THE INFLUENCE OF RACE, AGE, COMORBIDITIES, AND BMI ON DISABILITY FOLLOWING STROKE IN ELDERLY PEOPLE LIVING IN THEIR OWN HOME

A Paper
Submitted to the Graduate Faculty
of the
North Dakota State University
of Agriculture and Applied Science

By

Izumi Endo

In Partial Fulfillment of the Requirements
for the Degree of
MASTER OF SCIENCE

Major Department:
Statistics

May 2020

Fargo, North Dakota

North Dakota State University
Graduate School

Title

THE INFLUENCE OF RACE, AGE, COMORBIDITIES, AND BMI ON DISABILITY FOLLOWING STROKE IN ELDERLY PEOPLE LIVING IN THEIR OWN HOME

By

Izumi Endo

The Supervisory Committee certifies that this ***disquisition*** complies with North Dakota State University's regulations and meets the accepted standards for the degree of

MASTER OF SCIENCE

SUPERVISORY COMMITTEE:

Megan Orr
Chair

Bong-Jin Choi

Andrea Huseth-Zosel

Approved:

06/26/2020 — Date

Rhonda Magel — Department Chair

ABSTRACT

Stroke is one of the major health issues in the United States. I explored different aspects of disability based on a history of stroke, race, comorbidities, age, and body mass index for the population of community dwelling stroke survivors.

Using a dataset drawn from the first wave of the longitudinal study of the National Social Life, Health, and Aging Project (Waite et al., 2019), analysis was performed. The dataset consists of a nationally representative sample of 3,005 community dwelling people between the ages of 57 to 86 years old at the time of recruitment.

The results demonstrated that the history of stroke, presence of comorbidities such as arthritis, chronic obstructive pulmonary disease, asthma, and heart failure, age, and body mass index significantly influenced the amount of disability an elderly person had. Performing screening and addressing the issues are essential to lower the amount of disability in the elderly population.

ACKNOWLEDGEMENTS

I would like to express my gratitude to the following people:

Dr. Megan Orr, my advisor, thank you for your help, instruction, advice, and encouragement over the course of my graduate study and the preparation for this paper.

Dr. Bong-Jin Choi, thank you for your help, instruction, and advice over the course of my graduate study.

Dr. Andrea Huseth-Zosel, thank you for giving me valuable ideas and comments.

Tammi Neville, thank you for your guidance and advice on my academic paper.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

BMI .... body mass index

COPD .... chronic obstructive pulmonary disease

NSHAP .... the National Social Life, Health, and Aging Project

1. INTRODUCTION

Stroke is one of the major health problems in the United States resulting in permanent disability and even death; however, a majority of stroke could be prevented. 795,000 Americans have a new or recurrent stroke each year (Benjamin et al., 2017). Stroke can cause permanent disability. It is estimated by Lloyd-Jones et al. (2010) that 22.5% of people who experience stroke are permanently disabled. In addition, stroke is the fifth leading cause of death in the United States according to Heron (2017). Because a majority of stroke happen without advance notice, the research suggests it is important to prevent a stroke event before it happens.

Studies have revealed that the amount of disability an elderly person has largely depends on a history of stroke, age, comorbidities, and body mass index (BMI). According to Lloyd-Jones et al. (2010), 20% of stroke survivors require institutional care at three months following a stroke because of disabilities caused by the stroke event. It is widely known that as a person ages, the functional capability to perform basic activities of daily living becomes limited. Fong (2019) stated that older persons with major chronic diseases had higher rates of incident disability across all activities of daily living. According to Vincent et al. (2010), the activities of walking, stair climbing, and chair rise were compromised in obese individuals, especially in individuals with a BMI greater than 35.

In this paper, I examined the population of stroke survivors in their home because not much research had been conducted on this population. I explored different aspects of disability based on a history of stroke, race, age, comorbidities, and BMI in the population. I hypothesized that a history of stroke, race, age, comorbidities, and BMI are associated with the amount of disability an elderly American has. In order to test the hypothesis, I first compared stroke survivors with age and race matched stroke free controls. Next, I compared black stroke subjects with white stroke subjects. Lastly, a Poisson regression model predicting the number of disability items on other variables was developed.

2. LITERATURE REVIEW

2.1. Stroke

Stroke has been one of the major health issues for a long time in the United States resulting in permanent disability or even death. Benjamin et al. (2017) estimated about 795,000 people in the United States had a new or recurrent stroke each year. Furthermore, stroke remains the fifth leading cause of death (Heron, 2017) and causes an estimated 140,000 deaths per year (Yang et al., 2017). There are about seven million stroke survivors in this country according to Virani et al. (2016).

There are different types of stroke, and the classification depends on what part of the body is involved and if the blood vessel is occluded, ischemic, or ruptures, hemorrhagic. Of all strokes, ischemic stroke, which occurs when blood vessels are occluded, is most prevalent in the United States. According to Koton et al. (2014), the mortality rate of ischemic stroke is 57.4%, and that of hemorrhagic stroke is 67.9%.

There are two categories of risk factors for stroke: modifiable and non-modifiable risk factors. Well-documented modifiable risk factors are hypertension, cigarette smoking, diabetes, dyslipidemia, atrial fibrillation, asymptomatic carotid stenosis, sickle cell disease, postmenopausal hormone therapy, oral contraceptives, diet, physical inactivity, and obesity (Goldstein et al., 2011). According to the same authors, non-modifiable risk factors are age, sex, low birth weight, race/ethnicity, and genetic factors.

Prevention has three classifications, and primary prevention is deemed imperative for stroke because stroke can cause death and permanent disability. The prevention classifications are primary, secondary, and tertiary prevention. Primary prevention is intervening before health issues occur. Secondary prevention is screening and identifying health issues in the earliest stage prior to the onset of signs and symptoms. Tertiary prevention is managing health issues after diagnosis to slow or stop the progression. For stroke, primary intervention is especially important for two reasons. The first is that more than 77% of strokes are first events (Lloyd-Jones et al., 2010). The second is that stroke survivors often end up functionally incomplete during recovery because adults lack the ability to regenerate the damaged portion of the brain completely (Cramer et al., 2007). Scientists have conducted research to find a way to prevent stroke. Larsson et al. (2015) found pursuing a healthy lifestyle is beneficial to prevent stroke in a high-risk group. The researchers described a healthy lifestyle as maintaining a BMI of 18.5 to 25, moderately exercising for 150 minutes or more a week, modest alcohol consumption, discontinuing smoking, and consuming a low-risk diet.

Stroke is also known to cause serious, long-term disability. This disability affects not only the patients themselves but also their family members. Twenty percent of stroke survivors require institutional care three months after a stroke according to Lloyd-Jones et al. (2010). According to Buntin (2010), more than two thirds of people who have experienced stroke utilize rehabilitation services following hospitalization. Lloyd-Jones et al. (2010) also estimated that 22.5% of stroke survivors were permanently disabled.

When a person has an ischemic stroke, the person can get treated with thrombolytic therapy if the treatment is initiated within three hours after onset of the stroke, so the goal is to transport a person with a suspected stroke to an emergency department in a timely manner. The researchers at the Massachusetts Medical Society (1995) compared acute ischemic stroke patients treated with thrombolytic therapy and those treated with placebo as a control. The researchers found the patients treated with thrombolytics were 30% more likely to have less disability three months after the stroke incident. Tong et al. (2012) reported more than one fourth of acute ischemic stroke patients were brought to an emergency department within three hours after the onset of the stroke. It is imperative to call 911 and use emergency medical services if an incidence of stroke is suspected. That is because the staff in the emergency department can prepare for patient arrival, and it results in decreased time to initiate thrombolytic therapy. Adeoye et al. (2009) pointed out only half of stroke patients utilize emergency medical services.

