Bayesian Analysis
of TICS Study

by Irene Hsueh

Introduction

The Telephone Interview for Cognitive (TICS) study is a standardized test measuring cognitive functioning in adults between 60-98 years of age. TICS was designed to be administered over the phone, which allows researchers to be able to examine those with visual impairments or people who can't read or write. Since information is simply collected over the phone, TICS is useful for studies with large samples or where in-person clinic follow-up is inconvenient or expensive. More information about TICS can be found on its product page.

Boston University recruited offspring of centenarians and controls and collected 5 longitudinal measures of TICS scores as part of their study on longevity and healthy aging. The purpose of this study is to use Bayesian modeling to determine whether there are significant differences in cognitive functioning between offspring of centenarians and the rest of the population.

Methods

Study Sample and Measurements

The main outcomes of interest were the five longitudinal TICS score measurements, ranging from 0 to 30 with lower scores corresponding to more severe cognitive impairment. Chi-squared and t-tests were used to determine if there were significant differences in baseline characteristics between the two groups.

The given dataset had 762 subjects, 40 of which were omitted due to having missing covariate data. Missing data was ignored because it was assumed to be MCAR or MAR. The remaining 722 subjects, 473 centenarian offspring and 247 controls, were used for subsequent analysis.

Bayesian Statistical Analysis (n=722)

Visual inspection of the 5 TICS score showed that they were approximately normally distributed. Therefore, normal prior distributions with gamma distributed standard deviations were chosen to model TICS score. All other predicted covariates were assumed to be normally distributed with a standard deviation of 0.0001.

Treating the first TICS measurement as the baseline TICS score, a crude linear regression was run between baseline score and subject group. Then, adjusted linear regressions with only one baseline covariate were used to detect confounders of the relationship between subject group and baseline TICS score. A final multivariate adjusted linear regression that included all identified confounders was run.

The graph of the first 50 subjects' change in TICS scores over the five longitudinal measurements (Figure 1) showed wildly varying intercepts and slopes, so a hierarchical model with random intercepts and slopes for time was chosen to represent the data. The interaction between subject group and time indicated whether there were different rates of changes in TICS scores over time between subject groups. Hierarchical models with random intercepts and slopes

were also ran for each binary covariate to investigate whether there were different rates of changes in TICS for groups of subjects.

To investigate whether missing data would change the results of the analysis, all the models described previously were repeated using the 300 complete cases and the full dataset of 762 subjects whose missing data was replaced by the last observations carried forward method.

Results

<u>Baseline TICS Score</u>

Table 1 displays the covariates measured at baseline for the overall study sample and for centenarian offspring and controls. A crude logistic regression between baseline score and subject group found that offspring of centenarians had on average 0.7025 higher TICS score than the controls (95% Credible Interval: 0.0978, 1.3130) (Table 2). Geweske Z-scores comparing the first 10% and last 50% of the simulations showed that the model converged (Table 5).

The adjusted linear regressions found that out of the 12 baseline covariates, 9 of them caused the adjusted subject group parameter to be 10% greater or less than the crude estimate parameter of 0.7025, and thus were considered confounders of the relationship between baseline TICS score and subject group (Table 2).

The results of the final multiple linear regression model, adjusting for sex, age at enrollment, BMI, smoking status, history of diabetes mellitus, history of hypertension, history of coronary artery disease, history of heart attack, and years of education, are in Table 3 and convergence diagnostics are displayed in Table 6. For subjects with the same exact covariates, offspring of centenarians had on average 0.16215 higher TICS score than the controls (95% Credible Interval: -0.5157, 0.75327). This association was no longer significant, as the credible interval contains 0. The distributions of baseline TICS score were not different between children of centenarians and controls after adjusting for potential confounders.

<u>Rate of Change in TICS Score</u>

Table 4 shows the results of the interaction analysis from the longitudinal models. Convergence diagnostics are in Table 6. There was no difference in rates of change of TICS scores over time between offspring of centenarians and controls. There was also no difference in rates when comparing participants of different sex, smoking status, aspirin use, history of stroke, history of diabetes mellitus, history of hypertension, history of cancer, and history of heart attack. However, there was a difference in rates of change in TICS scores when looking at history of coronary artery disease. On average, offspring of centenarians had 0.08871 higher TICS scores at each subsequent follow-up (95% Credible Interval: 0.01137, 0.1474).

