to determine the best model predicting the Likert-scale versions of Question 6 and Question 7 while treating them as ordinal variables. For both Question 6 and Question 7, their final ordinal logistic model will determine which potential covariates are significantly associated with the outcome while adjusting for other significant variables. The results of these model selections will later be compared to those of the binary logistic (Section 3.3) and the nominal logistic model selection procedure (Section 3.5).

Table 7A and **Table 7B** depict the selection procedures for Likert-scale Question 6 and Question 7, respectively. These tables are organized similarly to **Tables 5A-5D**. All possible covariates are listed along the leftmost column, and the remaining columns summarize the results of each round of the model selection process. For each round of selection, the Model AIC and the Wald Chi-Square p-value are listed. If a variable is selected in a round, the variable's p-value is marked with an asterisk and will not be listed again for any subsequent round.

Table 7A – Multivariable Ordinal Logistic Model Selection Procedure for Question 6:
6. To what extend do you agree or disagree with the following statements?
"As a result of attending this event, I have a better understanding of why older African Americans need to get involved in Alzheimer's research."
[ ] Strongly Disagree [ ] Disagree [ ] Neither Disagree or Agree [ ] Agree [ ] Strongly Agree

	Round 1 Containing Intercept and Each Given Variable		**Round 2** Containing Intercept, Survey Site, and Each Given Variable		**Round 3** Containing Intercept, Survey Site, Question 8, and Each Given Variable	
	Model AIC	**Wald P-value**	**Model AIC**	**Wald P-value**	**Model AIC**	**Wald P-value**
Age **18-99**	1838.250	0.0177	1730.460	0.7441	1639.811	0.8689
Gender **Female** **Male**	1810.722	0.2019	1702.031	0.7913	1614.105	0.7693
Question 8 **Yes** **No**	1752.920	<.0001	1637.838	<.0001*	-----	-----
Race** **AA** **Not AA**	1802.527	0.2007	1693.079	0.0603	1601.176	0.0656
Survey Site **Atlanta** **Chicago** **Houston** **Indianapolis** **Tampa**	1728.566	<.0001*	-----	-----	-----	-----

* Represents the significant Wald Chi-Square p-values resulting in a variable being selected.
** P-values for Race compare African American to non-African American.

As seen in **Table 7A**, the final multivariable ordinal logistic model for Question 6 contains the survey site and Question 8 variables. In the first round of selection, when the survey site is the only variable in the model, survey site is significantly associated with ordinal Question 6 (p-value < .0001). Both Question 8 (p-value < .0001) and age (p-value = 0.0177) are also significant in the first round, but survey site is the most significant. The survey site variable is therefore retained in all "Round 2" models. During the second round, Question 8 (p-value < .0001) is the only variable significantly associated with ordinal Question 6, and it is retained in all "Round 3" models along with the survey site variable. The

third round produces no additional significant variables, so the final model contains only the survey site and Question 8 variables determined to be significant in the first two rounds.

Table 7B – Multivariable Ordinal Logistic Model Selection Procedure for Question 7:
7. To what extent do you agree or disagree with the following statements?
"I would like to learn more about clinical trials and other opportunities that may be available to me."
[ ] Strongly Disagree [ ] Disagree [ ] Neither Disagree or Agree [ ] Agree [ ] Strongly Agree

	Round 1 Containing Intercept and Each Given Variable		**Round 2** Containing Intercept, Survey Site, and Each Given Variable		**Round 3** Containing Intercept, Survey Site, Question 8, and Each Given Variable	
	Model AIC	**Wald P-value**	**Model AIC**	**Wald P-value**	**Model AIC**	**Wald P-value**
Age						
18-99	2156.877	0.0031	2123.344	0.2528	2011.653	0.2139
Gender						
Female	2130.625	0.6958	2091.682	0.5245	1981.496	0.2782
Male						
Question 8						
Yes	2051.415	<.0001	2011.226	<.0001*	-----	-----
No						
Race**						
AA	2121.845	0.0802	2082.478	0.0487	1976.940	0.1676
Not AA						
Survey Site						
Atlanta						
Chicago						
Houston	2122.650	<.0001*	-----	-----	-----	-----
Indianapolis						
Tampa						

* Represents the significant Wald Chi-Square p-values resulting in a variable being selected.
** P-values for Race compare African American to non-African American.

Based on the results of the model selection procedure summarized in **Table 7B**, the survey site and Question 8 variables are together significantly associated with ordinal Question 7. In the first round, age (p-value = 0.0031), Question 8 (p-value < .0001), race (p-value = 0.0802), and survey site (p-value < .0001) are all significantly associated with ordinal Question 7. The survey site variable is the most significant, and it is therefore retained in all second round models. In the second round, Question 8 is the most significant variable (p-value < .0001), and it - along with the survey site variable – is retained in the

round three models. No additional significant covariates are found in "Round 3". The final ordinal logistic model for Question 7 therefore contains only survey site and Question 8.

Table 8A displays the resulting best ordinal logistic model for Likert-scale Question 6, and **Table 8B** displays the resulting best model for Likert-scale Question 7. In both tables, the covariates remaining in the final model are listed down the leftmost column. The second column represents the estimated beta coefficient and standard error for each covariate. The third column represents the odds ratio (OR) with 95% confidence interval of responding one category higher to the question (as an example, responding "Disagree" instead of "Strongly Disagree" or responding "Agree" instead of "Neither Disagree or Agree"). The last column represents the p-value of each covariate's Wald Chi-Square test from the ordinal logistic regression model.

Table 8A – Final Ordinal Logistic Model for Question 6:
6. To what extend do you agree or disagree with the following statements?
"As a result of attending this event, I have a better understanding of why older African Americans need to get involved in Alzheimer's research."
[ ] Strongly Disagree [ ] Disagree [ ] Neither Disagree or Agree [ ] Agree [ ] Strongly Agree

	β (SE)	OR (95% CI)	P-value
Question 8			
Yes	0.934 (0.231)	2.545 (1.619, 4.001)	<.0001
No (reference)			
Survey Site			
Atlanta (reference)	---	---	---
Chicago	1.626 (0.172)	5.084 (3.628, 7.125)	<.0001
Houston	1.786 (0.262)	5.962 (3.565, 9.971)	<.0001
Indianapolis	1.055 (0.277)	2.873 (1.670, 4.940)	0.0001
Tampa	0.335 (0.227)	1.399 (0.897, 2.180)	0.1387

AIC = 1637.838, Likelihood Ratio = 139.7948, LR P-value < .0001, SC = 1624.741, Score Test P-value <.0001***
****** P-values for Race compare African American to non-African American.**
******* Significant p-value for the Score Test indicates that the proportional odds assumption is not met.**

Based on the results summarized in **Table 8A**, the final ordinal logistic model for Likert-scale Question 6 contains the responses to Question 8 and the survey site – meaning these variables are together significantly associated with Question 6. Those who respond "Yes" to Question 8 (indicating a willingness to learn how to advocate for Alzheimer's) have 2.55 (95% CI: 1.62-4.00) times higher odds of

responding one category higher in ordinal Question 6 compared to those who responded "No" to Question 8. Compared to those surveyed in Atlanta, those surveyed in Chicago have 5.08 (95% CI: 3.63-7.13) times higher odds of responding one category higher in ordinal Question 6; those surveyed in Houston have 5.96 (95% CI: 3.57-9.97) times higher odds of responding one category higher; those surveyed in Indianapolis have 2.87 (95% CI: 1.67-4.94) times higher odds of responding one category higher; and those surveyed in Tampa have 1.40 (95% CI: 0.90-2.18) times higher odds of responding one category higher. Indicated by a significant Score test p-value, this model did not meet the proportional hazards assumption.