2.2. Disability Risk Factors

Various research studies have been conducted to investigate the association between disability and comorbidities using data sets collected in the United States. The findings consistently show a positive association. Jindai et al. (2016) performed analysis on data of community dwelling Americans aged 65 years or older. They found a prevalence of having two or more concurrent health issues is 67% in this group of people. They also noted an increase in the number of functional limitations associated with each additional chronic condition. Fong (2019) compared groups of Americans aged 80 years or older with and without major chronic diseases. It was found by the researcher that older persons with major chronic diseases have higher rates of incident disability across all activities of daily living.

The association between disability and BMI is an emerging topic. BMI is a person's weight in kilograms divided by the square of height in meters (World Health Organization, 1995). It is categorized into underweight (15-19.9), healthy weight (20-24.9), overweight (25.0-29.9), and obese (30.0 and above). Himes and Reynolds (2011) analyzed data collected from persons aged 65 years and older in the United States. The researchers concluded obesity was associated with greater risk of falling in the elderly population and a higher risk of greater activity of daily living disability after a fall. Vincent et al. (2010) systematically reviewed a total of 13 cross-sectional and 15 longitudinal studies based on physical assessments of mobility in the obese older population. A consistent finding made by the scientists was that the activities of walking, stair climbing, and chair rise were compromised in obese individual, especially in individuals with a BMI greater than 35.

2.3. Racial Difference in Comorbidity

It is well recognized that patterns of comorbidity differ in races, and minority populations may not fully benefit from research findings. According to Go et al. (2013), black Americans have the highest mortality rates because of cardiovascular disease. The same authors also found that the prevalence rate of hypertension is 44% which is the world's highest among all countries' black residents. Brown et al. (2017) analyzed the dataset of National Health and Nutrition Examination Survey. They found that 42.8% of black individuals born in the United States had hypertension and only 27.4% of foreign-born blacks had the same medical issue.

2.4. Activity of Daily Living Limitation in Nursing Home Residents

According to Lair and Lefkowitz (1990), the percentages of nursing home residents requiring personal assistance for walking, dressing, bathing, feeding, bed/chair transfer, and using the toilet are 63.2%, 76.0%, 87.5%, 33.8%, 61.7%, and 63.6% respectively. Walking is included as a measure even though it is not considered an activity of daily living.

3. METHOD

3.1. Data Set

In order to conduct a database analysis, I used the data set drawn from the first wave of the longitudinal study of the National Social Life, Health, and Aging Project (NSHAP, Waite et al., 2019). The NSHAP is a population-based study in the United States. It consists of a nationally representative sample of persons living in their home between the ages of 57 to 86 years at the time of recruitment. 3,005 subjects provided written informed consent and successfully participated in the study. The overall response rate was 75.5%. Detailed in-home interviews were completed by a trained professional between July 2005 and March 2006 for the first wave. The NSHAP protocol was reviewed and approved by the institutional review boards of the National Opinion Research Center, the University of Chicago, and North Dakota State University.

3.2. Variables

In this analysis, black and white persons who previously had a stroke were selected as stroke subjects. Stroke diagnosis was based on a response to the following question: "Has a medical doctor ever told you that you have any of the following conditions: stroke, cerebrovascular accident, blood clot or bleeding in the brain, or transient ischemic attack?" In the question, medical doctors include only specialists, general practitioners, and osteopaths, but not chiropractors, dentists, nurses, or nurse practitioners.

Demographic information, medical conditions, and disability measures were collected from the participants in NSHAP by self-report. For demographic information, NSHAP had age, educational level, marital status, insurance coverage, household income, and household assets. All the demographic information was included in the current study. In NSHAP, educational level was recorded in four levels (less than high school, high school/equivalent, vocational certificate/some college/associate, and bachelors/more). In the current study, educational level was dichotomized as high school graduate or not. For marital status, NSHAP had six levels of responses: married, living with a partner, separated, divorced, widowed, or never married. In the current study, marital status was dichotomized as married or not. NSHAP had four insurance coverage questions (Medicare, Medicaid, private insurance, and veteran's administration), and the participants answered yes or no to each question. In the current study, insurance coverage was dichotomized as yes or not. If a NSHAP participant answered yes to at least one insurance coverage question, the response was counted as yes to insurance coverage in the current study. For household income, NSHAP had four levels of responses (0-24,999, 25,000-49,999, 50,000-99,999, or 100k/higher). In the current study, these responses were dichotomized as >=25,000 or not. For household assets, NSHAP had five levels of responses (0-9,999, 10,000-49,999, 50,000-99,999, 100,000-499,999, or 500k/higher). In the current study the responses were dichotomized as >=100,000 or not. The variables of educational level, marital status, insurance coverage, household income, and household assets were dichotomized in the current study so that the data analysis was simplified without losing information.

BMI, heart attack, heart failure, arthritis, ulcers, chronic obstructive pulmonary disease (COPD), asthma, hypertension, and diabetes were recorded in NSHAP. In the current analysis, all these variables were included as comorbid conditions.

The primary exposure of interest in this study was disability measures. NSHAP recorded walking one block, walking across a room, dressing, bathing/showering, eating, getting in and out of bed, using the toilet, and health status compared to same-age-peers, so all these variables were also included in the current study. In NSHAP, the responses to walking one block, walking across a room, dressing, bathing/showering, eating, getting in and out of bed, and using the toilet in NSHAP were no difficulty, some difficulty, much difficulty, unable to do, or if volunteered – have never done. These measures were dichotomized as with or without difficulty in the current study. For example, if an NSHAP participant's response to dressing was no difficulty, the response in the current study was the person had no difficulty in dressing. If an NSHAP participant's response to dressing was one of the responses other than no difficulty, the response in the current study was the person had difficulty in dressing. In NSHAP, health status compared to same-age-peers was recorded as much worse, somewhat worse, about the same, somewhat better, or much better. These responses were dichotomized as worse or better/same in the current study. All the disability measures were dichotomized in the current study so that the data analysis was simplified without losing information.

3.3. Statistical Analysis

It is well known that functional capability to perform basic activities of daily living becomes limited as people get older. In order to minimize the influence of age-related disability, the first analysis was a comparison of demographic information, comorbid conditions, and disability measures between stroke subjects and age- and race-matched controls. The statistical analysis was performed using paired-samples t-tests for continuous variables and using McNemar's test for categorical variables.

The second analysis was performed by comparing black stroke subjects and white stroke counterparts. Demographic information, comorbid conditions, and disability measures of the two groups were compared using two sample t-tests for continuous variables and chi-square tests for categorical variables.

Lastly, the number of disability items was predicted using a generalized linear mixed model. The number of disability items included walking one block, walking across room, dressing, bathing/showering, eating, getting in and out of bed, and using the toilet. The independent variables of age, arthritis, ulcers, COPD, asthma, stroke, hypertension, diabetes, race, high school graduates, married, heart attack, heart failure, and BMI were assumed to have fixed effects. The variable of case/control pair was assumed to have random effects. I excluded the variables from the model if they had missing values of more than 10% of the observations. In this analysis, number of disability items was assumed to follow a Poisson distribution with a log link function. Variance inflation factors were calculated to detect potential multicollinearity. SAS Statistical Analysis System Version 9.4. was used for the study.

