

# Project 0 final report

Shuai Zhu

2024-10-01

---

## 1 Introduction

The data for this project is on a set of resting state functional connectivity(RSFC) and a set of measures of physical function. The investigator collected 30 resting state functional connectivity and 15 physical functions of 300 participants. The objective of this report is to outline an analysis plan aimed at addressing the following research question:

- what measures of resting state functional connectivity are associated with individual physical function.

To achieve this, we employ a series of separate linear regression models, one for each measure of physical function.

## **2 Method**

### **2.1 Data collection and Data cleaning**

The study collected data from 300 participants. For each participant, two sets of data were obtained. 30 measures of resting-state functional connectivity (RSFC): These measures, labeled as X.1 to X.30, were used as predictors in the analysis. 15 measures of physical function: These measures, labeled Y.1 to Y.15, were used as outcomes in the study. No formal data cleaning procedures were applied to the dataset. All available data from the 300 participants were included in the analysis without the removal of outliers,

### **2.2 Data analysis**

A total of 15 multivariable linear regression models were conducted to examine the associations between 30 resting-state functional connectivity predictors and 15 physical function outcomes, resulting in 450 tests. 480 parameters were estimated across all models, but the analysis focused on 450 parameters, excluding the intercept.

#### **2.2.1 Multiple Testing and Family-Wise Error Rate (FWER)**

Given the large number of tests conducted, controlling for the family-wise Type I error rate (FWER) was essential to reduce the risk of false discoveries. The FWER represents the probability of making at least one false positive across all 450 tests. To address this, the Bonferroni correction was applied. By dividing the desired significance level (0.05) by the total number of tests (450), the adjusted threshold for significance was set to a p-value of less than 0.000111. Any p-value below this threshold was considered statistically significant.

While the Bonferroni correction is straightforward to implement, it is known to be overly conservative, especially when handling a large number of tests, as it reduces statistical power by inflating

the threshold for significance. Therefore, it may lead to the exclusion of predictors that are truly significant.

### **2.2.2 Alternative Correction Approach: Bootstrap-Based Adjustment**

An alternative bootstrap-based approach was employed to address the limitations of the Bonferroni correction. This method leverages resampling to estimate the empirical distribution of the regression coefficients. Specifically, bootstrap sampling was used to generate distributions for the estimated beta coefficients, which were then used to compute 95% confidence intervals. The bootstrap method involved repeatedly resampling the data with replacement and refitting the models to generate a distribution of coefficient estimates for each predictor.

Predictors with confidence intervals that did not include zero were considered statistically significant. Unlike the Bonferroni method, the bootstrap-based approach does not rely on p-values and provides a measure of significance by directly assessing the variability of the coefficients. The results of this method, including the significant predictors and their corresponding confidence intervals, are presented in Table 1.

All analyses were conducted using R version 4.3.3.

## **3 Results**

Fifteen multivariable regression models were fitted to identify associations between 30 resting-state functional connectivity predictors and each of the 15 physical function outcomes. The significance of each predictor across the models is presented in Table 2 (Appendix). Without adjustment, 79 predictors were significantly associated with the physical function outcomes. Among these, predictor X.5 was significant for 9 outcomes.

Applying the Bonferroni correction reduced the number of significant predictors to 6 (Table 3). As mentioned in the Methods section, the Bonferroni correction was overly conservative. An alter-

native bootstrap-based approach increased the number of significant predictors to 44 (Appendix Table 3). P-values were not calculated with this method; instead, a 95% confidence interval that did not include 0 was considered significant.

## 4 Conclusions

Based on the results of the bootstrap-based adjustment, predictor X.5 was found to be significantly associated with most physical function outcomes, suggesting a potential relationship between this resting-state functional connectivity measure and physical function. For example, the relationship between X.5 and Y.1 can be interpreted as follows: for each one-unit increase in resting-state functional connectivity (X.5), the physical function Y.1 increases by 1.31 units, with a 95% confidence interval ranging from 0.0657 to 2.55. This indicates that the effect of X.5 on Y.1 is likely positive, and the association is statistically significant, as the confidence interval does not include zero. One advantage of the bootstrap method is its ability to control the family-wise error rate (FWER) more effectively compared to the Bonferroni correction. This approach is less conservative than Bonferroni's, allowing for greater statistical power while still controlling the FWER.