<u>Investigating Missing Data – Complete Case Only (n=300)</u>

Using only complete cases, 227 offspring of centenarians and 73 controls, the crude logistic regression found that offspring of centenarians had on average 0.1916 lower TICS score than the controls (95% Credible Interval: -1.097, 0.8941) (Table 7). Ten confounders were

identified: sex, smoking status, aspirin use, history of stroke, history of diabetes mellitus, history of hypertension, history of coronary artery disease, history of cancer, history of heart attack, and years of education (Table 8). The adjusted linear regression found that for subjects with the same exact covariates, offspring of centenarians on average had 0.60832 lower TICS score than the controls (95% Credible Interval: -1.67504, 0.4579). Neither of these associations were significant, so the distributions of baseline TICS score were not different between children of centenarians and controls, even after adjusting for potential confounders.

Table 9 shows the results the interaction analysis from the longitudinal models. There was no difference in rates of change of TICS scores over time between offspring of centenarians and controls and when comparing participants of different smoking status, aspirin use, history of diabetes mellitus, history of hypertension, history of coronary artery disease, and history of cancer. There was a difference in rate of change when comparing sex, history of stroke, and history of heart attack.

Investigating Missing Data – Last Observation Carried Forward (n=762)

The crude logistic regression of 492 centenarians and 270 controls found that offspring of centenarians had on average 0.5467 higher TIC score than the controls (95% Credible Interval: -0.04737, 1.192) (Table 10). Eight confounders were identified: sex, BMI, smoking status, aspirin use, history of stroke, history of diabetes, history of hypertension, history of and coronary artery disease (Table 11). The adjusted linear regression found that for subjects with the same exact covariates, offspring of centenarians had on average 0.00309 higher TICS score than the controls (95% Credible Interval: -0.447, 0.75561). This association was no longer significant, as the credible interval contains 0. The distributions of baseline TICS score were not different between children of centenarians and controls after adjusting for potential confounders.

Table 12 shows the results of the interaction analysis from the longitudinal models. There was no difference in the rates of change of TICS scores over time between offspring of centenarians and controls and when comparing participants of different sex, smoking status, aspirin use, history of diabetes mellitus, history of hypertension, history of coronary artery disease, history of cancer, and history of heart attack. However, there was a difference in rates of change in TICS scores when looking at history of coronary stroke. On average, offspring of centenarians had 0.1053 higher TICS scores at each subsequent follow-up (95% Credible Interval: 0.01528, 0.2936).

Discussions & Conclusions

This study found that there was no different in baseline TICS scores between children of centenarians and controls after adjusting for sex, age at enrollment, BMI, smoking status, history of diabetes mellitus, history of hypertension, history of coronary artery disease, history of heart attack, and years of education. There was no difference in rates of change of TICS scores over time between offspring of centenarians and controls and comparing every other group except for history of coronary artery disease.

This study was not without its limitations. In this dataset, age at enrollment had a wide range between 45 and 90. Cognitive decline is not the same for middle-aged adults and the elderly, so a future study should have a fixed age at enrollment. Also, the offspring of centenarian sample was heavily skewed towards women. Cognitive decline is different in men and women, so that might cause a bias in the analysis.

Most of the missing data were in the outcome measurements, not in the covariates. 38 out of the 40 missing covariate cases were because of unknown BMI. The later the TICS measurement, the more missing entries there were, which is expected in longitudinal studies. When only the 300 complete cases were considered, the parameter estimate for difference in baseline TICS score went in the opposite direction, which shows how ignoring missing data can cause . Imputing the missing data using last observation carried forward gave very similar results as the complete covariate cases analysis, which indicates that the results are robust to missing data mechanisms.

Appendix

Table 1. Descriptive Statistics of Dataset at Baseline

State Medical Marijuana Law Passed	Controls n=247 (31.80%)	Centenarian Offspring n=473 (68.20%)	Overall n=720 (100%)	p-value
Sex				
Male	135 (54.66%)	171 (36.15%)	306 (42.50%)	<0.0001
Female	112 (45.34%)	302 (63.85%)	414 (57.50%)	
Age at Enrollment	71.52 (6.83)	70.77 (7.23)	71.0 (7.10)	0.1673
BMI	27.7 (4.63)	27.0 (4.91)	27.3 (4.83)	0.0281
Smoker				
No	86 (34.82%)	225 (47.57%)	311 (43.19%)	0.0014
Yes	161 (65.18%)	248 (52.43%)	409 (56.81%)	
Aspirin				
No	115 (46.56%)	262 (55.39%)	377 (52.36%)	0.0297
Yes	132 (53.44%)	211 (44.61%)	343 (47.64%)	
Stroke				
No	236 (95.55%)	466 (98.52%)	702 (97.50%)	0.0296
Yes	11 (4.45%)	7 (1.48%)	18 (2.50%)	
Diabetes Mellitus				
No	220 (89.07%)	448 (94.71%)	668 (92.78%)	0.0086
Yes	27 (10.93%)	25 (5.29%)	52 (7.22%)	
Hypertension				
No	135 (54.66%)	332 (70.19%)	467 (64.86%)	<0.0001
Yes	112 (45.34%)	141 (29.81%)	253 (35.14%)	
Coronary Artery Disease				
No	220 (89.07%)	457 (96.62%)	677 (94.03%)	<0.0001
Yes	27 (10.93%)	16 (3.38%)	43 (5.97%)	
Cancer				
No	180 (72.87%)	360 (76.11%)	540 (75.00%)	0.3892
Yes	67 (27.13%)	113 (23.89%)	180 (25.00%)	
Heart Attack				
No	231 (93.52%)	452 (95.56%)	683 (94.86%)	0.3183

