Impact of Vascular Risk Factors on Parkinson's Disease

Yue Cao, Shulin Qing, Gege Gui

1 Introduction

The goal of the study is to determine the impact of vascular risk factors on the development of Parkinson's disease (PD) using the National Institute of Neurological Disease (NINDS) data. The core assessments are motor impairment (MDS-UPDRS) and cognitive functioning (MoCA). Normal MoCA score is anything greater than or equal to 26; Mild Cognitive Impairment (MCI) is generally greater than or equal to 20; Dementia is 19 or less. The first aim is to find out the relationship between Framingham Score, a summary of modifiable vascular risk factors, and cognitive impairment; the second aim is to identify the relationship between vascular risk factors (such as hypertension, obesity or smoking) and cognitive impairment. We focus on the baseline observations in this study.

2 Data Overview

The data contain information collected from 9 sites over 5 years. There are 29 variables and 2961 observations in total. We selected 16 variables, including ID, age, PD/control, gender, race, education, ever smoke, Hamilton Depression Score, motor score, height, weight, average systolic blood pressure (SBP), average diastolic blood pressure (DBP), hypertension medication, diabetes and MoCA score. We excluded 168 atypical parkinsonism syndromes patients, 96 observations from University of Florida (Gainesville) without height and weight information, 1362 visits other than baseline and 158 uncompleted cases.

Figure 1 for MoCA score distribution of data collected from different sites showed that there was no significant site difference in the collected data.

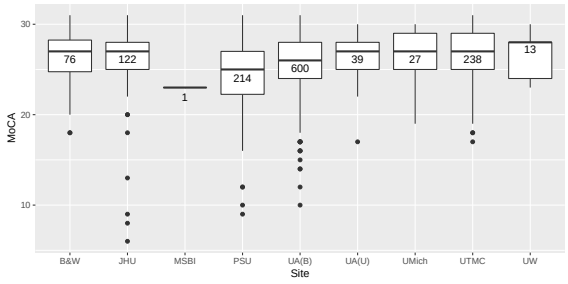

Figure 1: MoCA score collected at differnet sites. B&W: Brigham and Women's; JHU: Johns Hopkins University; PSU: Pennsylvania State University (Hershey); UA(B): University of Alabama (Birmingham); UA(U): University of Alabama (Udall); UMich: University of Michigan; UTMC: UT-Southwestern Medical Center; UW: University of Washington.

Table 1: The mean and median MoCA score in the missing and non-missing groups of each variable

(missing/non-missing)	Mean	Median	Missing Percentage
Depression Score	24.44/25.69	26/26	8.9%
Height	24.29/25.70	26/26	8.6%
SBP/DBP	26.28/25.53	27/26	5.7%
Diabetes	21.69/25.80	23/26	5.5%

To check if the data are missing at random, we summarized the MoCA score mean and median in the missing group and non-missing group for each variable that has more than 100 missing observations in Table 1. Some variables had a similar pattern of missing. All people without diabetes record (158) did not have medication indicator; most people without height (250) and weight information (235) did not have depression score (259), the overlapping group consisting of 205 individuals; all people without SBP (166) did not have DBP. The table showed that the only unbalanced variable was diabetes, with median 23 for missing group compared to 26 for non-missing group. Therefore, we may consider remove this variable from the study.

Table 2: The percentage of categorical variables with indicator = 1 in each MoCA score group

Percentage	Dementia (MoCA <20)	MCI (MoCA 20 - 25)	Normal (MoCA >25)
PD	0.902	0.700	0.543
Female	0.213	0.371	0.525
Ever smoke	0.443	0.419	0.380
Medication	0.459	0.324	0.216
Diabetes	0.115	0.060	0.051

We created another two tables for the percentage or median of each variable in three different groups, Normal MoCA score, MCI and Dementia. These groups had 61, 420 and 690 observations respectively. As more than 90% of the people were Caucasian, we did not include race in the analysis. There were 31 individuals with education 12th grade or below, so we combined this category with high school graduate, which had 161 people. The other three education groups had 290, 351 and 338 people respectively. The descriptive analysis from Table 2 for other categorical variables showed that PD percentage and gender percentage were related with MoCA score, with more PD patients in low MoCA score group and more female in high MoCA score group. The percentages of diabetes were around or less than 1% in each MoCA score group, so it might be appropriate to delete this covariate in the study considering this and the check of missingness.