4. RESULTS

Of the 3,005 subjects who participated in the NSHAP, 268 subjects previously had a stroke. Among the 268 stroke subjects, 62 were black and 187 were white. The demographic information, clinical characteristics, and disability measures of stroke subjects and the age- and race-matched stroke free controls are detailed in table 1. For the demographic information, there were no statistically significant differences in mean age, percentage of individuals who were high school graduates, percentage of individuals who were married, or percentage of individuals covered by insurance between the stroke subjects and controls. Significant differences were noted in household income and assets with controls having a higher percentage of both a household income of at least $25,000 and household assets of at least $100,000. For the clinical characteristics, significant differences were seen in all health issues other than mean BMI. That is, stroke survivors consistently had more medical issues proportional to the controls. Figure 1 describes that the differences in BMI of stroke subjects and the age- and race-matched stroke controls were approximately normally distributed, indicating the major assumption of paired t-test was met. Table 2 shows the correlation coefficient of BMI between stroke subjects and controls. The correlation indicates negative weak linear relationship ($r = -0.13$) and was not significant at a 0.05 significance level. Regarding disability measures, significant differences were identified for all items, with a higher percentage of stroke subjects experiencing each disability compared to controls. That is, stroke subjects had more evident disability. This result is consistent with findings from the literature review.

Table 1. Comparisons of Stroke Cases versus Age and Race Matched Stroke Free Controls

	Cases (n=249)	Controls (n=249)	P Value
Demographic Information			
Mean Age (years)	71.8	71.8	1
High School Graduates (%)	73.1 (182/249)	77.5 (193/249)	0.2
Married (%)	51.0 (127/249)	55.8 (139/249)	0.3
Covered by Insurance (%)	100 (196/196)	100 (199/199)	1
Household Income (>=$25,000) (%)	54.1 (92/170)	67.7 (115/170)	0.02*
Household Asset (>=$100,000) (%)	54.2 (78/144)	69.1 (105/152)	0.02*
Clinical Characteristics			
Mean BMI (kg/m^2)	28.9	28.2	0.3
Heart Attack (%)	29.8 (74/248)	9.3 (23/249)	<0.0001*
Heart Failure (%)	25.3 (62/245)	7.3 (19/247)	<0.0001*
Arthritis (%)	65.06 (162/249)	53.0 (132/249)	0.007*
Ulcers (%)	22.5 (56/249)	10.4 (26/249)	<0.0001*
COPD (%)	19.7 (49/249)	8.4 (21/249)	0.0005*
Asthma (%)	12.9 (32/249)	6.8 (17/249)	0.02*
Hypertension (%)	73.5 (183/249)	56.2 (140/249)	<0.0001*
Diabetes (%)	30.1 (75/249)	14.1 (35/249)	<0.0001*
Disability Measures			
Difficulty Walking One Block (%)	53.8 (134/249)	28.1 (70/249)	<0.0001*
Difficulty Walking across Room (%)	29.7 (74/249)	14.1 (35/249)	<0.0001*
Difficulty Dressing (%)	32.7 (81/248)	12.5 (31/249)	<0.0001*
Difficulty Bathing/Showering (%)	24.9 (62/249)	9.6 (24/249)	<0.0001*
Difficulty Eating (%)	14.1 (35/249)	5.2 (13/249)	0.0007*
Difficulty Getting in/out of Bed (%)	18.5 (46/249)	10.8 (27/249)	0.02*
Difficulty Using Toilet (%)	21.7 (54/249)	10.4 (26/249)	0.0007*
Worse Health Status Comparing to Same Age Peers (%)	26.2 (50/191)	5.8 (12/206)	<0.0001*

*: p-value < 0.05

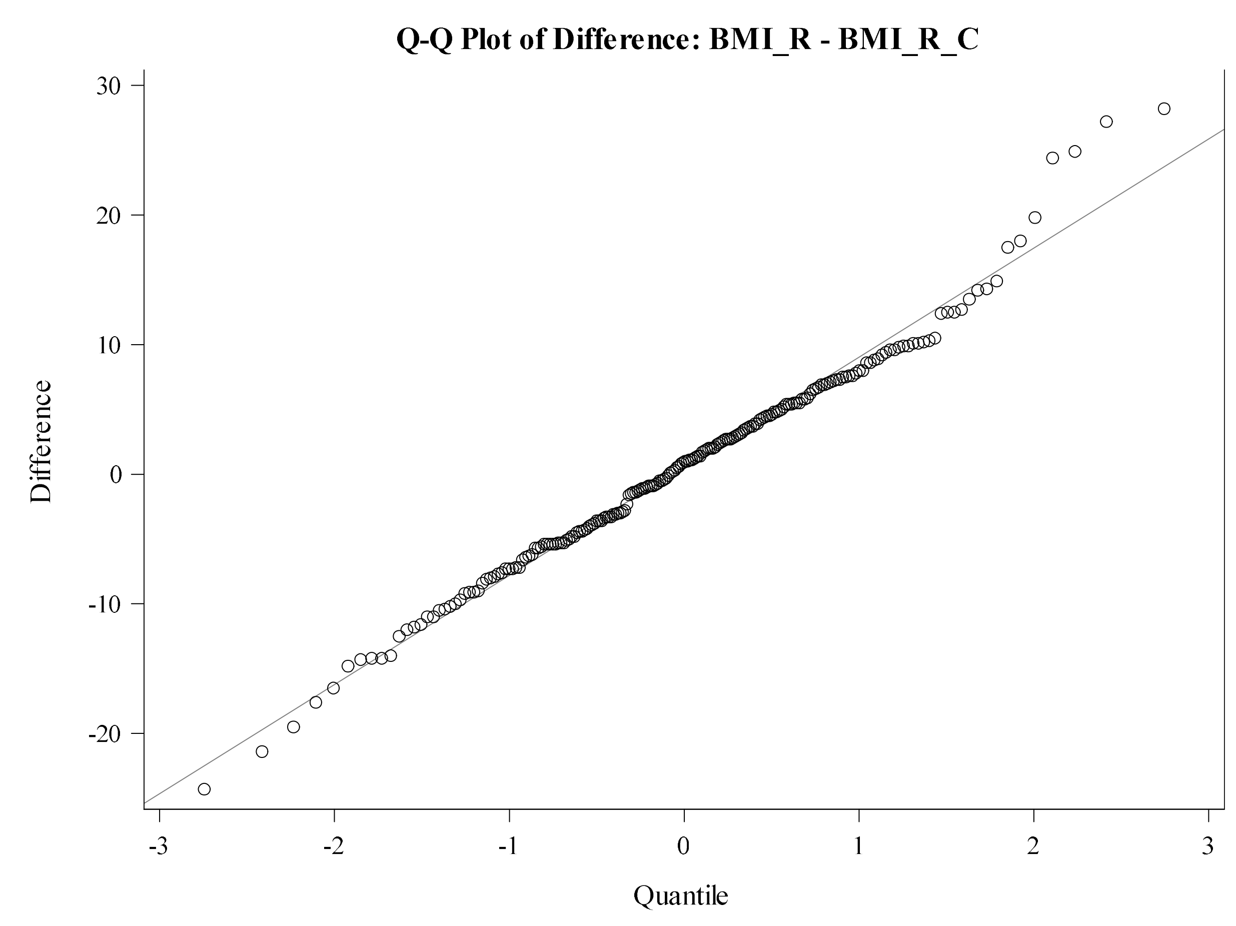

Figure 1. Q-Q Plots of BMI Difference between Stroke Subjects and Controls

Table 2. Correlation Coefficient of BMI between Stroke Subjects and Controls

Pearson Correlation Coefficients Prob > |r| under H0: Rho=0 Number of Observations		
	BMI_R	**BMI_R_C**
BMI_R	1.00000 222	-0.13023 0.0614 207
BMI_R_C	-0.13023 0.0614 207	1.00000 231

In table 3, demographic information, clinical characteristics, and disability measures of the stroke subjects by race are detailed. Regarding demographic information, black stroke subjects were significantly younger, less educated, and less likely to be married compared to white stroke subjects. Additionally, black stroke subjects had a significantly lower household mean income and mean household assets than their white counterparts. In figure 2, both Q-Q plots have an S shape indicating a distribution with heavier tails than a normal distribution; however, the sample mean age is still considered to be approximately normal for each group because of the large sample size. The significant differences in clinical characteristics were that black stroke subjects were more likely to have heart failure, hypertension, and diabetes. Other health issues do not show a significant difference between the groups. Figure 3 shows the normality assumption of BMI was met. For disability measures, difficulty in walking a block, eating, and using the bathroom were statistically more likely for black stroke subjects compare to white stroke subjects. There were no significant differences in the likelihood of difficulty in walking across the room, dressing, bathing/showering, getting in and out of bed, and health status between black and white stroke subjects. The sample size was sufficiently large for all chi-square tests, as no cell in the table had an expected count of less than one or no more than 20% of the cells had an expected count less than five.