Table 8B – Final Ordinal Logistic Model for Question 7:
7. To what extent do you agree or disagree with the following statements?
"I would like to learn more about clinical trials and other opportunities that may be available to me."
[ ] Strongly Disagree [ ] Disagree [ ] Neither Disagree or Agree [ ] Agree [ ] Strongly Agree

	β (SE)	**OR (95% CI)**	**P-value**
Question 8			
Yes	1.596 (0.241)	4.932 (3.078, 7.904)	<.0001
No (reference)			
Survey Site			
Atlanta (reference)	---	---	---
Chicago	0.931 (0.157)	2.538 (1.864, 3.454)	<.0001
Houston	0.697 (0.211)	2.008 (1.329, 3.035)	0.0009
Indianapolis	0.753 (0.272)	2.123 (1.246, 3.618)	0.0056
Tampa	-0.146 (0.224)	0.864 (0.557, 1.340)	0.5133

AIC = 2011.226, Likelihood Ratio = 101.1306, LR P-value < .0001, SC = 1990.779, Score Test P-value <.0001**
**** Significant p-value for the Score Test indicates that the proportional odds assumption is not met.**

As seen in **Table 8B**, the final ordinal logistic model for Likert-scale Question 7 contains the responses to Question 8 and the survey site. The Question 8 and survey site variables are therefore together significantly associated with ordinal Question 7. Those who responded "Yes" to Question 8 have 4.93 (95% CI: 3.08-7.90) times higher odds of responding one category higher in ordinal Question 7 compared to those who responded "No" to Question 8. Compared to those surveyed in Atlanta, those surveyed in Chicago have 2.54 (95% CI: 1.86-3.45) times higher odds of responding one category higher in ordinal Question 7; those surveyed in Houston have 2.01 (95% CI: 1.33-3.04) times higher odds of

responding one category higher; those surveyed in Indianapolis have 2.12 (95% CI: 1.25-3.62) times higher odds of responding one category higher; and those surveyed in Tampa have 0.86 (95% CI: 0.56-1.34) times lower odds of responding one category higher. Just as in the ordinal logistic model for Question 6, this model also did not meet the proportional hazards assumption based on a significant Score test p-value.

3.5 Results of Nominal Logistic Model Selection

The purpose of these nominal logistic model selection procedures, the final portion of the analysis, is to determine the best model predicting the Likert scale versions of Questions 6 and 7 while treating them as nominal variables. For both Question 6 and Question 7, their final nominal logistic model will determine which potential covariates are significantly associated with the outcome while adjusting for other significant variables. The results of these model selections will later be compared to those of the binary logistic (Section 3.3) and the ordinal logistic model selection procedure (Section 3.4).

Table 9A and **Table 9B** depict the selection procedures for nominal Question 6 and Question 7, respectively. These tables are organized similarly to **Tables 5A-5D** and **Tables 7A-7B**. All possible covariates are listed along the leftmost column, and the remaining columns summarize the results of each round of the model selection process. For each round of selection, the Model AIC and the Wald Chi-Square p-value are listed. If a variable is selected in a round, the variable's p-value is marked with an asterisk and will not be listed again for any subsequent round.

Based on the model selection procedure summarized in **Table 9A**, the survey site, Question 8, and race variables are together significantly associated with nominal Question 6. All three variables are therefore in nominal Question 6's final nominal logistic model. In the first round of selection, Question 8 (p-value < .0001), race (p-value = 0.0181), and survey site (p-value <.0001) are significantly associated with Question 6. The survey site variable is the most significant variable, and so it is retained in all other

rounds of selection. In the second round, Question 8 (p-value < .0001) and race (p-value = 0.0089) are significant. The Question 8 variable is the most significant, so it is retained for "Round 3". In the third round, only the race variable (p-value = 0.0279) is significantly associated with Question 6, and it is also retained. No variables are found to be significant in the fourth round. The final model then contains only the survey site, Question 8, and race variables.

Table 9A – Multivariable Nominal Logistic Model Selection Procedure for Question 6:
6. To what extend do you agree or disagree with the following statements?
"As a result of attending this event, I have a better understanding of why older African Americans need to get involved in Alzheimer's research."
[ ] Strongly Disagree [ ] Disagree [ ] Neither Disagree or Agree [ ] Agree [ ] Strongly Agree

	Round 1 Containing Intercept and Each Given Variable		**Round 2** Containing Intercept, Survey Site, and Each Given Variable		**Round 3** Containing Intercept, Survey Site, Question 8, and Each Given Variable		**Round 4** Containing Intercept, Survey Site, Question 8, Race, and Each Given Variable	
	Model AIC	**Wald P-value**	**Model AIC**	**Wald P-value**	**Model AIC**	**Wald P-value**	**Model AIC**	**Wald P-value**
Age **18-99**	1841.055	0.0631	1689.170	0.6120	1601.552	0.5356	1565.009	0.3770
Gender **Female** **Male**	1815.247	0.6182	1663.228	0.7515	1579.475	0.8055	1547.633	0.6676
Question 8 **Yes** **No**	1750.027	<.0001	1596.962	<.0001*	-----	-----	-----	-----
Race** **AA** **Not AA**	1800.339	0.0181	1561.582	0.0089	1561.582	0.0279*	-----	-----
Survey Site **Atlanta** **Chicago** **Houston** **Indianap.** **Tampa**	1684.071	<.0001*	-----	-----	-----	-----	-----	-----

* Represents the significant Wald Chi-Square p-values resulting in a variable being selected.
** P-values for Race compare African American to non-African American.

As seen in **Table 9B**, Question 8, survey site, and age are all represented in the final nominal logistic model for nominal Question 7. During the first round of selection, age (p-value = 0.0081), Question 8 (p-value < .0001), and survey site (p-value < .0001) are significantly associated with Question

7. As the Question 8 variable is the most significant, it is retained for the second round. In "Round 2", age (p-value = 0.0039) and survey site (p-value < .0001) are significant. The survey site variable is retained along with the Question 8 variables for the third round. In "Round 3", only the age variable is significant (p-value = 0.0447), so it is also retained. Without any additional significant covariates in the fourth and final round, only the Question 8, survey site, and age variables remain the final nominal logistic model for Question 7.

Table 9B – Multivariable Nominal Logistic Model Selection Procedure for Question 7:
7. To what extent do you agree or disagree with the following statements?
"I would like to learn more about clinical trials and other opportunities that may be available to me."
[ ] Strongly Disagree [ ] Disagree [ ] Neither Disagree or Agree [ ] Agree [ ] Strongly Agree

	Round 1 Containing Intercept and Each Given Variable		**Round 2** Containing Intercept, Question 8, and Each Given Variable		**Round 3** Containing Intercept, Question 8, Survey Site, and Each Given Variable		**Round 4** Containing Intercept, Question 8, Survey Site, Age, and Each Given Variable	
	Model AIC	**Wald P-value**	**Model AIC**	**Wald P-value**	**Model AIC**	**Wald P-value**	**Model AIC**	**Wald P-value**
Age **18-99**	2158.105	0.0081	2017.287	0.0039	1966.777	0.0447*	-----	-----
Gender **Female** **Male**	2136.396	0.9829	2000.740	0.9991	1947.452	0.8840	1946.270	0.8858
Question 8 **Yes** **No**	2024.604	<.0001*	------	-----	-----	-----	-----	-----
Race** **AA** **Not AA**	2124.610	0.1341	1991.421	0.3826	1939.736	0.5084	1939.720	0.5548
Survey Site **Atlanta** **Chicago** **Houston** **Indianap.** **Tampa**	2103.728	<.0001	1968.450	<.0001*	-----	-----	-----	-----

* Represents the significant Wald Chi-Square p-values resulting in a variable being selected.
** P-values for Race compare African American to non-African American.