Yes	16 (6.48%)	21 (4.44%)	37 (5.14%)	
Years of Education	15.6 (2.82)	15.9 (2.95)	15.8 (2.90)	0.2907
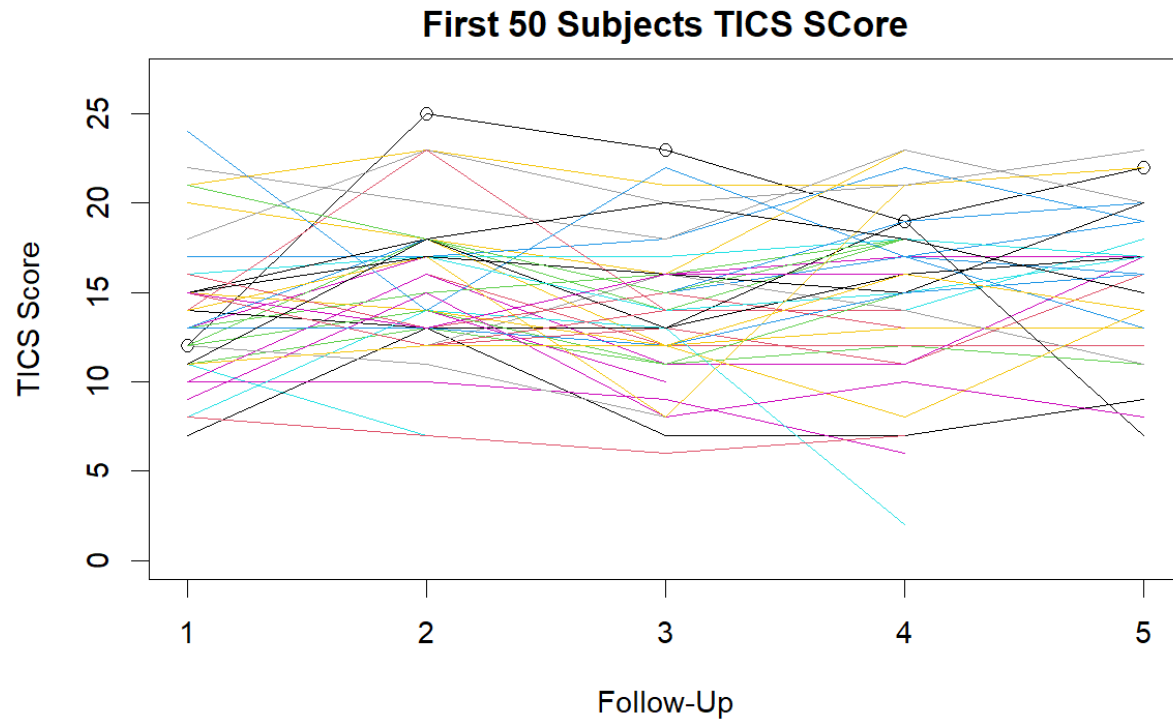
Figure 1. TICS Score of First 50 Subjects