Table 3. Comparisons of Stroke Cases by Race

	Black (n=62)	White (n=187)	P Value
Demographic Information			
Mean Age (years)	69.9	72.4	0.03*
High School Graduates (%)	46.8 (29/62)	81.8 (153/187)	<0.0001*
Married (%)	37.1 (23/62)	55.6 (104/187)	0.01*
Covered by Insurance (%)	100 (40/40)	100 (156/156)	1
Household Income (>=$25,000) (%)	31.0 (13/42)	61.7 (79/128)	0.0005*
Household Asset (>=$100,000) (%)	19.0 (7/37)	66.4 (71/107)	<0.0001*
Clinical Characteristics			
Mean BMI (kg/m^2)	30.1	28.5	0.2
Heart Attack (%)	33.9 (21/62)	28.5 (53/186)	0.4
Heart Failure (%)	35.5 (22/62)	21.9 (40/183)	0.03*
Arthritis (%)	72.6 (45/62)	62.6 (117/187)	0.2
Ulcers (%)	22.6 (14/62)	22.5 (42/187)	1
COPD (%)	14.5 (9/62)	21.4 (40/187)	0.2
Asthma (%)	11.3 (7/62)	13.4 (25/187)	0.7
Hypertension (%)	83.9 (52/62)	70.1 (131/187)	0.03*
Diabetes (%)	40.3 (25/62)	26.7 (50/187)	0.04*
Disability Measures			
Difficulty Walking One Block (%)	69.4 (43/62)	48.7 (91/187)	0.005*
Difficulty Walking across Room (%)	32.3 (20/62)	28.9 (54/187)	0.6
Difficulty Dressing (%)	35.5 (22/62)	31.7 (59/186)	0.6
Difficulty Bathing/Showering (%)	32.3 (20/62)	22.5 (42/187)	0.1
Difficulty Eating (%)	22.6 (14/62)	11.2 (21/187)	0.03*
Difficulty Getting in/out of Bed (%)	22.6 (14/62)	17.1 (32/187)	0.3
Difficulty Using Toilet (%)	33.9 (21/62)	17.7 (33/187)	0.007*
Worse Health Status Comparing to Same Age Peers (%)	35.1 (13/37)	24.0 (37/154)	0.2

*: p-value < 0.05

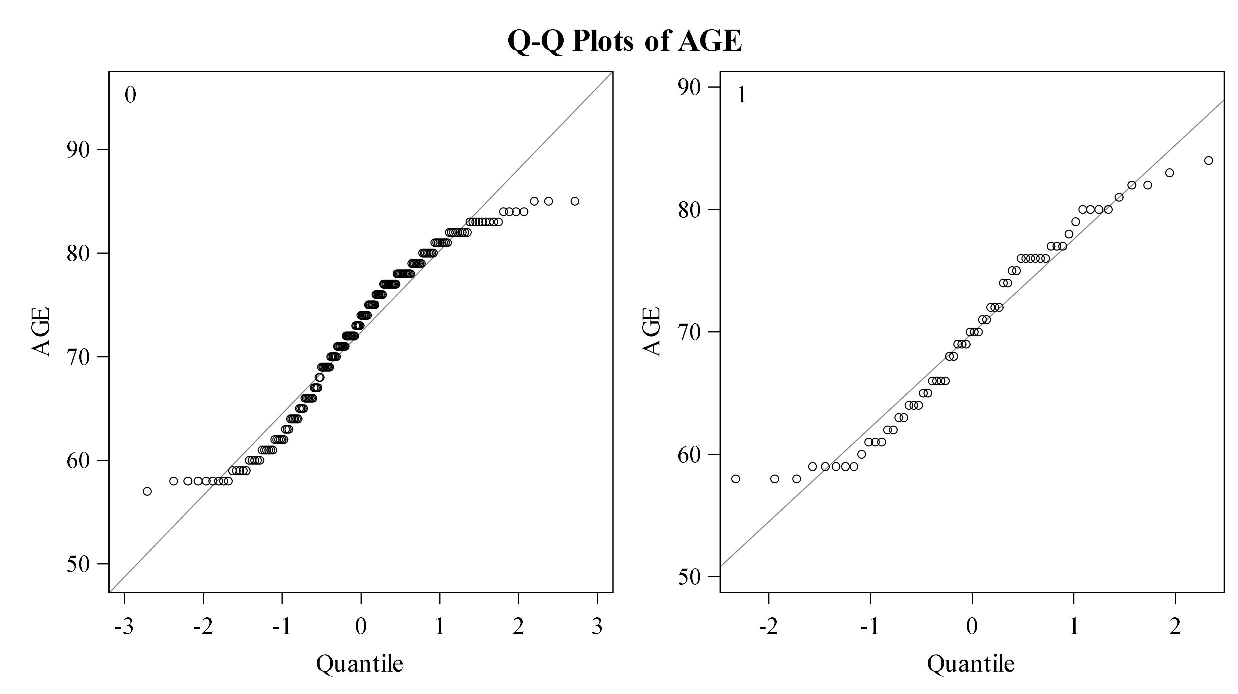

Figure 2. Q-Q Plots of Age for Black Subjects (left) and White Counterparts (right)