Table 10A displays the resulting best nominal logistic model for the Likert-scale version of Question 6, and **Table 10B** displays the resulting best model for the Likert-scale version of Question 7. In

both tables, the covariates remaining in the final model are listed down the leftmost column. If a covariate is categorical with more than two categories (program location is the only such covariate), each non-reference category fills its own row. Because the nominal logit model treats the outcome variables as nominal, the effect of each covariate is estimated for every category of the outcome variable compared to a reference group. For that reason, the second column represents the different categories of the Likert scale variables, and estimates are made using "Strongly Disagree" as a reference. The third column displays the estimated beta coefficient and standard error for each covariate by outcome category. The forth column represents the odds ratio (OR) with 95% confidence interval of responding within a response category of the outcome. The last column represents the p-value of each estimate's Wald Chi-Square test for the nominal logistic regression model.

Based on the results in **Table 10A**, the final nominal logistic model for Likert scale Question 6 contains the responses to Question 8, race, and survey site – indicating that these three covariates are together statistically significant predictors of Question 6. Although we have used selection procedures to find this final model, our model has quasi-complete separation of data points. Quasi-complete separation happens in a logistic regression when an outcome variable separates a predictor variable or a combination of predictor variables to a certain degree. This quasi-complete separation may have occurred because, while using several categorical variables whose categories are coded as indicators, there may be subgroups in these variables that unanimously responded to a questionnaire item. This failure for all parameters to converge results in non-existent maximum likelihood estimates under special conditions in the sample. This failure suggest that the estimates of this model may not be appropriate to report as all infinite odds ratio estimates are likely a result of quasi-separation. All such infinite odds ratio estimates are highlighted in the table (Heinze, 2002).

Table 10A – Final Nominal Logistic Model for Question 6:

6. To what extend do you agree or disagree with the following statements?

"As a result of attending this event, I have a better understanding of why older African Americans need to get involved in Alzheimer's research."

[ ] Strongly Disagree [ ] Disagree [ ] Neither Disagree or Agree [ ] Agree [ ] Strongly Agree

		β (SE)	**OR (95% CI)**	**P-value**
Question 8	Strongly Agree	0.838 (0.464)	2.313 (0.931, 5.742)	0.0708
Yes	Agree	-0.333 (0.447)	0.717 (0.298, 1.722)	0.4564
No (reference)	Neither	-1.489 (0.646)	0.226 (0.064, 0.799)	0.0211
	Disagree	9.700 (309.4)	>999. (<.001, >999)	0.9750
	Strongly Disagree (reference)	---	---	---
Race**	Strongly Agree	0.389 (0.574)	1.476 (0.479, 4.545)	0.4978
African American	Agree	0.043 (0.584)	1.044 (0.333, 3.276)	0.9410
Not AA	Neither	-1.676 (0.799)	0.187 (0.039, 0.897)	0.0360
(reference)	Disagree	8.143 (302.3)	>999. (<.001, >999)	0.9785
	Strongly Disagree (reference)	---	---	---
Survey Site	Strongly Agree	11.51 (28.84)	>999. (<.001, >999)	0.6899
Atlanta (reference)	Agree	10.65 (28.84)	>999. (<.001, >999)	0.7120
Chicago	Neither	9.62 (28.85)	>999. (<.001, >999)	0.7389
	Disagree	-0.05 (181.3)	0.949 (<.001, >999)	0.9998
	Strongly Disagree (reference)	---	---	---
Survey Site	Strongly Agree	3.813 (1.019)	45.28 (6.145, 333.7)	0.0002
Atlanta (reference)	Agree	2.664 (1.042)	14.35 (1.864, 110.5)	0.0105
Houston	Neither	2.128 (1.287)	8.401 (0.675, 104.6)	0.0981
	Disagree	-7.713 (298.2)	<0.001 (<.001, >999)	0.9794
	Strongly Disagree (reference)	---	---	---
Survey Site	Strongly Agree	11.399 (62.44)	>999. (<.001, >999)	0.8552
Atlanta (reference)	Agree	11.015 (62.44)	>999. (<.001, >999)	0.8600
Indianapolis	Neither	10.261 (62.45)	>999. (<.001, >999)	0.8695
	Disagree	0.130 (387.5)	1.139 (<.001, >999)	0.9997
	Strongly Disagree (reference)	---	---	---
Survey Site	Strongly Agree	0.444 (0.354)	1.559 (0.779, 3.120)	0.2096
Atlanta (reference)	Agree	0.285 (0.385)	1.329 (0.626, 2.825)	0.4591
Tampa	Neither	-1.945 (1.193)	0.143 (0.014, 1.483)	0.1031
	Disagree	-10.529 (333.3)	<0.001 (<.001, >999)	0.9748
	Strongly Disagree (reference)	---	---	---

AIC = 1561.582, Likelihood Ratio = 214.0899, LR P-value < .0001, SC = 1696.542

Model Convergence Status: Quasi-Complete Separation of Data Points Detected

** P-values for Race compare African American to non-African American.

Table 10B – Final Nominal Logistic Model for Question 7:

7. To what extent do you agree or disagree with the following statements?

"I would like to learn more about clinical trials and other opportunities that may be available to me."

[ ] Strongly Disagree [ ] Disagree [ ] Neither Disagree or Agree [ ] Agree [ ] Strongly Agree

		β (SE)	OR (95% CI)	P-value
Age	Strongly Agree	-0.001 (0.011)	0.999 (0.978, 1.020)	0.8905
18-99 Years Old	Agree	-0.001 (0.011)	0.999 (0.978, 1.020)	0.9040
	Neither	-0.021 (0.013)	0.979 (0.956, 1.003)	0.0921
	Disagree	-0.031 (0.020)	0.969 (0.933, 1.007)	0.1117
	Strongly Disagree (reference)	---	---	---
Question 8	Strongly Agree	1.538 (0.764)	6.02 (1.505, 24.05)	0.0441
Yes	Agree	-0.393 (0.646)	0.83 (0.270, 2.575)	0.5425
No (reference)	Neither	-2.086 (0.653)	0.15 (0.049, 0.483)	0.0014
	Disagree	-2.199 (0.823)	0.14 (0.031, 0.637)	0.0075
	Strongly Disagree (reference)	---	---	---
Survey Site	Strongly Agree	2.892 (0.544)	18.03 (6.211, 52.37)	<.0001
Atlanta (reference)	Agree	2.330 (0.548)	10.27 (3.509, 30.07)	<.0001
Chicago	Neither	2.579 (0.610)	13.19 (3.988, 43.62)	<.0001
	Disagree	1.992 (0.914)	7.33 (1.222, 44.00)	0.0293
	Strongly Disagree (reference)	---	---	---
Survey Site	Strongly Agree	2.470 (0.741)	11.82 (2.768, 50.49)	0.0009
Atlanta (reference)	Agree	2.067 (0.747)	7.90 (1.829, 34.13)	0.0056
Houston	Neither	2.478 (0.803)	11.92 (2.473, 57.45)	0.0020
	Disagree	1.340 (1.307)	3.82 (0.295, 49.43)	0.3051
	Strongly Disagree (reference)	---	---	---
Survey Site	Strongly Agree	15.141 (540.7)	>999.0 (<.001, >999)	0.9777
Atlanta (reference)	Agree	14.814 (540.7)	>999.0 (<.001, >999)	0.9781
Indianapolis	Neither	15.173 (540.7)	>999.0 (<.001, >999)	0.9776
	Disagree	0.994 (1167.)	2.70 (<.001, >999)	0.9993
	Strongly Disagree (reference)	---	---	---
Survey Site	Strongly Agree	0.010 (0.401)	1.01 (0.460, 2.217)	0.9796
Atlanta (reference)	Agree	0.150 (0.394)	1.16 (0.536, 2.517)	0.7034
Tampa	Neither	0.602 (0.490)	1.83 (0.699, 4.772)	0.2192
	Disagree	0.786 (0.745)	2.20 (0.510, 9.451)	0.2910
	Strongly Disagree (reference)	---	---	---

