Final Project

Zhuodiao Kuang(zk2275)

Prompt Two – Sex and Genetic Risk Factors for Cognitive Impairment in Late Middle-Aged Adults

Statistical Analysis: (214 words)

A meticulous process was used to analyze cognitive impairment data, specifically addressing missing values and ensuring the accuracy of statistical models. Variables used in the analysis are all presented and explained in **Table 1**. Due to the categorical character of important variables such as APOE-ε4 carrier statuses and sex, nonparametric procedures like the Wilcoxon Two-sample test and the Kruskal-Wallis test were mainly employed to compare distributions among different groups. Before conducting ANOVA tests, the data was assessed for normality and equality of variances to confirm the assumptions of parametric testing. The Chi-squared test was employed to examine the correlation between diabetes and Aβ positive, offering insights into potential comorbidities that may impact cognitive health. The significance levels were set at $p < 0.05$ for all statistical tests, indicating that results with a probability of occurring by chance less than 5% were considered statistically significant (See **Table2**). In addition, a generalized linear model (GLM) was created to investigate the correlation between sex, genetic risk factors, and the presence of Alzheimer's disease (AD) neuropathology. This model systematically integrated specific covariates to account for potential confounding factors revealed through a comprehensive examination. The statistical analyses were performed using SAS Studio, a web-based interface allowing quick data management and execution of complex models.

Results (250 words):

The analysis revealed significant variations in mean Aβ levels by APOE-ε4 genotype statuses (F-value = 14.01, $p < 0.001$, **Figure 1**). Post-hoc comparisons using Dunnett's Test indicated that individuals with two copies of the ε4 allele had significantly higher mean Aβ levels than those with one or no copies ($p < 0.05$). Besides, significant differences were found between sexes in mean Aβ levels (F-value = 42.85, $p < 0.001$).

The chi-squared analysis supported no association between diabetes and Aβ positivity (p-value = 0.2888). There is no strong evidence to support that, depending on APOE-ε4 genotype carrier statuses, the cortical thickness is significantly different (p-value =0.6162). However, the one-way ANOVA test provided strong evidence to support that, depending on sexes, cortical thickness is significantly different (p-value = 0.0012). Additionally, interaction effects between genotype and sex were not statistically significant (F-value = 0.24 p = 0.7879).

Lastly, GLM models were built to evaluate risk factors for vascular dementia. Two outcomes, mean Aβ levels and cortical thickness, are chosen for creating different models. A correlation matrix was made to eliminate multicollinearity between covariates first. When mean Aβ levels were the response variable,

APOE-ε4 genotype carrier status and sex were considered as main covariates; when the cortical thickness was treated as the response, age could be the confounder, sexes still have a significant impact with several covariates we are interested in(**Figure 2**). The use of log transformation could potentially lead to more accurate and robust modeling results, and both models fit well with the data.

Discussion (90 words):

The analysis underscores the profound impact of genetic factors, particularly APOE-ε4 alleles, on cognitive functions, aligning with the hypothesis that genetic predispositions significantly contribute to AD neuropathology risks. The study's methodology, particularly the use of GLM to account for confounders, provides a robust framework for examining these relationships. However, the results are constrained by the inherent limitations of observational data and the selected cohort's demographic characteristics. Future research should expand to longitudinal studies and include a more comprehensive range of genetic and environmental factors to enhance understanding of AD pathogenesis.

Table 1. Variables explored for analysis of sex and genetic risk factors for cognitive impairment in late middle-aged adults.

Variable	Description	Levels/Summary
id	Participant ID	Total of 266 participants
sex	Sex of the participant	1 = Male, 2 = Female
IPAQtotal	Total physical activity score from the IPAQ (International Physical Activity Questionnaire), measured in MET-minutes/week	7777 = Unknown, 8888 = Refusal, 9999 = Not Applicable
IPAQcat	IPAQ categorical score	1 = Low; 2 = Moderate; 3 = High; 7777 = Unknown; 8888 = Refusal; 9999 = Not Applicable
phq9	PHQ-9 depression assessment	Numeric scale: 0 = no symptoms, 1 = minimal symptoms, 2 = mild symptoms, 3 = moderate symptoms, 4 = severe symptoms
APOEgeno	APOE-ε4 carrier statuses	0 = no copy of ε4; 1 = single copy of ε4; 2 = double copies of ε4
neuro-degeneration	Cortical thickness in Alzheimer's disease signature regions: average of several measures related to cortical thickness	Min=2.40 median=2.71 mean=2.71 max=2.91 std=0.09
global_mean_ab	Global mean amyloid β (Aβ) SUVR: average of several measures related to Aβ SUVR from different parts	Min=1.00 median=1.11 mean=1.10 max=1.35 std=0.05
education	Years of education completed by the participant	Min=0.00 median=11.00 mean=10.44 max=19.00 std=3.73
age	Age of the participant in years	Min=56.26 median=64.00 mean=64.00 max=70.56 std=3.43
waist	Waist circumference (cm)	Min=69.40 median=96.40 mean=96.05 max=134.70 std=10.75

Table 2. Results of most tests in the project.