Table 1: Significant Predictors after Adjustment of bootstrap

Outcome	Predicator	Estimate	Lower confident interval	Upper confident interval
1	5	1.307	0.066	2.549
2	1	-1.017	-2.023	-0.012
2	5	1.823	0.565	3.081
2	17	-1.714	-3.384	-0.043
2	23	0.947	0.439	1.456
3	26	0.829	0.059	1.599
4	5	1.825	0.555	3.094

Outcome	Predicator	Estimate	Lower confident interval	Upper confident interval
4	29	-1.600	-2.878	-0.322
5	7	-0.934	-1.823	-0.045
5	23	0.827	0.290	1.364
6	5	1.405	0.104	2.706
6	7	-0.882	-1.753	-0.011
6	20	-2.069	-3.928	-0.211
6	29	-1.756	-3.065	-0.447
7	15	1.772	0.237	3.307
7	23	0.837	0.314	1.360
7	27	-1.805	-3.505	-0.104
7	29	-1.686	-2.988	-0.384
8	4	-1.253	-2.330	-0.176
8	5	1.654	0.348	2.960
8	17	-1.796	-3.530	-0.062
8	23	1.157	0.629	1.684
8	26	0.997	0.215	1.780
8	27	-1.953	-3.669	-0.236
9	7	-1.016	-1.862	-0.171
9	23	0.609	0.098	1.119
10	5	1.407	0.143	2.671
10	7	-0.965	-1.811	-0.119
10	26	1.082	0.325	1.839
11	4	-1.253	-2.318	-0.188
11	26	1.410	0.636	2.184
11	27	-1.821	-3.519	-0.123

Outcome	Predicator	Estimate	Lower confident interval	Upper confident interval
11	29	-1.561	-2.860	-0.261
12	4	-1.168	-2.278	-0.059
12	5	1.434	0.089	2.779
12	15	1.686	0.090	3.282
12	23	0.680	0.137	1.224
13	5	1.592	0.228	2.956
13	15	1.830	0.211	3.449
13	22	1.424	0.054	2.794
14	23	1.004	0.452	1.556
15	1	-1.129	-2.193	-0.066
15	22	1.395	0.059	2.731
15	29	-1.663	-3.001	-0.324

## 5 Appendix

Table 2: Significant Predictors without adjustment

Outcome	Predictor	Estimate	Lower confident interval	Upper confident interval
Y.1	X.1	-0.867	-1.612	-1.612
Y.1	X.5	1.307	0.375	0.375
Y.1	X.9	-1.739	-3.334	-3.334
Y.1	X.15	1.284	0.179	0.179
Y.1	X.17	-1.334	-2.571	-2.571
Y.2	X.1	-1.017	-1.772	-1.772
Y.2	X.5	1.823	0.879	0.879
Y.2	X.7	-0.732	-1.364	-1.364
Y.2	X.17	-1.714	-2.967	-2.967
Y.2	X.23	0.947	0.566	0.566
Y.3	X.1	-0.890	-1.662	-1.662
Y.3	X.3	2.427	0.517	0.517
Y.3	X.9	-2.033	-3.684	-3.684
Y.3	X.20	-1.615	-2.992	-2.992
Y.3	X.26	0.829	0.251	0.251
Y.4	X.5	1.825	0.872	0.872
Y.4	X.21	2.729	0.091	0.091
Y.4	X.29	-1.600	-2.559	-2.559
Y.5	X.7	-0.934	-1.601	-1.601
Y.5	X.15	1.479	0.296	0.296
Y.5	X.20	-1.690	-3.114	-3.114
Y.5	X.23	0.827	0.424	0.424