AIC = 1909.172, Likelihood Ratio = 171.0076, LR P-value < .0001, SC = 2041.607

** P-values for Race compare African American to non-African American.

Finally, as seen in **Table 10B**, the final nominal logistic model for Likert-scale Question 7 contains age, the responses to Question 8, and survey site – indicating that these three covariates are together statistically significant predictors of Question 7. Unlike the nominal logit model fitted for Question 6, the model in **Table 10B** does not show any signs of complete or quasi-complete separation of data points. Our model estimates may therefore be appropriate.

As an example of how to interpret the results of **Table 10B**, consider the row representing estimates for the Question 8 variable. Those that respond "Yes" to Question 8 have 6.02 (95% CI: 1.51-24.05) times higher odds of responding "Strongly Agree" over "Strongly Disagree" compared to those who respond "No" to Question 8. Those that respond "Yes" to Question 8 also have 0.83 (95% CI: 0.27-2.58) times lower odds of responding "Agree" over "Strongly Disagree" compared to those who respond "No". They also have 0.15 (95% CI: 0.05-0.48) times lower odds of responding "Neither" over "Strongly Disagree", and 0.14 (95% CI: 0.03-0.64) times lower odds of responding "Disagree" over "Strongly Disagree". All rows can be interpreted in this manner. Also, the occasional infinite odds ratio estimate in this table suggest some level of incomplete separation, although not enough to warrant a convergence error.

3.6 Comparison of Three Multivariable Logistic Models

As stated in the Introduction, the third aim of this study is to determine the best method of analyzing outcome variables that are in a Likert scale. In this last aim, we question whether analyzing these variables as ordinal, nominal, or binary is most appropriate.

Table 11 summarizes independent variables that are significantly associated with the dependent, outcome variables in the multivariable binary, ordinal, and nominal logistic regression models for both Question 6 and Question 7. In this table, all possible covariates are listed in the leftmost column. The second column indicates if that variable is significant in the final multivariable binary logistic model for

either Question 6 or 7 (Seen in section 3.3). The third column indicates whether the variable is significant in the final multivariable ordinal logistic model for either question (Seen in Section 3.4). The fourth column indicates whether the variable is significant in the final multivariable nominal logistic model (Seen in Section 3.5). In columns 2-4, if a variable is significant in that model, it is marked with an "X". The last column represents how many of the three final logistic models contain each possible covariate for both questions.

Table 11 – Comparison of Three Multivariable Logistic Regression Models

	Question 6			
	Significant in Binary Logistic Model?	**Significant in Ordinal Logistic Model?**	**Significant in Nominal Logistic Model?**	**Number of Models Significant**
Age				0
Gender				0
Question 8		X	X	2
Race	X		X	2
Survey Site	X	X	X	3
	Question 7			
	Significant in Binary Logistic Model?	**Significant in Ordinal Logistic Model?**	**Significant in Nominal Logistic Model?**	**Number of Models Significant**
Age	X		X	2
Gender				0
Question 8	X	X	X	3
Race				0
Survey Site	X	X	X	3

Based on the results seen in **Table 11**, all three models (binary, ordinal, and nominal) contain the survey site variable for both Question 6 and Question 7. For Question 6, both the ordinal and the nominal

logistic models contain the Question 8 variable, and both the binary and nominal logistic models contain the race variable. For Question 7, all three models contain the Question 8 variable. Also for Question 7, both the binary and nominal models contain the age variable.

To further compare the three logistic models fitted for Question 6, the goodness of fit statistics for each model should be inspected. First, the multivariable binary logistic model has the following goodness of fit (GOF) statistic: AIC = 522.4, LR test statistic = 129.1, LR p-value < 0.0001, and SC = 581.6. The ordinal logistic model has the following GOF statistics: AIC = 1637.8, LR test statistic = 139.8, LR p-value < 0.0001, and SC = 1624.7. It should again be noted that this ordinal logistic model does not conform to the proportional hazards assumption, making it less credible. The nominal logistic model has the following GOF statistics: AIC = 1561.6, LR test statistic = 214.1, LR p-value < 0.0001, and SC = 1696. 5. It should also be noted that Question 6's nominal model shows quasi-complete separation of data points. For the ease of interpretation, and because both the ordinal and nominal models either do not meet their model's assumptions or do not fully converge, the binary logistic model is more appropriate in this circumstance to predict the responses to Question 6.

The three logistic models fitted for Question 7 can similarly be compared. First, the multivariable binary logistic model has the following goodness of fit (GOF) statistics: AIC = 762.9, LR test statistic = 88.8, LR p-value < 0.0001, and SC = 796.2. The ordinal logistic model has the following GOF statistics: AIC = 2011.2, LR test statistic = 101.1, LR p-value < 0.0001, and SC = 1990.8. This ordinal logistic model also does not conform to the proportional hazards assumption, making it less credible. The nominal logistic model has the following GOF statistics: AIC = 1909.2, LR test statistic = 171.0, LR p-value < 0.0001, and SC = 2041.6. Unlike that for Question 6, Question 7's nominal logistic model does not experience quasi-complete separation of data points. Because the binary logistic model has a lower AIC and a lower SC, the binary logistic model is more appropriate in this circumstance to predict the responses to Question 7.

4 Discussion

4.1 Discussion of Results

From the results summarized in **Table 3**, the vast majority of program attendees are female (87.8%) and African American (92.5%). These proportions are fairly consistent across survey sites. As the Forget Me Not Project seeks to address the disproportionate affect Alzheimer's has on African Americans, members of this minority are actively recruited to attend the program. Although recruitment strategies did not specifically target women, Alzheimer's disproportionately affects women. "60% of caregivers are women, and at the age of 65, women have 1 in 6 chance of developing Alzheimer's, compared to a 1 in 11 chance in men" – a fact that may influence the amount of women in attendance (Why, 2015).

Also seen in **Table 3**, the responses to questionnaire items often vary by survey site. For example, when answering Question 1, which asks participants how they heard about the event, most respondents claim they heard about the event from church or from a friend. Participants from Atlanta, Chicago, and Indianapolis echo these results, but responses from Houston and Tampa differ. For both Houston and Tampa, over 44% claim they heard about the event from a friend, over 29% claim that they heard about it from "other" sources, and less than 8% claim they heard about it from church. Perhaps these differences – and other differences between survey sites - can be attributed to the differences in the African American communities of these areas. For instance, programs in Chicago and Atlanta are supported by universities and strong engagement in their communities, while the Tampa program seems to lack a hub for this community. The event from Atlanta was held in conjunction with a large faith-based conference which could explain increased referral from church. But as participants were advised to select only one referral option, these differences may also be attributed to uncertainty of how to respond. As an example,

someone who was referred by a friend from church could have selected either "Friend" or "Church". This result needs to be further explored in future research.