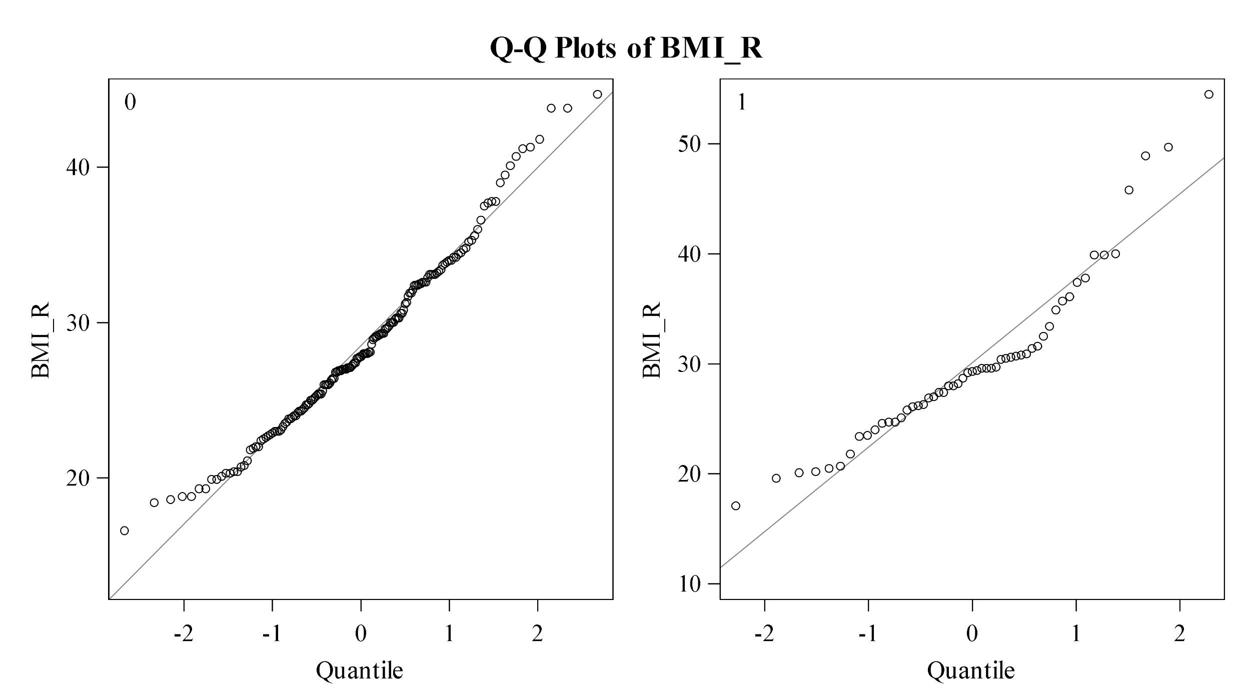

Figure 3. Q-Q Plots of BMI for Black Subjects (left) and White Counterparts (right)

Generalized linear mixed modeling was performed to predict the number of disability items. The number of disability items was assumed to follow a Poisson distribution with a log link function. Age, arthritis, ulcers, COPD, asthma, stroke, hypertension, diabetes, race, high school graduate, married, heart attack, heart failure, and BMI were included as independent variables in this model. The variables of insurance, income, and asset were excluded from the model because they had missing values of more than 10% of the observations. The case/control pair an observation came from was included as a random effect. The total sample size for this analysis was 447 after observations with missing values were removed. The established model is below.

$$\begin{aligned}\log(\textit{number of disability items}) \\ &= -2.5011 + 0.01815x_1 + 0.2988x_2 + 0.1441x_3 + 0.3584x_4 + 0.3944x_5 \\ &+ 0.4615x_6 + 0.1683x_7 + 0.07083x_8 + 0.009557x_9 - 0.2753x_{10} \\ &- 0.2525x_{11} - 0.1183x_{12} + 0.321x_{13} + 0.02629x_{14}\end{aligned}$$

X1 is age (years), X2 is arthritis (0: no, 1: yes), X3 is ulcers (0: no, 1: yes), X4 is COPD (0: no, 1: yes), X5 is asthma (0: no, 1: yes), X6 is stroke (0: no, 1: yes), X7 is hypertension (0: no, 1: yes), X8 is diabetes (0: no, 1: yes), X9 is race (0: white, 1: black), X10 is high school graduate (0: no, 1: yes), X11 is married (0: no, 1: yes), X12 is heart attack (0: no, 1: yes), X13 is heart failure (0: no, 1: yes), and X14 is BMI (kg/m^2). The details of the estimates are presented in table 4. According to the p values, the variables of age, arthritis, COPD, asthma, stroke, heart failure, high school graduates, and married have significant effects at a significance level of 0.05. That is, a senior is expected to have more disability items if he or she is older, has arthritis, has COPD, has asthma, or has had a stroke. A senior who is a high school graduate or is married is associated with a decrease in disability items. It is found that the variable of race does not have significant effects on the number of disability items an elderly individual has.

Furthermore, an interpretation of a significant coefficient using an example of an estimate of stroke is as follows. The estimate of the variable of stroke from the model is 0.4615. That is, the estimated difference in the log of the number of disability items between a person with and without a history of stroke. If all other independent variables in the model are the same and the only difference is with or without a history of stroke, the estimated number of disability items can be expressed as below (let y1 be the estimated number of disability items of a person with a history of stroke and y2 be that of one without).

$$\log(y1) - \log(y2) = 0.4615$$

Then, the equation above can be modified into the equation below using the properties of logs.

$$\log\left(\frac{y1}{y2}\right) = 0.4615$$

Exponentiating both sides of the equation above, and then the ratio of y1/y2 is 1.59. That is, a senior with history of stroke has an estimated 1.59 times the number of disability items compared to one without history of stroke when all other independent variables remain the same. Additionally, the 95% confident interval for the ratio in table 4 is in log, so the ratio of interval in number of disability items is from 1.314 to 1.916 after exponentiating. That means, with 95% confidence, individuals with history of stroke have 1.314 to 1.916 times the number of disability items compared to those without history of stroke when all other factors remain the same.

In order to assess multicollinearity, variance inflation factors were calculated. The largest variance inflation factor among the independent variables was 1.36060 corresponding to the variable of heart failure. That is, the independent variables in the model did not show any moderate or severe multicollinearity. Lastly, the value of Pearson chi-square / degree of freedom was 1.07, indicating the dataset fits the model sufficiently.

Table 4. Solution for Fixed Effects of Generalized Linear Mixed Model Analysis

Effect	Estimate	Standard Error	t Value	P Value	95% Confidence Interval	
Intercept	-2.5011	0.6489	-3.85	0.0001	-3.7793	-1.2230
Age (years)	0.01815	0.007003	2.59	0.0103	0.004338	0.03197
Arthritis (0: no, 1: yes)	0.2988	0.1145	2.61	0.0098	0.07295	0.5247
Ulcers (0: no, 1: yes)	0.1441	0.1284	1.12	0.2634	-0.1093	0.3974
COPD (0: no, 1: yes)	0.3584	0.1508	2.38	0.0185	0.06096	0.6558
Asthma (0: no, 1: yes)	0.3944	0.1593	2.48	0.0142	0.08004	0.7087
Stroke (0: no, 1: yes)	0.4615	0.09560	4.83	<.0001	0.2729	0.6501
Hypertension (0: no, 1: yes)	0.1683	0.1207	1.39	0.1648	-0.06978	0.4065
Diabetes (0: no, 1: yes)	0.07083	0.1302	0.54	0.5872	-0.1861	0.3278
Race (0: white, 1: black)	0.009557	0.1311	0.07	0.9420	-0.2491	0.2682
High School Graduates (0: no, 1: yes)	-0.2753	0.1216	-2.26	0.0248	-0.5152	-0.03532
Married (0: no, 1: yes)	-0.2525	0.1098	-2.30	0.0226	-0.4691	-0.03582
Heart Attack (0: no, 1: yes)	-0.1183	0.1487	-0.80	0.4273	-0.4116	0.1750
Heart Failure (0: no, 1: yes)	0.3210	0.1509	2.13	0.0347	0.02331	0.6187
BMI (kg/m^2)	0.02629	0.009185	2.86	0.0047	0.008172	0.04441

5. CONCLUSION

In this paper, I explored different facets of disability based on history of stroke, race, comorbidities, age, and BMI of American seniors living in their own home. Even though I had hypothesized that those factors determined the amount of disability, the results of the analysis showed that history of stroke, presence of comorbidities such as arthritis, COPD, asthma, and heart failure, age, and BMI were significantly associated with the amount of disability a senior had. It was found that the variable of race did not have a significant association with the amount of disability for seniors.

In the current study, the percentages of difficulty performing activity of daily living in the population of community-dwelling stroke survivors for walking one block, walking across room, dressing, bathing/showering, eating, getting in/out of bed, and using the toilet were 53.8%, 29.7%, 32.7%, 24.9%, 14.1%, 18.5%, and 21.7% respectively. The percentages of nursing home residents requiring personal assistance for walking, dressing, bathing, feeding, bed/chair transfer, and using the toilet were 63.2%, 76.0%, 87.5%, 33.8%, 61.7%, and 63.6% respectively, according to Lair and Lefkowitz (1990). These comparisons suggest that the limitations of stroke survivors living in their own homes are proportionally less than nursing home residents in all activity measures. However, more updated data on performing these activities among nursing home residents are warranted.