Table 2. Adjusted Linear Regressions to Check Confounders

Covariate	Parameter Estimate for Group Difference	95% Credible Interval	% Change from Crude Estimate	Confounder?
Crude Regression (none)	0.7025	(0.0978, 1.3130)	-	-
Sex	0.3716	(-0.2878, 1.0040)	47.10%	Yes
Age at Enrollment	0.6036	(0.0551, 1.1669)	10.01%	Yes
BMI	0.5664	(-0.2387, 1.2537)	19.37%	Yes
Smoker	0.7735	(0.1625, 1.3459)	-10.11%	Yes
Aspirin	0.7420	(-0.0745, 1.4622)	-5.62%	No
Stroke	0.6718	(0.0648, 1.1897)	4.37%	No
Diabetes Mellitus	0.6200	(0.0342, 1.4060)	11.74%	Yes
Hypertension	0.5705	(-0.0581, 1.2450)	18.79%	Yes
Coronary Artery Disease	0.5457	(-0.0186, 1.0310)	22.32%	Yes
Cancer	0.6967	(0.0697, 1.2166)	0.83%	No
Heart Attack	0.6273	(0.0166, 1.1990)	10.70%	Yes
Years of Education	0.5250	(-0.0217, 1.0967)	25.27%	Yes

Table 3. Adjusted Linear Regression

Covariate	Parameter Estimate	95% Credible Interval
Subject Group	0.16215	(-0.5157, 0.75327)
Sex	1.63655	(1.028, 2.19946)
Age Enrollment	-0.1798	(-0.2213, -0.14366)
BMI	-0.02673	(-0.0826, 0.03256)
Smoker	0.41606	(-0.199, 1.0063)
Diabetes Mellitus	-0.65141	(-1.9483, 0.45396)
Hypertension	-0.26139	(-0.9064, 0.32782)
Coronary Artery Disease	0.12773	(-1.2504, 1.42102)
Heart Attack	-1.87129	(-3.317, -0.49551)
Years Education	0.24043	(0.154, 0.35369)

Table 4. Rates of Change in TICS Scores

Group	Interaction Parameter Estimate	95% Credible Interval	Different Rate of Change?
Subject Group	-0.0005602	(-0.09807, 0.04968)	No
Sex	0.02633	(-0.0141, 0.07229)	No
Smoker	0.01271	(-0.02724, 0.06224)	No
Aspirin	0.003785	(-0.0465, 0.07384)	No
Stroke	- 0.05302	(-0.2092, 0.148365)	No
Diabetes Mellitus	0.01835	(-0.03595, 0.1067)	No
Hypertension	0.03083	(-0.03139, 0.08379)	No
Coronary Artery Disease	0.08871	(0.01137, 0.1474)	Yes
Cancer	-0.005835	(-0.05153, 0.04493)	No
Heart Attack	0.009271	(-0.05711, 0.09184)	No

Table 5. Convergence Diagnostics for Crude Linear Regression

	Geweke Z-Score		
	First Arm	**Second Arm**	**Third Arm**
Subject Group	1.420	0.5484	1.367

Table 6. Convergence Diagnostics for Adjusted Linear Regression

	Geweke Z-Score		
	First Arm	**Second Arm**	**Third Arm**
Subject Group	0.3504	1.4689	0.1101
Sex	-1.2663	0.1957	-1.2268
Age Enrollment	-6.2174	2.8598	2.6997
BMI	-1.5271	1.4810	0.2735
Smoker	-0.8097	0.6426	0.9787
Diabetes Mellitus	-1.1314	1.2725	0.9732
Hypertension	0.9479	-0.4897	1.6412
Coronary Artery Disease	0.1197	-0.5286	-3.5476
Heart Attack	2.3020	-2.0989	3.3210
Years Education	-4.1324	-1.9026	1.6037

Table 6. Convergence Diagnostics for Longitudinal Analysis

	Geweke Z-Score		
	First Arm	**Second Arm**	**Third Arm**
Subject Group	2.8321	3.7445	3.467
Group*Time	-0.9091	0.3126	2.059

Table 7. Adjusted Linear Regressions to Check Confounders

Covariate	Parameter Estimate for Group Difference	95% Credible Interval	% Change from Crude Estimate	Confounder?
Crude Regression (none)	-0.1916	(-1.097, 0.8941)	-	-
Sex	-0.273	-1.2251, 0.8227	-42.48%	Yes
Age at Enrollment	-0.1849	-1.0510, 0.90299	3.50%	No
BMI	-0.19967	-1.0343, 0.92327	-4.21%	No
Smoker	-0.08773	-1.0341, 0.8059	54.21%	Yes
Aspirin	-0.1256	-1.185, 0.9147	34.45%	Yes
Stroke	-0.3095	-1.224, 0.9358	-61.53%	Yes
Diabetes Mellitus	-0.1559	-1.149, 1.04125	18.63%	Yes
Hypertension	-0.2484	-1.150, 0.7017	-29.65%	Yes
Coronary Artery Disease	-0.237	-1.011, 0.8014	-23.70%	Yes
Cancer	-0.1685	-1.263, 0.7995	12.06%	Yes
Heart Attack	-0.1647	-1.176, 0.8541	14.04%	Yes
Years of Education	0.0003149	-1.01922, 0.9536	100.16%	Yes