While responses to several items in the questionnaire vary by survey site, often responses are unanimous across these locations. Almost all respondents, by answering "Yes" to Questions 2-5, claimed they enjoyed the play, enjoyed the panel discussion, thought the length of the program was appropriate, and thought the information was helpful to them. The majority also agreed, that they have a better understanding of why older African Americans need to get involved in Alzheimer's research after attending the event and that they would like to learn more about clinical trials. These consistent results imply a great success for the Forget Me Not Project. Based on their responses, participants appear to be enjoying the events, feel that they are learning from the experience, and would like to learn more about opportunities to get involved. It is possible, however, that participants with negative responses tend to forgo answering these questions. It is also possible that those most willing to be engaged are the only ones participating in the event and filling out the survey.

For both the univariate analyses and the model selection procedures, we were interested in determining whether responses to questions of interest were associated with other items in the questionnaire, specifically with age, race, gender, survey site, and having expressed interest to learn how to advocate for Alzheimer's research. Questions of interests include those regarding willingness to engage in Alzheimer's research, willingness to advocate for this research, and whether participants felt they learned anything from the program.

When determining which items in the questionnaire are associated with Question 6, the univariate analysis and the three model selection procedures generated different results. Question 6 is used to measure the extent to which a participant agrees that they gained a better understanding of why older African Americans need to be involved in Alzheimer's research. Based on the results of the univariate, seen in **Table 4A**, the responses to Question 8, age, and the program location were all associated with the responses to Question 6. The multivariable binary logistic model, seen in **Table 6A**, suggests that the

responses to Question 6 are significantly associated with the responses to the program location and race – but not with age or with the responses to Question 8. These results differ with those of the multivariable ordinal and nominal logistic model selection procedures, seen in **Table 8A** and **Table 10A**. **Table 8A** suggests that Question 6 is associated with Question 8 and survey site, while **Table 10A** suggests it is associated with Question 8, survey site, and race. It should be noted that the ordinal model does not meet the proportional hazards assumption, and the nominal model evokes quasi-complete separation of data points. The multivariable binary logistic model is then the most appropriate for this dataset. Therefore, we propose that Question 6 is associated with only race and survey site in this study population. Specifically, African Americans have 2.52 times higher odds of agreeing that they gained a better understanding of why older African Americans need to be involved in Alzheimer's research than those who are not African American. Compared to those surveyed in Atlanta, those surveyed in Chicago have 29.27 times higher odds of agreeing; those surveyed in Houston have 15.02 times higher odds of agreeing; and those surveyed in Indianapolis have 10.84 times higher odds of agreeing.

The results of the four analyses also differ when determining which items in the questionnaire are associated with Question 7. Question 7 is used to measure the extent to which participants agree that they would like to learn more about clinical trials. In the univariate analysis summarized in **Table 4B** the responses to Question 8, age, race, and program location are found to be associated with the responses to binary Question 7. Both the binary logistic and the nominal logistic models, seen in **Table 6B** and **Table 10B**, suggest that age, Question 8 and survey site are significantly associated with Question 7. These results differ from the ordinal logistic regression model selection, seen in **Table 8B**, suggesting that only the responses to Question 8 and the program location were associated with Question 7. It should be noted however that this ordinal logistic model, again, does not conform to the proportional hazards assumption, making it the least credible. For the ease of interpretation, and because of its goodness of fit statistics, the binary logistic model is more appropriate in this circumstance. We therefore propose that three covariates are associated with Question 7 in this study population – responses to Question 8, age, and the survey site.

Specifically, every one year increase in age corresponds to a 1.01 times increase in odds of a participant agreeing that they would like to learn more about clinical trials. Also, those that respond "Yes" to Question 8 – indicating a willingness to learn how to advocate for Alzheimer's research – have 7.06 times higher odds of agreeing that they would like to learn more about clinical trials compared to those who respond "No" to Question 8. Lastly, compared to those surveyed in Atlanta, those surveyed in Chicago have 2.80 times higher odds of agreeing; those surveyed in Houston have 2.15 times higher odds of agreeing; those surveyed in Indianapolis have 2.85 times higher odds of agreeing; and those surveyed in Tampa have 0.83 times lower odds of agreeing.

Unlike Questions 6 and 7, the other two outcomes of interest, Questions 8 and 10, are only studied through univariate analyses and through multivariable binary logistic model selection. Question 8 asks if a program attendee is willing to learn how to advocate for Alzheimer's research, and Question 10 asks if the attendee is willing to learn about the Brain Health Registry. For the analysis using Question 8 as the outcome variable, both the univariate analysis (seen in **Table 4C**) and the binary logistic model (seen in **Table 6C**) find that the responses to Question 8 are associated with participants' survey location. We conclude that compared to those surveyed in Atlanta, those surveyed in Chicago have 0.82 times lower odds of responding that yes, they would like to receive information about actions they can take to persuade Congress to provide more funding for Alzheimer's; those surveyed in Houston have 1.00 times the odds of responding "Yes"; those surveyed in Indianapolis have 2.01 times higher odds of responding "Yes"; and those surveyed in Tampa have 0.43 times lower odds of responding "Yes". The univariate analysis summarized in **Table 4D** suggests that the responses to Question 10 are associated with both the responses to Question 8 and the participant's race, while the binary logistic model in **Table 6D** suggests that Question 10 is only associated with the responses to Question 8. In this case, because the multivariable model adjusts for other significant covariates, we propose that Question 10 is only associated with the responses to Question 8 in the study population. Those that respond "Yes" to Question 8 – indicating a willingness to learn how to advocate for Alzheimer's research – have 9.51 times higher

odds of being willing to participate with the Brain Health Registry website compared to those who respond “No” to Question 8.

4.2 Suggestions to Improve Questionnaire

After completing this study, there are ways to improve the Forget Me Not questionnaire. First, when surveying the birth date of each participant, the questionnaire should ask for the full four digit year. The current questionnaire allows a participant to report a two digit year. When only reporting a two digit year, complications can arise. Without reporting the specific year, for example, a researcher who sees a birth date of 05/05/05 does not know whether a participant is born in the year 1905 or 2005. While in some situations a researcher may be able to deduce the specific year, in this study, this was not the case as there could be adolescents attending the play as well as older senior citizens. For this reason, we only considered attendees 18-99 years old.

Secondly, it is also unclear whether participants attended both the “Forget Me Not” play and the expert panel discussion. Many participants, while filling out a questionnaire, did not answer every question. Several of these individuals would write on the questionnaire that unanswered items were not applicable to them. Because the program allowed individuals to attend all or part of the event, it may be beneficial to add a question asking if the participant attended the entire event. Using this information, we can decide whether an individual who only attended part of the program may be allowed in the study. As the panel was presented during intermission, and the surveys were administered at the conclusion of the second act of they play, it is possible that those who completed the survey did attend the entire program.

Lastly, it may be useful to include additional questions in the survey regarding a participant’s background, why they are at the event, and their views on Alzheimer’s research. With only a dozen questions, it is difficult to get a true understanding of an individual’s experience at the Forget Me Not Program or the motivation behind each question’s response. As an example, it could be informative to

know if an individual has any family history of Alzheimer's or other dementias in order to understand their motivation for wanting or not wanting to be involved in clinical trials. It also could be advantageous to include clarifying questions after certain items in the questionnaire. For instance, Question 4 asks if participant thought that the event format and length were appropriate. An additional question could be added asking if they answered "No" to Question 4, did they think the event was too long or too short.