I believe the results of this study are beneficial for health policy making. The variables significantly associated with the amount of disability include ones which are non-modifiable; however, preventing a stroke to occur, preventing to have comorbidities, and maintaining a BMI within a normal range are the keys to lower the amount of disability an elderly person has. Performing screening and addressing the issues are deemed essential.

6. FUTURE RESEARCH

For future work, I recommend analyzing changes in disability of stroke survivors living in their own home over a period of time using the data sets of all three waves of NSHAP. The demographic information or comorbid conditions may potentially influence the changes in disability of the group of population. Also, use of propensity score matching (Rosenbaum & Rubin, 1983) will be beneficial to compare two groups.

If complete information on comorbid conditions, such as full medical records or other documents, is able to be obtained along with demographic information and disability measures, it may be beneficial to calculate the Charlson comorbidity score (Charlson et al., 1987) to evaluate its potential influence on the disability measures.

REFERENCES

Adeoye, O., Lindsell, C., Broderick, J., Jauch, E., Alwell, K., Moomaw, C., Flaherty, M., Pancioli, A., Kissela, B., & Kleindorfer, D. (2009). Emergency medical services utilization by stroke patients: population-based Study. *The American Journal of Emergency Medicine*, 27(2), 141-145.

Benjamin, E. J., Blaha, M. J., Chiuve, S. E., Cushman, M., Das, S. R., Deo, R., Ferranti, S. D., Floyd, J., Fornage, M., Gillespie, C., Isasi, C. R., Jiménez, M. C., Jordan, L. C., Judd, S. E., Lackland, D., Lichtman, J. H., Lisabeth, L., Liu, S., Longenecker, C. T., Mackey, R. H., Matsushita, K., Mozaffarian, D., Mussolino, M. E., Nasir, K., Neumar, R. W., Palaniappan, L., Pandey, D. K., Thiagarajan, R. R., Reeves, M. J., Ritchey, M., Rodriguez, C. J., Roth, G. A., Rosamond, W. D., Sasson, C., Towfighi, A., Tsao, C. W., Turner, M. B., Virani, S. S., Voeks, J. H., Willey, J. Z., Wilkins, J. T., Wu, J. H., Alger, H. M., Wong, S. S., & Muntner, P. (2017). Heart disease and stroke statistics – 2017 update: report from the American Heart Association. *Circulation,* 135, 146-603.

Brown, A. G. M., Houser, R. F., Mattei, J., Mozaffarian, D., Lichtenstein, A. H., & Folta, S. (2017). Hypertension among US-born and foreign-born non -Hispanic blacks. *Journal of Hypertension,* 35(12), 2380-2387.

Buntin, M. B., Colla, C. H., Deb, P., Sood, N., & Escarce, J. J. (2010). Medicare spending and outcomes after postacute care for stroke and hip fracture. Medical Care, 4(9), 776-784.

Charlson, M. E., Pompei, P., Ales, K. L., & MacKenzie, C. R. (1987). A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *Journal of Chronic Diseases*, 40(5), 373-383.

Fong, J.H. (2019). Disability incidence and functional decline among older adults with major chronic diseases. *BMC Geriatric* 19, Article 323. https://doi.org/10.1186/s12877-019-1348-z

Go, A. S., Mozaffarian, D., Roger, V. L., Benjamin, E. J., Berry, J. D., Borden, W. B., Bravata, D. M., Dai, S., Fort, E. S., Fox, C. S., Franco, S., Fullerton, H. J., Gillespie, C., Hailpern, S. M., Heit, J. A., Howard, V. J., Huffman, M. D., Kissela, B. M., Kittner, S. J., Lackland, D. T., Lichtman, J. H., Lisabeth, L. D., Magid, D., Marcus, G. M., Marelli, A., Matchar, D. B., McGuire, D. K., Mohler, E. R., Moy, C. S., Mussolino, M. E., Nichol, G., Paynter, N. P. Schreiner, P. J., Sorlie, P. D., Stein, J., Turan, T. N., Virani, S. S., Wong, N. D., Woo, D., & Turner, M. B. (2013). Executive summary: heart disease and stroke statistics – 2013 update. *Circulation*, 127, 143-152.

Goldstein, L. B., Bushnell, C. D., Adams, R. J., Appel, L. J., Braun, L. T., Chaturvedi, S., Creager, M. A., Culebras, A., Eckel, R. H., Hart, R. G., Hinchey, J. A., Howard, V. A., Jauch, E. C., Levine, S. R., Meschia, J. F., Moore, W. S., Nixon, J. V., & Pearson, T. A. (2011). Guidelines for the primary prevention of stroke. *Stroke*, 42(2), 517-584.

Heron, M. (2019). Deaths: Leading causes for 2017. *National Vistal Statistics Reports,* 68, 1-77.

Himes, C, & Reynolds, S. L. (2011). Effect of obesity on falls, injury, and disability. *Journal of the American Geriatric Society*, 60(1), 124-129.

Jindai, K., Nielson, C. M., Vorderstrasse, B. A., & Quiñones, A. R. (2016). Multimorbidity and functional limitations among adults 65 or older, NHANES 2005–2012. *Preventing Chronic Disease*, 13.

Koton, S., Schneider, A. L. C., Rosamond, W. D., Shahar, E., Sang, Y., Gottesman, R. F., & Coresh, J. (2014). Stroke incidence and mortality trends in US communities, 1987 to 2011. *Journal of the American Medical Association*, 312(3), 259-268.

Larsson, S. C., Akesson, A., & Wolk, A. (2015). Primary prevention of stroke by a healthy lifestyle in a high-risk group. *American Academy of Neurology.* https://doi.org/10.1212/WNL.0000000000001637

Lair, T. & Lefkowitz, D. (1990). *Mental health and functional status of residents of nursing and personal care homes.* Department of Health and Human Services.

Lloyd-Jones, D., Adams, R. J., Brown, T. M., Carnethon, M., Dai, S., Simone, G. D., Ferguson, T. B., Ford, E., Furie, K., Gillespie, C., Go, A., Greenlund, K., Haase, N., Hailpern, S., Ho, P. M., Howard, V., Kissela, B., Kittner, S., Lackland, D., Lisabeth, L., Marelli, A., McDermott, M. M., Meigs, J., Mozaffarian, D., Mussolino, M., Nichol, G., Roger, V. L., Rosamond, W., Sacco, R., Sorlie, P., Stafford, R., Thom, T., Wasserthiel-Smoller, S., Wong, N. D., & Wylie-Rosett, J. (2010). Heart disease and stroke statistics—2010 update: a report from the American Heart Association. *Circulation*, 121, 46–215.

The Massachusetts Medical Society. (1995). Tissue plasminogen activator for acute ischemic stroke. *The New England Journal of Medicine*, 333(24), 1681-1587.

Tong, D., Reeves, M. J., Hernandez, A. F., Zhao, X., Olson, D. M., Fonarow, G. C., Schwamm, L. H., & Smith, E. E. (2012). Times from symptom onset to hospital arrival in the get with the guidelines – stroke program 2002 to 2009. *Stroke,* 43, 1912-1917.

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika,* 70(1), 41-55.

Vincent, H. K., Vincent, K. R., & Lamb, K. M. (2010). Obesity and mobility disability in the older adult. *World Obesity*, 11(8), 553-626.