Outcome	Predictor	Estimate	Lower confident interval	Upper confident interval
Y.5	X.28	2.894	0.101	0.101
Y.6	X.5	1.405	0.428	0.428
Y.6	X.7	-0.882	-1.536	-1.536
Y.6	X.9	-1.733	-3.405	-3.405
Y.6	X.19	3.338	0.139	0.139
Y.6	X.20	-2.069	-3.464	-3.464
Y.6	X.21	2.781	0.078	0.078
Y.6	X.23	0.464	0.069	0.069
Y.6	X.28	2.810	0.075	0.075
Y.6	X.29	-1.756	-2.738	-2.738
Y.7	X.1	-0.783	-1.559	-1.559
Y.7	X.5	0.994	0.023	0.023
Y.7	X.12	-1.229	-2.435	-2.435
Y.7	X.13	-2.461	-4.888	-4.888
Y.7	X.15	1.772	0.620	0.620
Y.7	X.18	-2.099	-4.123	-4.123
Y.7	X.23	0.837	0.444	0.444
Y.7	X.26	0.706	0.124	0.124
Y.7	X.27	-1.805	-3.081	-3.081
Y.7	X.29	-1.686	-2.663	-2.663
Y.8	X.4	-1.253	-2.062	-2.062
Y.8	X.5	1.654	0.674	0.674
Y.8	X.17	-1.796	-3.098	-3.098
Y.8	X.20	-1.565	-2.965	-2.965
Y.8	X.23	1.157	0.760	0.760



Outcome	Predictor	Estimate	Lower confident interval	Upper confident interval
Y.8	X.26	0.997	0.410	0.410
Y.8	X.27	-1.953	-3.241	-3.241
Y.9	X.7	-1.016	-1.651	-1.651
Y.9	X.23	0.609	0.226	0.226
Y.10	X.1	-0.876	-1.634	-1.634
Y.10	X.5	1.407	0.459	0.459
Y.10	X.7	-0.965	-1.600	-1.600
Y.10	X.9	-1.945	-3.569	-3.569
Y.10	X.26	1.082	0.514	0.514
Y.11	X.3	2.276	0.357	0.357
Y.11	X.4	-1.253	-2.053	-2.053
Y.11	X.17	-1.508	-2.795	-2.795
Y.11	X.26	1.410	0.829	0.829
Y.11	X.27	-1.821	-3.095	-3.095
Y.11	X.29	-1.561	-2.536	-2.536
Y.12	X.4	-1.168	-2.001	-2.001
Y.12	X.5	1.434	0.424	0.424
Y.12	X.15	1.686	0.488	0.488
Y.12	X.22	1.283	0.269	0.269
Y.12	X.23	0.680	0.272	0.272
Y.12	X.26	0.731	0.126	0.126
Y.13	X.3	2.087	0.060	0.060
Y.13	X.5	1.592	0.568	0.568
Y.13	X.15	1.830	0.615	0.615
Y.13	X.22	1.424	0.396	0.396

Outcome	Predictor	Estimate	Lower confident interval	Upper confident interval
Y.14	X.23	1.004	0.590	0.590
Y.15	X.1	-1.129	-1.928	-1.928
Y.15	X.18	-2.280	-4.361	-4.361
Y.15	X.19	3.788	0.516	0.516
Y.15	X.22	1.395	0.392	0.392
Y.15	X.25	-2.367	-4.347	-4.347
Y.15	X.29	-1.663	-2.667	-2.667

Table 3: Significant Predictors with Bonferroni correlation

Outcome	Predictor	Estimate	Lower confident interval	Upper confident interval
Y.2	X.23	0.947	0.566	0.566
Y.5	X.23	0.827	0.424	0.424
Y.7	X.23	0.837	0.444	0.444
Y.8	X.23	1.157	0.760	0.760
Y.11	X.26	1.410	0.829	0.829
Y.14	X.23	1.004	0.590	0.590