4.3 Future Analysis and Limitations

One limitation in this study is the voluntary nature of the program. In the Forget Me Not events, anyone who is interested can attend. Recruiting individuals to attend the event is therefore strictly voluntary and we assume that everyone in the audience chooses to be there with few exception. By electing to attend, participants have already displayed interest in Alzheimer's research and may have other motivators that make them more engaged in the program. This selection bias may skew the results to appear more positive than they would in the general population. Also, those that were coerced into attending and do not truly want to be at the event may respond more negatively – this would also skew the results. For this type of campaign, of course, it is not possible to force unwilling participants to attend these events.

Another limitation in this study is that participants responded almost unanimously to many of the items in the questionnaire. This lack of response diversity is most clear in Questions 2-5 where over 96% of participants responded "Yes" to each of these questions. When participants respond unanimously to a question, that question does not provide discernable information. While many of the questions in this study did not have unanimous responses, participants did highly favor positive responses. This trend led to convergence issues in our study when conducting the multivariable nominal logistic model selection procedures. In addition to causing computational complications, this unanimous positive feedback

prevents the program from benefitting from constructive criticism. It therefore may be beneficial to incorporate more detailed questions that could lend some diversity to the responses.

In future studies, we suggest addressing these limitations and making improvements the questionnaire. We also advise expanding the program to a wider audience. By expanding the program, and by incorporating individuals from different backgrounds, the program can assess how individuals in other communities respond to the program and can reach other minorities that are disproportionately represented in Alzheimer's research.

4.4 Conclusion

In conclusion, the study population is predominately African American and female. Across survey sites, the participants responded positively to the program and reported that they would be willing to learn more about opportunities to get involved in clinical trials and to advocate for Alzheimer's research. A participant's belief that they now have a better understanding of why older African Americans need to get involved in Alzheimer's research is significantly associated with the participant's race taken together with their survey site. A participant's age, where they were surveyed, and their willingness to learn how they can advocate for Alzheimer's research are together significantly associated with their willingness to learn about clinical trials. In addition, both the survey site and an individual's willingness to participate in the Brain Health Registry are each singly associated with the individual's willingness to learn how they can advocate for Alzheimer's research. When analyzing Likert-scale variables in this study, we found that - between binary, ordinal, or nominal logistic models – the binary logistic model lends itself best to our data. In future studies, we suggest implementing changes to the questionnaire, addressing the limitations of this study, and expanding the program to a wider net of communities.

References

African-Americans and Alzheimer's Disease: The Silent Epidemic [Brochure]. Chicago, IL: Alzheimer's Association.

African Americans Network Against Alzheimer's (AANAA). (2012). Retrieved March 14, 2016, from http://www.usagainstalzheimers.org/networks/african-americans

Alzheimer's Association. 2015 Alzheimer's disease facts and figures. (2015). *Alzheimer's & Dementia, 11*(3), 332-384.

Alzheimer's Discriminates: We Have the Power to End It [Brochure]. African American Network Against Alzheimer's.

Alzheimer's Disease and Its Impact on African Americans: At a Glance [Brochure]. Chicago, IL: Alzheimer's Association.

Basics of Alzheimer's Disease: What It Is and What You Can Do [Brochure]. Chicago, IL: Alzheimer's Association.

Heinze, Georg, and Michael Schemper. "A solution to the problem of separation in logistic regression." *Statistics in medicine* 21.16 (2002): 2409-2419.

Kleinbaum, David G., Mitchel Klein, and Erica Rihl. Pryor. *Logistic Regression: A Self-learning Text*. New York: Springer, 2010. Print.

Official Website of the Stage Play Forget Me Not. (n.d.). Retrieved March 14, 2016, from http://www.forgetmenotplay.com/

"Why Is Alzheimer's More Likely in Women?" *Alzheimers.net*. 2015. Web. 03 Mar. 2016. <http://www.alzheimers.net/8-12-15-why-is-alzheimers-more-likely-in-women/>.

5 Appendices

A Forget Me Not Questionnaire

Forget Me Not

Thursday, July 16, 2015 *Indianapolis, IN

HOW DID WE DO?

Please take a moment to evaluate today's event.

Name: Phone:

Email:

Race: ☐ African American ☐ Caucasian ☐ Other | Gender: ☐ Female ☐ Male | Date of Birth: _____/_____/_____ Month / Date / Year

1. How did you hear about the event? ☐ Church ☐ Friend ☐ Flier ☐ Postcard ☐ Website ☐ Event ☐ Word of Mouth ☐ EventBrite ☐ Other [*please write answer*]		
2. Did you enjoy the play?	☐ Yes	☐ No
3. Did you enjoy the panel discussion?	☐ Yes	☐ No
4. Was the event format and length appropriate?	☐ Yes	☐ No
5. Was the information helpful to you?	☐ Yes	☐ No

To what extent do you agree or disagree with the following statements?

6. "As a result of attending this event, I have a better understanding of why older African Americans need to get involved in Alzheimer's research."
[ ] **Strongly Disagree** [ ] **Disagree** [ ] **Neither Disagree or Agree** [ ] **Agree** [ ] **Strongly Agree**

7. "I would like to learn more about clinical trials and other opportunities that may be available to me."
[ ] **Strongly Disagree** [ ] **Disagree** [ ] **Neither Disagree or Agree** [ ] **Agree** [ ] **Strongly Agree**

8. Would you like to receive information about actions you can take to persuade Congress to provide more funding for Alzheimer's?	☐ Yes	☐ No
9. Do you want to be contacted about other events?	☐ Yes	☐ No
10. Every 3-6 months, would you be willing to go to the Brain Health Registry website, play simple online brain games and provide basic health and lifestyle information to help researchers understand the brain as it develops and ages? We will send you information on how if you check 'yes'.	☐ Yes	☐ No
11. What did you like most about the play and panel discussion?		
12. What did you learn that you did not know before?		

Thank You!