Table 3: The median of continuous variables in each MoCA score group

Median	Dementia (MoCA <20)	MCI (MoCA 20 - 25)	Normal (MoCA >25)
Age	71	68.92	63.08
Height	178	173	170
Weight	83.9	81.6	79.4
BMI	26.48	27.37	27.16
Depression score	4	3	3
Motor score	31	19	9
SBP135	135	132	130
DBP	78	79	79
Framingham score	0.42	0.26	0.17

From Table 3 for continuous variables, we noticed that people with lower motor score were more likely to be in normal MoCA group, and so did people with lower Framingham score. As the calculated Framingham score was between 0 and 1, we rescaled it to be from 10 to 100 in model building.

3 Methods

We modeled the relationship between MoCA score and potential risk factors using multiple linear regression where the MoCA score was treated as a continuous outcome, and multinomial logistic regression where it was treated as a categorical outcome. The selection of covariates for our models was first determined by likelihood ratio test which evaluates how likely the data are under an alternative model than a null model. We started from the model including all covariates. We deleted one variable at each time and compared the current null model with the model from last step by likelihood ratio test. If p-value is less than 0.05,

there is evidence against the current null model in favor of the model containing one more covariate. The selected covariates in the multiple linear regression for aim 1 are Framingham score, PD/Control, education and Hamilton Depression Score. The selected covariates in the multiple linear regression for aim 2 are age, PD/Control, gender, education, Hamilton Depression Score, BMI and motor score. Secondly, collinearity between predictor variables was checked. Each pair of predictor variables was plotted against each other and a loess regression was used to fit a smooth curve through points of the scatter plot. Correlation was also calculated for the linear relationship between variables. We found that the correlation between motor score and PD/Control is relatively high so this implies that individuals with Parkinson disease may have worse motor ability and their relationship is positively linear. Thus, we excluded motor score from our regression analysis. The coefficient of the predictor BMI in the temporary multiple linear regression model is around 0 and we found there is no much difference of BMI among 3 MoCA score groups, so BMI was also removed from our regression model. We conducted model diagnosis by plotting the histogram of residuals and QQ plot. We applied same predictors to both multiple linear regression and multinomial logistic regression models, and the final regression models are:

Aim 1:

$$MoCA \sim Framingham + PD/Control + Education + Depression\ Score + \epsilon$$

a.

$$\log(\frac{P(MoCA = Dementia)}{p(MoCA = Normal)}) \sim Framingham + PD/Control + Education + Depression\ Score + \epsilon$$

b.

$$\log(\frac{P(MoCA = Mild)}{p(MoCA = Normal)}) \sim Framingham + PD/Control + Education + Depression\ Score + \epsilon$$

Aim 2:

$$MoCA \sim Age + PD/Control + Gender + Education + Depression\ Score + \epsilon$$

a.

$$\log(\frac{P(MoCA = Dementia)}{p(MoCA = Normal)}) \sim Age + PD/Control + Gender + Education + Depression\ Score + \epsilon$$

b.

$$\log(\frac{P(MoCA = Mild)}{p(MoCA = Normal)}) \sim Age + PD/Control + Gender + Education + Depression\ Score + \epsilon$$

4 Results

4.1 Aim 1

4.1.1 Multiple Linear Regression for Aim 1

To explore the impact of vascular risk factors on cognitive impairment, we first conduct a multiple linear regression to explore the relationship between MoCA score and Framingham score, adjusting for other risk factors. As shown in Table-4, Framingham score, PD/Control, most of education categories and Hamilton Depression Score are of statistically significance, with p-values less than 0.05. With the Framingham score increasing by 0.1, the MoCA score would decrease by 0.45 on average, with a 95% confidence interval of [-0.45, -0.37]. This means individuals with a higher Framingham score would have a lower MoCA score, i.e. greater

Table 4: Multiple linear regression results for the association between MoCA score and Framingham Score, adjusting for PD/Control, Education level and Hamilton Depression Score

		Estimate	95%CI	P-value
(Intercept)		26.63	[26.09, 27.17]	<0.01
Framingham score		-0.45	[-0.54, -0.37]	<0.01
PD		-0.79	[-1.15, -0.42]	<0.01
Education	Associate degree	0.48	[-0.07, 1.02]	0.09
	Bachelors	1.59	[1.06, 2.11]	<0.01
	Professional degree	1.68	[1.15, 2.21]	<0.01
Depression score		-0.11	[-0.16, -0.06]	<0.01

cognitive impairment. Compared with the control group, the MoCA score of PD patients would decrease by 0.79 on average, with a 95% confidence interval of [-1.15, -0.42]. Compared to the high school graduates or below, the more advance education level would have a higher MoCA score, which means the increasing of education level has a positive influence on cognition. With the increasing of Hamilton Depression Score by 1, the MoCA score would decrease by 0.11 on average, with a 95% confidence interval of [-0.16, -0.06].