Virani, S. S., Alonso, A., Benjamin, E. J., Bittencourt, M. S., Callaway, C. W., Carson, A. P., Chamberlain, A. M., Chang, A. R., Cheng, S., Delling, F. N., Djousse, L., Elkind, M. S.V., Ferguson, J. F., Fornage, M., Khan, S. S., Kissela, B. M., Knutson, K. L., Kwan, T. W., Lackland, D. T., Lewis, T. T., Lichtman, J. H., Longenecker, C. T., Loop, M. S., Lutsey, P. L., Martin, S. S., Matsushita, K., Moran, A. E., Mussolino, M. E., Perak, A. M., Rosamond, W. D., Roth, G. A., Sampson, U. K.A., Satou, G. M., Schroeder, E. B., Shah, S. H., Shay, C. M., Spartano, N. L., Stokes, A., Tirschwell, D. L., VanWagner, L B., & Tsao, C. W. (2020). Heart disease and stroke statistics – 2020 update. *Circulation*, 141, 139-596.

Waite, L. J., Laumann, E. O., Levinson, W. S., Lindau, S. T., & O'Muircheartaigh, C. A. (2019). National Social Life, Health, and Aging Project: Wave 1, [United States], 2005-2006. *Inter-university Consortium for Political and Social Research*. https://doi.org/10.3886/ICPSR20541.v9

World Health Organization. (1995). Physical status: the use and interpretation of anthropometry. *World Health Organization.*

Yang, Q., Tong, X., Schieb, L., Vayghan, A., Gillespie, C., Wiltz, J. L., King, S. C., Odom, E., Merritt, R., Hong, Y., & George, M. G. (2017). Vital signs: Recent trends in stroke death rates – United States, 2000-2015. *Morbidity and Mortality Weekly Report*, 66(35), 933-939.