B SAS Code

```
/** Atlanta Code Identical to Every Survey Site Code **/

libname thesis "H:\THESIS";
proc format;
value yn        0 = "No"
                1 = "Yes"
                ;
value agree     0 = "Strongly Disagree"
                1 = "Disagree"
                2 = "Neither Disagree or Agree"
                3 = "Agree"
                4 = "Strongly Agree"
                ;
value race      1 = "African American"
                2 = "Caucasian"
                3 = "Other"
                ;
value gender    0 = "Male"
                    1 = "Female"
                    ;
value hear      1 = "Church"
                2 = "Friend"
                3 = "Flier"
                4 = "Postcard"
                5 = "Website"
                6 = "Event"
                7 = "Word of Mouth"
                8 = "EventBrite"
                9 = "Other"
                ;
run;
data thesis.atlanta2;
        set thesis.atlanta;
        rename Question_1 = Q1 Question_2 = Q2 Question_3=Q3 Question_4=Q4 Question_5=Q5
Question_6=Q6 Question_7=Q7
                Question_8=Q8 Question_9=Q9 Question_10 = Q10 Date_of_Birth = DOB;
run;
data thesis.atlanta2;
        set thesis.atlanta2;
        if Q1 = "Church" then Q1 = 1;
        else if Q1 = "Friend" then Q1 = 2;
        else if Q1 = "Flier" then Q1 = 3;
        else if Q1 = "Postcard" then Q1 =4;
        else if Q1 = "Website" then Q1 = 5;
        else if Q1 = "Event" then Q1 = 6;
        else if Q1 = "Word of Mouth" then Q1 = 7;
```

```
        else if Q1 = "EventBrite" then Q1 = 8;
        else if Q1 = "Eventbrite" then Q1 = 8;
        else if Q1 = "Other" then Q1 =9;
        else if Q1 = "Not Provided" then Q1 =.;
        else if Q1 = "not provided" then Q1 =.;
        Q_1 = Q1*1;
        format Q_1 hear.;
        label Q_1 = "Question 1";
run;
data thesis.atlanta2;
        set thesis.atlanta2;
        if Q2 = "Yes" then Q2 = 1;
        else if Q2 = "yes" then Q2 = 1;
        else if Q2 = "Not Provided" then Q2=.;
        else if Q2 = "No" then Q2 = 0;
        Q_2 = Q2*1;
        format Q_2 yn.;
        label Q_2 = "Question 2";
run;
data thesis.atlanta2;
        set thesis.atlanta2;
        if Q3 = "Yes" then Q3 = 1;
        else if Q3 = "No" then Q3 = 0;
        else if Q3 = "Not Provided" then Q3=.;
        Q_3 = Q3*1;
        format Q_3 yn.;
        label Q_3 = "Question 3";
run;
data thesis.atlanta2;
        set thesis.atlanta2;
        if Q4 = "Yes" then Q4 = 1;
        else if Q4 = "yes" then Q4 =1;
        else if Q4 = "No" then Q4 = 0;
        else if Q4 = "Not Provided" then Q4 = .;
        Q_4 = Q4*1;
        format Q_4 yn.;
        label Q_4 = "Question 4";
run;
data thesis.atlanta2;
        set thesis.atlanta2;
        if Q5 = "Yes" then Q5 = 1;
        else if Q5 = "yes" then Q5 = 1;
        else if Q5 = "Not Provided" then Q5 =.;
        Q_5 = Q5*1;
        format Q_5 yn.;
        label Q_5 = "Question 5";
run;
data thesis.atlanta2;
        set thesis.atlanta2;
        if Q6 = "Strongly Agree" then Q6 = 4;
        else if Q6 = "Agree" then Q6 = 3;
```

```
        else if Q6 = "Neither" then Q6 = 2;
        else if Q6 = "Disagree" then Q6 = 1;
        else if Q6 = "Strongly Disagree" then Q6 =0;
        else if Q6 = "Not Provided" then Q6 =.;
        Q_6 = Q6*1;
        format Q_6 agree.;
        label Q_6 = "Question 6";
run;
data thesis.atlanta2;
        set thesis.atlanta2;
        if Q7 = "Strongly Agree" then Q7 = 4;
        else if Q7 = "Agree" then Q7 = 3;
        else if Q7 = "Neither" then Q7 =2;
        else if Q7 = "Disagree" then Q7 =1;
        else if Q7 = "Strongly Disagree" then Q7 = 0;
        else if Q7 = "Not Provided" then Q7 =.;
        Q_7 = Q7*1;
        format Q_7 agree.;
        label Q_7 = "Question 7";
        if Q8 = "Yes" then Q8 = 1;
        else if Q8 = "No" then Q8 = 0;
        else if Q8 = "Not Provided" then Q8 =.;
        Q_8 = Q8*1;
        format Q_8 yn.;
        label Q_8 = "Question 8";
        if Q9 = "Yes" then Q9 = 1;
        else if Q9 = "No" then Q9 = 0;
        else if Q9 = "no" then Q9 =0;
        else if Q9 = "Not Provided" then Q9 =.;
        Q_9 = Q9*1;
        format Q_9 yn.;
        label Q_9 = "Question 9";
        if Q10 = "Yes" then Q10 = 1;
        else if Q10 = "yes" then Q10 = 1;
        else if Q10 = "No" then Q10 = 0;
        else if Q10 = "no" then Q10 = 0;
        else if Q10 = "Not Provided" then Q10 = .;
        Q_10 = Q10*1;
        format Q_10 yn.;
        label Q_10 = "Question 10";
        if Race = "African American" then Race = 1;
        else if Race = "Caucasian" then Race = 2;
        else if Race = "Other" then Race = 3;
        else if Race = "other" then Race =3;
        else if Race = "Not Provided" then Race = .;
        Race2 = Race*1;
        format Race2 race.;
        label Race2 = "Race";
        if Gender = "Female" then Gender = 1;
        else if Gender = "Male" then Gender = 0;
        else Gender = .;
```

```
        Gender2 = Gender*1;
        format Gender2 gender.;
        label Gender2 = "Gender";
run;
/*
proc contents data=thesis.atlanta2;
run;
proc freq data=thesis.atlanta2;
table Race;
run;
*/
data thesis.atlanta2;
        set thesis.atlanta2;
        drop Q1 Q2 Q3 Q4 Q5 Q6 Q7 Q8 Q9 Q10 Race Gender;
run;
data thesis.atlanta2;
        set thesis.atlanta2;
        rename Gender2 = Gender Race2 = Race;
run;
/**
proc contents data=thesis.atlanta2;
run;
proc freq data=thesis.atlanta;
        table Race;
run;
**/
/** Comparing newly coded dataset with original **/
/**
proc contents data=thesis.atlanta;
run;
proc contents data=thesis.atlanta2;
run;
proc freq data=thesis.atlanta;
        table Gender;
run;
proc freq data=thesis.atlanta2;
        table Gender;
run;
proc freq data=thesis.atlanta;
        table Race;
run;
proc freq data=thesis.atlanta2;
        table Race;
run;
proc freq data=thesis.atlanta;
        table Question_1;
run;
proc freq data=thesis.atlanta2;
        table Q_1;
run;
proc freq data=thesis.atlanta;
```

```
        table Question_2;
run;
proc freq data=thesis.atlanta2;
        table Q_2;
run;
proc freq data=thesis.atlanta;
        table Question_3;
run;
proc freq data=thesis.atlanta2;
        table Q_3;
run;
proc freq data=thesis.atlanta;
        table Question_4;
run;
proc freq data=thesis.atlanta2;
        table Q_4;
run;
proc freq data=thesis.atlanta;
        table Question_5;
run;
proc freq data=thesis.atlanta2;
        table Q_5;
run;
proc freq data=thesis.atlanta;
        table Question_6;
run;
proc freq data=thesis.atlanta2;
        table Q_6;
run;
proc freq data=thesis.atlanta;
        table Question_7;
run;
proc freq data=thesis.atlanta2;
        table Q_7;
run;
proc freq data=thesis.atlanta;
        table Question_8;
run;
proc freq data=thesis.atlanta2;
        table Q_8;
run;
proc freq data=thesis.atlanta;
        table Question_9;
run;
proc freq data=thesis.atlanta2;
        table Q_9;
run;
proc freq data=thesis.atlanta;
        table Question_10;
run;
proc freq data=thesis.atlanta2;
```

```
        table Q_10;
run;
**/
data thesis.atlanta2;
        set thesis.atlanta2;
        lecdate = "25jun2015"d;
        format DOB DDMMYY10.;

        age = floor((lecdate-DOB)/365);
run;
proc means data=thesis.atlanta2;
var age;
run;




/** MERGED DATASET CODE**/

LIBNAME thesis "H:\THESIS";
/* Combining datasets */
/*proc contents data=thesis.chicago2;
run;
proc contents data=thesis.houston2;
run;
proc contents data=thesis.indianapolis2;
run;
*/
proc format;
value yn        0 = "No"
                1 = "Yes"
