Final Project

BS849 - Bayesian Modeling in Public Health

Data

The data in the file `tics.data.2021.csv` report up to 5 longitudinally collected measures of "Telephone Interview for Cognitive Status" (`TICS`) in offspring of centenarians and controls enrolled in two studies of longevity and healthy aging conducted at Boston University. `TICS` is a test that consists of various questions that are designed to assess a person level of cognitive function and can be administered over the phone. Variables included in the file are:

- `ID`: patient ID
- `Fam.num`: family ID
- `Sex`: 0-Females, 1-Males
- `Ptype`: 0 = Controls, 2= Centenarian offspring
- `Age.at.Enrollment`: Age when subjects enrolled in the study
- `Age.Last.Contact`: Age when last seen, or age at death for diseased subjects
- `Deceased`: `yes` if participants has died
- `BMI`
- `SH.Ever.Smoked.`: `yes` for current or past smokers at enrollment
- `MC.Aspirin`: `yes` for regular aspirin use at enrollment
- `MC.Stroke`: `yes` for history of stroke at enrollment
- `MC.Diabetes.Mellitus`: `yes` for history of diabetes at enrollment
- `MC.HTN`: `yes` for history of hyper-tension at enrollment
- `MC.Coronary.Artery.Disease`: `yes` for history of CAD at enrollment
- `MC.Cancer`: `yes` for history of cancer, excluding skin cancer, at enrollment
- `MC.Heart.Attack`: - `yes` for history of heart atack
- `Years.of.Education`: at enrollment
- `TICS01`—`TICS05`: `TICS` score in 5 longitudinal follow-ups
- `Age01`—`Age05`: Age at `TICS` tests.

Report writing

Use R Markdown to write your report. Scripts and software output should be included separate from your report and can be uploaded to Blackboard. My suggestion is to keep the report brief (around 3 pages, at least 10pts fonts and 2 additional page of figures and tables). You can include details of the analysis as `PDF` or `HTML` file from `R` markdown.

In this project, you will use the data to try to address the following:

1) What are the variables that are different when comparing offspring of centenarians and controls at baseline? This analysis will be useful to select a set of possible confounders to include in the next steps. *Hint: Analyze variables one at a time.*
2) Are the distributions of `TICS` score at baseline different between offspring of centenarians and individuals without parental longevity after you adjust for possible confounders?
3) Use hierarchical models to investigate whether there are differences in the rate of change of `TICS` between offspring of centenarians and individuals without parental longevity
4) Are there groups of study participants with significantly different rates of change of `TICS`?
5) Report on the missing data. Are the results robust to missing data mechanism?

You should conduct the analyses that you think are most appropriate and then write a report that includes:

i) Introduction to the report and description of your research questions;
ii) Methods section that describes all statistical methods used in the analysis. Be complete in your presentation, including the choice of priors and the justification for those choices.
iii) Results of your analysis. Your results should include a table with summary of patients' characteristics at baseline, and a summary table that shows the results of the analyses. Provide some graphical displays of the results that you think are most useful to convey the results and convergence diagnostics.
iv) A brief paragraph with discussion and conclusions.