APPENDIX. SAS PROGRAMMING CODE

```
*library;
libname nshap 'C:\Users\izumi.endo\Downloads\Paper';

ods rtf title = 'Academic Paper';

*select observations which have a history of stroke;
data additional;
  set nshap.wave1;
  where CONDITNS_5=1;
  run;

*count the number of observations which have a hisotry of stroke;
proc freq data=additional;
  tables CONDITNS_5;
  run;

*select only black and white subjects, and clean the dataset;
data cleaning(rename=(CONDITNS_1=ARTHRITIS CONDITNS_2=ULCERS CONDITNS_3=COPD
CONDITNS_4=ASTHMA
                       CONDITNS_5=STROKE CONDITNS_6=HYPERTENSION
CONDITNS_7=DIABETES));
  set nshap.wave1(keep=CONDITNS_5 ETHGRP AGE EDUC MARITLST
                        INSURE_1 INSURE_2 INSURE_3 INSURE_4 INSURE_5
HEARN_RECODE
                        HSASSETS_RECODE BMI HRTPROB HRTFAIL CONDITNS_1
                        CONDITNS_2 CONDITNS_3 CONDITNS_4
                        CONDITNS_6 CONDITNS_7 WALKBLK WALKROOM DRESSING
                        BATHING EATING INOUTBED TOILET HEALTH);
  where Ethgrp=1 or Ethgrp=2;
  if Ethgrp=1 then RACE=0;
    else if Ethgrp=2 then RACE=1;
  if INSURE_1=1 or INSURE_2=1 or INSURE_3=1 or INSURE_4=1 or INSURE_5=1 then
INSURANCE=1;
    else if INSURE_1=0 and INSURE_2=0 and INSURE_3=0 and INSURE_4=0 and
INSURE_5=0 then INSURANCE=0;
    else if INSURE_1 in (-5 -4 -1) and INSURE_2 in (-5 -4 -1) and INSURE_3 in
(-5 -4 -1)
          and INSURE_4 in (-5 -4 -1) and INSURE_5 in (-5 -4 -1) then
INSURANCE=.;
  if EDUC=1 then HIGH_SCL=0;
    else if EDUC in (2 3 4) then HIGH_SCL=1;
  if MARITLST=1 then MARRIED=1;
    else if MARITLST in (2 3 4 5 6) then MARRIED=0;
  if HEARN_RECODE in (-8 -3 -2 -1) then INCOME=.;
    else if HEARN_RECODE in (1) then INCOME=0;
      else if HEARN_RECODE in (2 3 4) then INCOME=1;
  if HSASSETS_RECODE in (-8 -3 -2 -1) then ASSET=.;
    else if HSASSETS_RECODE in (1 2 3) then ASSET=0;
      else if HSASSETS_RECODE in (4 5) then ASSET=1;
  if BMI=-3 then BMI_R=.;
    else BMI_R=round(BMI,.1);
  if HRTPROB=-2 then HRTATK=.;
    else if HRTPROB=0 then HRTATK=0;
    else if HRTPROB=1 then HRTATK=1;
```

```
  if HRTFAIL=-2 then HRTFAILURE=.;
    else if HRTFAIL=0 then HRTFAILURE=0;
    else if HRTFAIL=1 then HRTFAILURE=1;
  if WALKBLK=-8 then BLKWALK=.;
    else if WALKBLK=0 then BLKWALK=0;
    else if WALKBLK in (1 2 3 4) then BLKWALK=1;
  if WALKROOM=-8 then ROOMWALK=.;
    else if WALKROOM=0 then ROOMWALK=0;
    else if WALKROOM in (1 2 3 4) then ROOMWALK=1;
  if DRESSING in (-8 -2) then DRESS=.;
    else if DRESSING=0 then DRESS=0;
    else if DRESSING in (1 2 3 4) then DRESS=1;
  if BATHING in (-8 -2) then BATH=.;
    else if BATHING=0 then BATH=0;
    else if BATHING in (1 2 3 4) then BATH=1;
  if EATING in (-8 -2) then EAT=.;
    else if EATING=0 then EAT=0;
    else if EATING in (1 2 3 4) then EAT=1;
  if INOUTBED in (-8 -2) then BED=.;
    else if INOUTBED=0 then BED=0;
    else if INOUTBED in (1 2 3 4) then BED=1;
  if TOILET in (-8 -2) then BATHRM=.;
    else if TOILET=0 then BATHRM=0;
    else if TOILET in (1 2 3 4) then BATHRM=1;
  if HEALTH in (-5 -4 -2) then PEER=.;
    else if HEALTH in (3 4 5) then PEER=0;
    else if HEALTH in (1 2 3 4) then PEER=1;
      ID=_N_;
      RANDOM=uniform(0);
  run;

*another step of cleaning. Selecting needed variables only;
data bw;
  set cleaning(Keep=ID RANDOM AGE RACE HIGH_SCL INSURANCE MARRIED INCOME
                    ASSET BMI_R HRTATK HRTFAILURE BLKWALK ROOMWALK DRESS
                    BATH EAT BED BATHRM PEER ARTHRITIS ULCERS COPD ASTHMA
                    STROKE HYPERTENSION DIABETES);
  run;

*subsetting the data of black and white stroke survivors;
data s_bw;
  set bw;
  if STROKE=1;
  run;

*Sorting the dataset by race;
proc sort data=s_bw;
  by RACE;
  run;

*t test of Age and BMI by race;
proc ttest data=s_bw sides=2 alpha=0.05 h0=0;;
  class RACE;
  var AGE BMI_R;
  run;

*Chi square test different characteristics and measures;
```

```
proc freq data=s_bw;
  tables RACE*INSURANCE RACE*HIGH_SCL RACE*MARRIED RACE*INCOME RACE*ASSET
         RACE*HRTATK RACE*HRTFAILURE RACE*ARTHRITIS RACE*ULCERS
         RACE*COPD RACE*ASTHMA RACE*HYPERTENSION RACE*DIABETES
         RACE*BLKWALK RACE*ROOMWALK RACE*DRESS RACE*BATH
         RACE*EAT RACE*BED RACE*BATHRM RACE*PEER / chisq expected;
  run;

*Create study and control datasets;
data study control;
  set bw;
  if STROKE=1 then output study;
    else output control;
  run;

*for each case, getting all control subjects that match case's race and age;
proc sql;
  create table control1
   as select
      ONE.ID as STUDY_ID,
      TWO.ID as CONTROL_ID,
      ONE.AGE as STUDY_AGE,
      TWO.AGE as CONTROL_AGE,
      ONE.RACE as STUDY_RACE,
      TWO.RACE as CONTROL_RACE,
      ONE.RANDOM as RANDOM
  from study one, control two
  where (ONE.AGE=TWO.AGE and ONE.RACE=TWO.RACE);

*delete duplicate controls;
proc sort data=control1 nodupkey;
  by CONTROL_ID RANDOM;
  run;

*sort data by study_id;
proc sort data=control1;
  by STUDY_ID;
  run;

*randomly select a control for each case;
proc surveyselect data=control1
  method=srs n=1
  seed=0 out=strata;
  strata STUDY_ID;
  run;

*create a dataset only with the variables of STUDY_ID and ID;
data control2(rename=(CONTROL_ID=ID));
  set strata(keep=STUDY_ID CONTROL_ID);
  run;

*sort the dataset (preparing to merge);
proc sort data=control2;
  by ID;
  run;

*sort the dataset (preparing to merge);
```

```
proc sort data=bw;
  by ID;
  run;

*merge the dataset of control and black and white subjects;
data control3;
  merge control2 bw;
  by ID;
  run;

*delete the observations which were not selected to be in control group;
data control4;
  set control3;
  if STUDY_ID=. then delete;
  run;

*change variable names (preparing to merge);
data control5(rename=(ID=CONTROL_ID AGE=AGE_C ARTHRITIS=ARTHRITIS_C
ARTHRITIS=ARTHRITIS_C ULCERS=ULCERS_C
                      COPD=COPD_C ASTHMA=ASTHMA_C STROKE=STROKE_C
HYPERTENSION=HYPERTENSION_C
                      DIABETES=DIABETES_C RACE=RACE_C INSURANCE=INSURANCE_C
HIGH_SCL=HIGH_SCL_C
                      MARRIED=MARRIED_C INCOME=INCOME_C ASSET=ASSET_C
BMI_R=BMI_R_C HRTATK=HRTATK_C
                      HRTFAILURE=HRTFAILURE_C BLKWALK=BLKWALK_C
ROOMWALK=ROOMWALK_C DRESS=DRESS_C
                      BATH=BATH_C EAT=EAT_C BED=BED_C BATHRM=BATHRM_C
PEER=PEER_C));
  set control4(drop=RANDOM);
  run;

*sort the dataset (preparing to merge);
proc sort data=control5;
  by STUDY_ID;
  run;

*change variable names and get rid of unecessary variable (preparing to
merge);
data prep(rename=(ID=STUDY_ID));
  set s_bw(drop=RANDOM);
  run;

*sort the dataset (preparing to merge);
proc sort data=prep;
  by STUDY_ID;
  run;

*merge the dataset;
data combined;
  merge prep control5;
  run;

*Perform paired t test on age and BMI;
proc ttest data=combined sides=2 alpha=0.05 h0=0;;
  paired AGE*AGE_C BMI_R*BMI_R_C;
  run;
```

```
*Perform McNemar's test;
proc freq data=combined;
  tables INSURANCE*INSURANCE_C HIGH_SCL*HIGH_SCL_C MARRIED*MARRIED_C
         INCOME*INCOME_C ASSET*ASSET_C HRTATK*HRTATK_C
HRTFAILURE*HRTFAILURE_C
         ARTHRITIS*ARTHRITIS_C ULCERS*ULCERS_C COPD*COPD_C ASTHMA*ASTHMA_C
         HYPERTENSION*HYPERTENSION_C DIABETES*DIABETES_C BLKWALK*BLKWALK_C
         ROOMWALK*ROOMWALK_C DRESS*DRESS_C BATH*BATH_C EAT*EAT_C
         BED*BED_C BATHRM*BATHRM_C PEER*PEER_C / agree expected;
  run;

*check correlation between BMI_R and BMI_R_C;
proc corr data=combined plots=scatter;
  var BMI_R BMI_R_C;
  run;

*create control dataset (cleaning);
data control6(rename=(CONTROL_ID=ID));
  set strata(keep=STUDY_ID CONTROL_ID);
  run;

*sort the dataset before merging;
proc sort data=control6;
  by ID;
  run;

*sort the black and white dataset before merging;
proc sort data=bw;
  by ID;
  run;

*merge the datasets;
data control7;
  merge control6 bw;
  by ID;
  run;

*delete the observations which are not selected to the control group and
remove unneeded variables;
data control8(drop=STUDY_ID RANDOM);
  set control7;
  if STUDY_ID=. then delete;
  run;

*Stack the datasets of control and stroke subjects;
data stacked;
  set s_bw control8;
  run;

*Sort the dataset;
proc sort data=stacked;
  by STROKE;
  run;

*Find mean of AGE and BMI_R;
proc means data=stacked;
```

```
  var AGE BMI_R;
  by STROKE;
  run;

*Count frequencies on the variables;
proc freq data=stacked;
  tables INSURANCE HIGH_SCL MARRIED INCOME ASSET HRTATK HRTFAILURE
         ARTHRITIS ULCERS COPD ASTHMA HYPERTENSION DIABETES BLKWALK
         ROOMWALK DRESS BATH EAT BED BATHRM PEER;
  by stroke;
  run;

*Prepare dataset;
data c;
  set stacked;
  DAB_C=BLKWALK+ROOMWALK+DRESS+BATH+EAT+BED+BATHRM;

HLTIS_C=HRTATK+HRTFAILURE+ARTHRITIS+ULCERS+COPD+ASTHMA+HYPERTENSION+DIABETES;
  run;

*Sort the dataset;
proc sort data=c;
  by STROKE;
  run;

*Prepare dataset for regression;
data reg(rename=(DAB_C=DISABILITY));
  set c;
  run;

*drop unecessary variables for regression;
data reg2(keep=AGE ARTHRITIS ULCERS COPD ASTHMA STROKE HYPERTENSION DIABETES
RACE
               HIGH_SCL MARRIED BMI_R HRTATK HRTFAILURE DISABILITY);
  set reg;
  run;

*create control dataset to insert pair ID;
data reg3;
  set reg2;
  where stroke=0;
  PAIR=_N_;
  run;

*create stroke dataset to insert pair ID;;
data reg4;
  set reg2;
  where stroke=1;
  PAIR=_N_;
  run;

*stack control and stroke datasets;
data reg5;
  set reg3 reg4;
  run;

*regression;
```

```
proc glimmix data=reg5 method=quad;
  class pair;
  model DISABILITY = AGE ARTHRITIS ULCERS COPD ASTHMA
                     STROKE HYPERTENSION DIABETES RACE HIGH_SCL MARRIED
                     HRTATK HRTFAILURE BMI_R/ link=log s dist=poisson cl;
  random intercept / subject=pair;
  run;

*check VIF;
proc reg data=reg5;
  model DISABILITY = AGE ARTHRITIS ULCERS COPD ASTHMA
                     STROKE HYPERTENSION DIABETES RACE HIGH_SCL MARRIED
                     HRTATK HRTFAILURE BMI_R
            / tol vif collin;
  run;


ods rtf close;
```