                ;
value agree     0 = "Strongly Disagree"
                1 = "Disagree"
                2 = "Neither Disagree or Agree"
                3 = "Agree"
                4 = "Strongly Agree"
                ;
value race      1 = "African American"
                2 = "Caucasian"
                3 = "Other"
                ;
value gender    0 = "Male"
                        1 = "Female"
                        ;
value hear      1 = "Church"
                2 = "Friend"
                3 = "Flier"
                4 = "Postcard"
                5 = "Website"
```

```
                        6 = "Event"
                        7 = "Word of Mouth"
                        8 = "EventBrite"
                        9 = "Other"
                        ;
value location   1 = "Chicago"
                                2 = "Houston"
                                3 = "Indianapolis"
                                4 = "Atlanta"
                                5 = "Tampa"
                                ;
run;
data thesis.chicago2;
        set thesis.chicago2;
        rename Number___Street = Address  eMail_Address = Email;
        Location = 1;
        format Location location.;
        ID = _N_;
run;
data thesis.houston2;
        set thesis.houston2;
        rename Number___Street = Address eMail_Address = Email;
        Location = 2;
        format Location location.;
        ID = _N_ +490;
run;
data thesis.indianapolis2;
        set thesis.indianapolis2;
        rename Number___Street = Address eMail_Address = Email;
        Location =3;
        format Location location.;
        ID = _N_ + 636;
run;
data thesis.atlanta2;
        set thesis.atlanta2;
        Location = 4;
        format Location location.;
        ID = _N_ + 740;
run;
data thesis.tampa2;
        set thesis.tampa2;
        Location =5;
        format Location location.;
        ID = _N_ + 1204;
run;
data thesis.combo5;
        merge thesis.chicago2 thesis.houston2 thesis.indianapolis2 thesis.atlanta2 thesis.tampa2;
        by ID;
        drop Number Phone_Number;
run;
proc format;
```

```
value age_group                    0 = "<10"
                                       1 = "10-17"
                                       2 = "18-29"
                                       3 = "30-39"
                                       4 = "40-49"
                                       5 = "50-59"
                                       6 = "60-69"
                                       7 = "70-79"
                                       8 = "80-89"
                                       9 = "90-99"
                                       10 = "100<="
                                       ;
run;
data thesis.combo6;
        set thesis.combo5;
        age_group = 10;
        if age lt 100 then age_group =9;
        if age lt 90 then age_group = 8;
        if age lt 80 then age_group =7;
        if age lt 70 then age_group =6;
        if age lt 60 then age_group =5;
        if age lt 50 then age_group =4;
        if age lt 40 then age_group =3;
        if age lt 30 then age_group =2;
        if age lt 18 then age_group =1;
        if age lt 10 then age_group =0;
        format age_group age_group.;
run;
/*
proc freq data=thesis.combo6;
        table Q_6*age_group;
run;
*/
data thesis.combo7;
        set thesis.combo6;

        if age_group = 10 then delete;
        else if age_group =1 then delete;
        else if age_group =0 then delete;
run;
proc format;
value race_two  0 = "Other"
                                   1 = "African American"
                                   ;
run;
data thesis.combo7;
        set thesis.combo7;
        if Race = 1 then Race2 = 1;
        else if Race = 2 then Race2 = 0;
        else if Race = 3 then Race2 = 0;
        else Race2 = .;
```

```
        format Race2 race_two.;
run;
proc format;
value agree_two                 0 = "Do Not Agree"
                                1 = "Agree"
                                ;
run;
data thesis.combo7;
        set thesis.combo7;
        if Q_6 = 0 then Q_6_2 = 0;
        else if Q_6 = 1 then Q_6_2 = 0;
        else if Q_6 = 2 then Q_6_2 = 0;
        else if Q_6 = 3 then Q_6_2 = 1;
        else if Q_6 = 4 then Q_6_2 = 1;
        else Q_6_2 = .;
        if Q_7 = 0 then Q_7_2 = 0;
        else if Q_7 = 1 then Q_7_2 = 0;
        else if Q_7 = 2 then Q_7_2 = 0;
        else if Q_7 = 3 then Q_7_2 = 1;
        else if Q_7 = 4 then Q_7_2 = 1;
        else Q_7_2 = .;
        format Q_6_2 agree_two.;
        format Q_7_2 agree_two.;
run;
data thesis.combo7;
        set thesis.combo7;
        Q_6 = Q_6 +1;
        Q_7 = Q_7 +1;
run;
proc format;
value agree_three       1 = "Strongly Disagree"
                        2 = "Disagree"
                        3 = "Neither Disagree or Agree"
                        4 = "Agree"
                        5 = "Strongly Agree"
                        ;
run;
data thesis.combo7;
        set thesis.combo7;
        format Q_6 agree_three.;
        format Q_7 agree_three.;
        label   ID = "ID Number"
                Location = "Site of Survey"
                Q_6_2 = "Bivariate Question 6"
                Q_7_2 = "Bivariate Question 7"
                Race2 = "Bivariate Race"
                age = "Calculated Continuous Age"
                age_group = "Categorical Age"
                lecdate = "Survey Date"
                ;
        rename lecdate = survey_date;
```

```
        rename Location = survey_site;
run;
data thesis.combo6;
        set thesis.combo6;
        if age < 16 then age2 = age+100;
        else age2 = age;
        label age2 = "Age Assuming Older Population";

        if age > 99 then age3 = age-100;
        else age3 = age;
        label age3 = "Age Assuming Younger Population";
run;/**
proc ttest data=thesis.combo7;
        class Q_10;
        var age;
run;
proc npar1way data = thesis.combo7 wilcoxon;
        class Q_10;
        var age;
run;
proc glm data=thesis.combo7;
        class Q_7;
        model age = Q_7;
        means Q_7;
run;
quit;
proc freq data = thesis.combo7;
        tables Q_10*age_group/ fisher;
run;
proc logistic data = "D:\ologit"  order = data;
class pared(ref='0') public(ref='0') / param=reference;
model apply = pared public gpa;
run;
proc sort data = thesis.combo7;
        by Q_7;
run;
proc print data = thesis.combo7;
run;
proc logistic data = thesis.combo7  order = data;
model Q_6 = Gender ;
run;
proc logistic data=thesis.combo7 order = data;
class survey_site / param = ref;
model Q_7 =  survey_site Q_8 Gender;
run;
proc logistic data=thesis.combo7 ;
        class survey_site / param = ref;
        model Q_10 = Q_8 survey_site/ clodds = wald;
run;
proc logistic data = thesis.combo7  order = data;
        class survey_site / param=ref;
```

```
        model Q_6 (descending) = survey_site  /link=glogit selection = s slentry = 0.1;
run;

proc logistic data = thesis.combo7  order = data;
        class survey_site / param=ref;
        model Q_7 (descending) = survey_site Gender age Race2 Q_8 / clodds = wald selection=s slentry
= 0.1 slstay = 0.1;
run;
proc logistic data = thesis.combo7  order = data;
        class survey_site (ref = "Atlanta" param=ref);
        model Q_7 (descending)  = survey_site Q_8 Gender age Race2 / clodds = wald selection=f
slentry= 0.1;
run;
proc logistic data = thesis.combo7  order = data;
        class survey_site (ref = "Atlanta" param=ref);
        model Q_7 (descending) = survey_site Q_8 Gender age Race2  /link=glogit selection = b slstay =
0.1;
run;
*/
```