Table 8. Adjusted Linear Regression

Covariate	Parameter Estimate	95% Credible Interval
Subject Group	-0.60832	(-1.67504, 0.4579)
Sex	2.193932	(1.21051, 3.1363)
Smoker	0.375123	(-0.4858, 1.2445)
Aspirin	0.002585	(-0.89268, 0.7972)
Stroke	-3.06683	(-7.04089, 0.6701)
Diabetes Mellitus	-0.75506	(-2.39524, 0.9052)
Hypertension	-1.03382	(-1.98, -0.0483)
Coronary Artery Disease	-1.12104	(-3.62867, 1.1187)
Cancer	-0.40542	(-1.47188, 0.6131)
Heart Attack	-0.49878	(-3.29576, 1.9123)
Years Education	0.183717	(0.03306, 0.33)

Table 9. Rates of Change in TICS Scores

Group	Interaction Parameter Estimate	95% Credible Interval	Different Rate of Change?
Subject Group	0.05195	(-0.08225, 0.12331)	No
Sex	0.1039	(0.01508, 0.1928)	Yes
Smoker	-0.001291	(-0.0640, 0.08355)	No
Aspirin	-0.001596	(-0.07912, 0.08660)	No
Stroke	0.245	(0.02898, 0.4485)	Yes

Diabetes Mellitus	0.01714	(-0.2389, 0.11317)		No
Hypertension	0.0458	(-0.03279, 0.1348)		No
Coronary Artery Disease	0.09907	(-0.1163, 0.18214)		No
Cancer	-0.08369	(-0.1788, -0.00616)		No
Heart Attack	0.3678	(0.1590, 0.4813)		Yes

Table 10. Adjusted Linear Regressions to Check Confounders

Covariate	Parameter Estimate for Group Difference	95% Credible Interval	% Change from Crude Estimate	Confounder?
Crude Regression (none)	0.5467	(-0.04737, 1.192)		-
Sex	0.2325	(-0.5382, 0.8062)	57.47%	Yes
Age at Enrollment	0.4941	(-0.1426, 1.0460)	9.62%	No
BMI	0.4515	(-0.1557, 1.11866)	17.41%	Yes
Smoker	0.4841	(-0.1482, 1.1401)	11.45%	Yes
Aspirin	0.4486	(-0.02895, 1.1980)	17.94%	Yes
Stroke	0.3742	(-0.1354, 0.9139)	31.55%	Yes
Diabetes Mellitus	0.3656	(-0.1476, 0.9438)	33.13%	Yes
Hypertension	0.4315	(-0.07104, 0.9541)	21.07%	Yes
Coronary Artery Disease	0.3477	(-0.2287, 1.086)	36.40%	Yes
Cancer	0.5080	(-0.08353, 1.1500)	7.08%	No
Heart Attack	0.4961	(0.008345, 1.085)	9.26%	No
Years of Education	0.5299	(-0.08037, 1.2548)	3.07%	No

Table 11. Adjusted Linear Regression

Covariate	Parameter Estimate	95% Credible Interval
Subject Group	0.03309	(-0.447, 0.75561)
Sex	1.67253	(1.1735, 2.27386)
BMI	-0.07659	(-0.1201, -0.02403)
Smoker	0.14204	(-0.4531, 0.76613)
Aspirin	-0.22042	(-0.7783, 0.27367)
Stroke	-0.99095	(-2.7793, 0.7724)
Diabetes Mellitus	-0.83828	(-1.935, 0.63439)
Hypertension	-0.36200	(-1.026, 0.21377)
Coronary Artery Disease	-0.74497	(-1.3828, -0.1401)

Table 12. Rates of Change in TICS Scores

Group	Interaction Parameter Estimate	95% Credible Interval	Different Rate of Change?
Subject Group	-0.04232	-0.1012, 0.02562	No
Sex	-0.02223	-0.07659, 0.04879	No
Smoker	-0.005719	-0.03983, 0.03447	No
Aspirin	0.0003725	-0.04473, 0.05892	No
Stroke	0.1053	0.01528, 0.2936	Yes
Diabetes Mellitus	0.0328	-0.03876, 0.1011	No
Hypertension	0.001753	-0.03721, 0.044642	No

Coronary Artery Disease	0.02285	-0.06541, 0.09976		No
Cancer	-0.0135	-0.1080, 0.03155		No
Heart Attack	-0.10517	-0.1680, 0.05261		No