In conclusion, the individuals with a higher Framingham score, Parkinson's Disease, lower education level, and higher Hamilton Depression Score would probably have greater cognitive impairment. Among these factors, the most influential covariate is Bachelor's degree (compared to high school graduates or below), with an estimate of 1.59, and the least influential covariate is Hamilton Depression Score, with an estimate of -0.11.

We show the multiple linear regression results in Table-4. Noted that education levels are in comparison with high school graduates or below.

We further add interaction terms into the linear regression. We use likelihood ratio test to prove that adding the interaction term of PD/Control with education level would result in a significantly different model. In this new model, the effect of education on MoCA score is different in the control group and the PD patients. At the same education level, the increase of MoCA score is generally greater in the PD patients than the control group.

4.1.2 Multinomial Logistic Regression for Aim 1

To address the difference between the association of dementia and MCI with Framingham score and other risk factors, we divided MoCA score into three categories, normal cognition, MCI and dementia. A multinomial logistic regression is conducted, with the outcome as the log odds of being dementia vs. normal cognition and the log odds of being MCI vs. normal cognition. As the exponentiated results shown in Table-5., a 0.1 increase of Framingham score is associated with the increase of the odds of being dementia in the amount of 1.55, with a 95% confidence interval of [1.37, 1.77], while the increase of the odds of being MCI is 1.27, with a confidence interval of [1.19, 1.37]. This means the effect of Framingham score is stronger in the odds of being dementia, than the odds of being MCI. Similar trends can be observed in the coefficients of PD/Control and Hamilton Depression Score. Compared to high school graduate or below, higher education level is associated with the decrease of odds of being dementia or being MCI, and the decreasing is greater in the odds of being dementia than the odds of being MCI.

We show the multinomial logistic regression results in Table-5. Noted that education levels are in comparison with high school graduates or below; the estimates and 95% confidence intervals are exponentiated from the original regression results.

Table 5: Multinomial logistic regression results for the association between cognition status and Framingham Score, adjusting for PD/Control, Education level and Hamilton Depression Score

		Dementia	95%CI	pval1	MCI	95%CI	pval2
(Intercept)		0.01	[0, 0.03]	<0.01	0.32	[0.22, 0.49]	<0.01
Framingham score		1.55	[1.37, 1.77]	<0.01	1.27	[1.19, 1.37]	<0.01
PD		5.21	[2.15, 12.69]	<0.01	1.75	[1.33, 2.31]	<0.01
Education	Associate degree	0.68	[0.33, 1.39]	0.29	1.01	[0.68, 1.5]	0.98
	Bachelors	0.18	[0.08, 0.42]	<0.01	0.51	[0.34, 0.75]	<0.01
	Professional degree	0.19	[0.08, 0.42]	<0.01	0.46	[0.31, 0.68]	<0.01
Depression score		1.11	[1.04, 1.17]	<0.01	1.01	[0.98, 1.05]	0.45

Table 6: Multiple linear regression results for the association between MoCA score and vascular risk factors

		Estimate	95%CI	P-value
(Intercept)		30.17	[28.96, 31.38]	<0.01
Age		-0.09	[-0.1, -0.07]	<0.01
PD		-0.63	[-1, -0.27]	<0.01
Female		1.27	[0.92, 1.61]	<0.01
Education	Associate degree	0.46	[-0.07, 0.99]	0.09
	Bachelors	1.74	[1.23, 2.26]	<0.01
	Professional degree	1.95	[1.43, 2.46]	<0.01
Depression score		-0.11	[-0.16, -0.07]	<0.01