Test	Description	P-value
ANOVA	Variations in mean Aβ levels by APOE-ε4 genotype statuses	F-value = 14.01, p < 0.001
Dunnett's Test	Individuals with two copies of the ε4 allele had significantly higher mean Aβ levels than those with one or no copies	P-value < 0.05
Wilcoxon Two-Sample Test and KRUSKAL-WALLIS TEST	Significant differences were found between sexes in mean Aβ levels	F-value = 42.85, p < 0.001
The chi-squared analysis	No association between diabetes and Aβ positivity	P-value = 0.2888
ANOVA	There is no strong evidence to support that, depending on APOE-ε4 genotype carrier statuses, the cortical thickness is significantly different	P-value =0.6162
F-test	Interaction effects between genotype and sex were not statistically significant	F-value = 0.24 p = 0.7879
The one-way ANOVA test	Depending on sexes, cortical thickness is significantly different	P-value = 0.0012

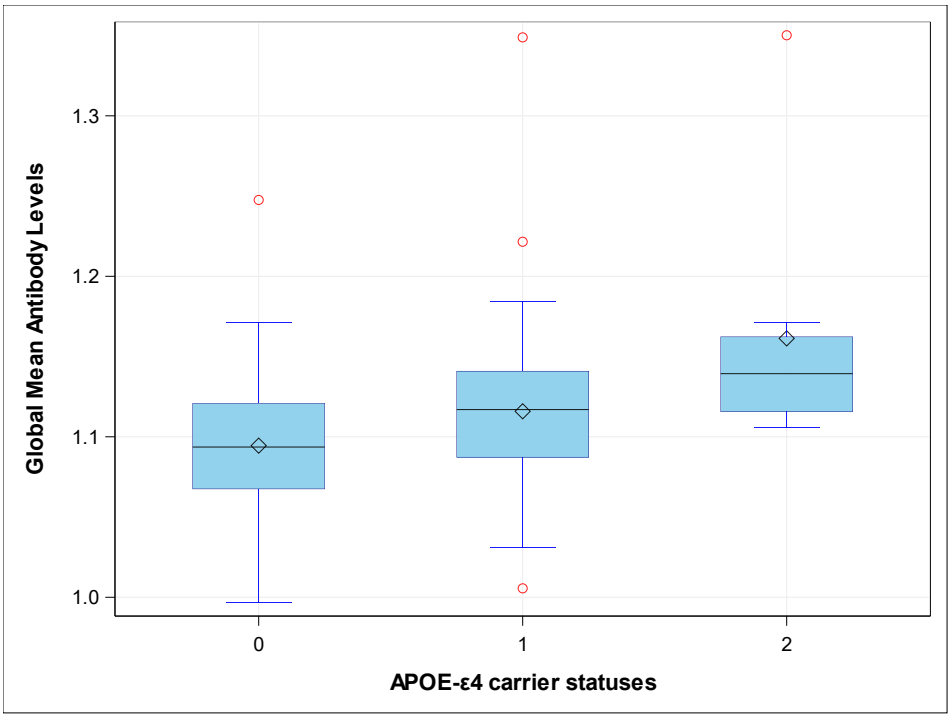

Figure1. Box Plot of Global Mean Amyloid Beta(β) Levels by APOE-ε4 Carrier Statuses

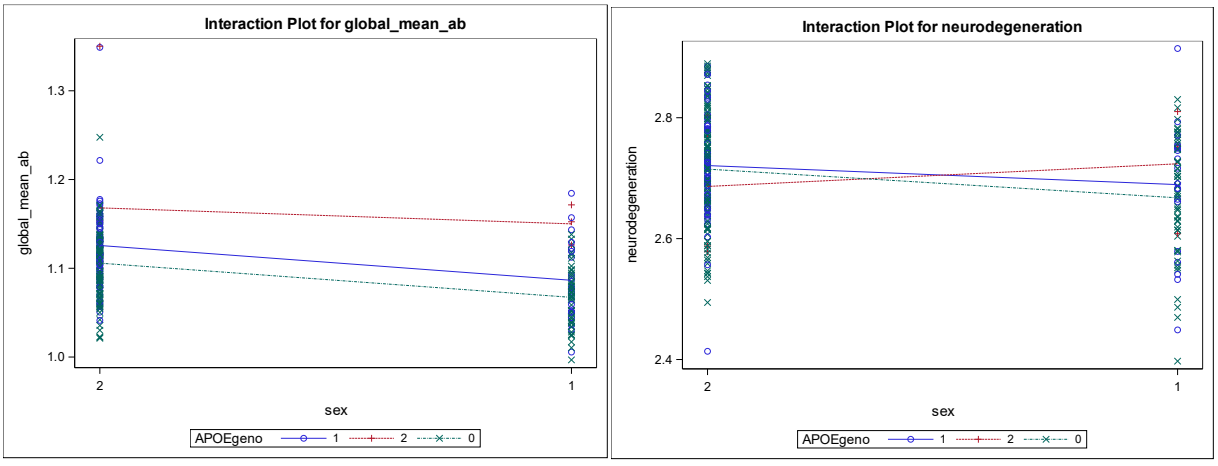

Figure2. Interaction Plots for Global_mean_ab and Neurodegeneration