<u>Comments on the paper authored by Group 09 as Refereed by Group 08</u>

- The questions of interest are well presented and the method is proper.

- As a general note, there are grammatical errors throughout the paper.

- The abbreviation "DSST" is used throughout the paper, but never defined in the paper. It should be made clear what this is.

- Phrasing in lines 31-33 regarding sex modification is a bit confusing.

- On page 3, lines 104-106, it is written that "cerebral atrophy was broken up into 3 equal groups which the author will refer to as minimal atrophy (0-29), mild atrophy (30-59), and severe atrophy (60+)". It is unclear how these three groups are equal. Additionally, are these groups divided for clinical reasons? Because it sounds like it was divided through both methods which is confusing.

- On page 4 line 127, it is written that "The categorical variables with multiple groups were fit using dummy variables, which include stroke status (0 = no stroke events, 1 = diagnosis of transient ischemic attack, 2 = diagnosis of stroke)". If you model smoke status as dummy variables, "0 = no stroke events, 1 = diagnosis of transient ischemic attack, 2 = diagnosis of stroke" is a little bit consusing

- On page 4 line 129, It is written "coronary heart disease history (0=no events, 1=angina, 2=myocardial infarction)". Same problem with above.

- The wording lines 133-140 in the methods is confusing.

- The sentence 147-149 has redundant phrases

- Lines 158-161 – why are they included? It seems redundant and unnecessary.

- Define what AAI is

- On page 5, lines 179-180, it is written that "735 subjects who were followed for death from any cause for a Kaplan-Meier estimated average of 5.33 years (range 5.00 to 5.91 years)". However, from Table 2, the minimum observation time is 0.19 years, so should 5.00 be replaced with 0.19?

- On page 5, lines 184-187, it is written that "At 5 years, there is a ~20% difference between mild to moderate atrophy score survival probabilities and moderate to severe atrophy score survival probabilities, with the highest in the mild atrophy category at 90.08% and the lowest in the severe atrophy category at 68.97%." However, saying "there is a ~20% difference between mild to moderate atrophy score survival probabilities and moderate to severe atrophy score survival probabilities" implies that there is a 20% difference between mild and moderate as well as a 20% difference between moderate and severe.

- On page 5, lines 187-190, it is written that "The subjects in the mild atrophy score group averaged 5.64 years of life, while those in the moderate category averaged 5.32 years of life, and the lowest in the severe atrophy group at 4.55 years of life." This is scientifically incorrect, since the sample has minimum age of 65, so all subjects averaged at least 65 years of life. The use of the restricted mean here should be clarified. In general, language describing survival in the results section is indistinct is a bit hard to follow.

- On page 6, lines 201-202, it is written that "we estimate each year increase in age is associated with an increase in mean atrophy score of 0.696", but this might be interpreted as referring to the increased mean atrophy score within one individual after a year has gone by, which is not the question that this analysis answered.

- Lines 221-222 Note what the covariates are

- On page 7, line 244, the phrase "good health scores" is used. What does this mean?

- Lines 250-254 in discussion is a run-on sentence, difficult to interpret.

- Overall, Discussion section has a lot of restating of methods and results, which is likely not necessary. Why not discuss the limitations or future study directions? You never discussed what the data was missing in your description of data so at least note it in discussion.

- Looking at the tables, Table 1 is visually hard to follow, with lots of gaps in it. Reformatting might be helpful.

- In figures 1, 2, and 3, there should be units included on the x-axis of the Kaplan-Meier plots, since "analysis time" is vague.

- Table 1: should we also give 95% CI for the proportion of categorical variable? For the two variables that have missing data, in which atrophy score category does the missing data exist?

- Table 2: The unit "years" could be removed to the first column. "Mean time of follow-up (years)"

- Table 3: Would be helpful to explain the column "atrophy". Not everyone understands the slope outputs from a regression model, and so this table is unclear. Table 3 should explain the estimate, otherwise it is confusing

- Table 4: again, visually hard to follow.

- In Tables 3 and 4, there are some confusing row variables, such as "Stroke TIA" and "CHD Angina". These should be made more clear.

- Table 5: probably unnecessary. You could just state the results of the multiple partial testing in the text, as the table only contains two values.

-