4.2 Aim 2

4.2.1 Multiple Linear Regression for Aim 2

To investigate the relationship between vascular risk factors (hypertension, obesity or smoking) and cognitive functioning evaluated by MoCA score, we first conducted multiple linear regression. The p-values for coefficients of hypertension and smoking are large (hypertension: p-value > 0.1, smoking: p-value > 0.5, diabetes: p-value > 0.5) so it implies that none of these vascular risk factors have significant association with cognitive impairment. The multiple linear regression is then adjusted for other risk factors. The results are shown in Table-6.. With every year increase in age, the MoCA score would decrease by 0.09 on average, with a 95% confidence interval of [-1.0, -0.07]. This means that as individuals get older, they are more likely to have cognitive impairment. The MoCA score for individuals with Parkinson disease is 0.63 units lower than the MoCA score for individuals in the control group, with a 95% confidence interval [-1, -0.27]. This means that people with Parkinson disease are more likely to have cognitive impairment. Females would have MoCA scores 1.27 units higher than males on average, with a 95% confidence interval [0.92,1.61]. This means that females are less likely to have cognitive impairment than males. Individuals with higher education tend to have higher MoCA score. For example, individuals with professional degree would have MoCA score 1.95 units higher than those with high school degree or below, with a 95% confidence interval [1.43,2.46]. This means that people receiving higher education are more likely to have well-functioning cognition. With the increase of Hamilton Depression Score by 1 unit, the MoCA score would decrease by 0.11, with a 95% confidence interval [-0.16, -0.07]. This means more depressed people would have lower MoCA score and therefore worse cognition functioning.

We show the multiple linear regression results in Table-6. Noted that education levels are in comparison with high school graduates or below.

Table 7: Multinomial logistic regression results for the association between cognition status and vascular risk factors

		Dementia	95%CI	pval1	MCI	95%CI	pval2
(Intercept)		0	[0, 0]	<0.01	0.02	[0.01, 0.06]	<0.01
Age		1.14	[1.09, 1.18]	<0.01	1.06	[1.04, 1.08]	<0.01
PD		4.95	[2.02, 12.21]	<0.01	1.68	[1.26, 2.22]	<0.01
Female		0.21	[0.1, 0.42]	<0.01	0.52	[0.39, 0.68]	<0.01
Education	Associate degree	0.76	[0.36, 1.6]	0.46	1.03	[0.69, 1.56]	0.87
	Bachelors	0.15	[0.06, 0.36]	<0.01	0.45	[0.3, 0.67]	<0.01
	Professional degree	0.13	[0.06, 0.31]	<0.01	0.38	[0.25, 0.57]	<0.01
Depression score		1.12	[1.05, 1.19]	<0.01	1.02	[0.99, 1.06]	0.25

4.2.2 Multinomial Logistic Regression for Aim 2

To get a closer look at the association of cognitive impairment with the risk factors, we treated the MoCA score as a categorical outcome with 3 categories as mentioned above. The multinomial logistic regression helps to compare the categories pairwise. As the exponentiated results shown in Table-7., one year increase in age is associated with the increase of the odds of being dementia vs. normal cognition in the amount of 1.14, with a 95% confidence interval [1.09,1.18], while the increase of the odds of being MCI vs. normal cognition is 1.06, with a 95% confidence interval [1.04, 1.08]. This means that the risk factor, age, has a greater impact on the odds of being dementia than the odds of being MCI. The relative risk for females getting dementia vs. normal cognition is 0.21 times the relative risk for males getting dementia vs. normal cognition, with a 95% confidence interval [0.1, 0.42], while the relative risk ratio comparing females with males is 0.52 for getting MCI vs. normal cognition. Compared to the relative risk of getting MCI, the relative risk for females getting dementia is much lower. Similar trends observed in the multiple linear regression are also showed here. Individuals with higher education degree are less likely to have dementia or MCI and more depressed individuals are more likely to have dementia or MCI.

In conclusion, the individuals who are females and older, with Parkinson's disease, lower education level, and higher Hamilton Depression Score would probably have greater cognitive impairment.

We show the multinomial logistic regression results in Table-7. Noted that education levels are in comparison with high school graduates or below; the estimates and 95% confidence intervals are exponentiated from the original regression results.

5 Discussion

The results of our study are generally in accordance with the current understandings of Parkinson's disease. However, as our study is based on the cross-sectional data at baseline visits, the major concern is the insufficient use of data and possible misleading results of the regression models. For example, in the multinomial logistic regression, the transition from normal cognition to MCI and the transition from MCI to dementia could be studied if we use longitudinal data for each subject and make the odds of transition as the outcome in the regression model. In addition, the influence of BMI on the cognition impairment is not clear since we excluded BMI in the analysis for Aim 2, while BMI is in fact a very important factor to evaluate the health condition of patients in clinical practice.