Comments on the paper authored by Group 07 as Refereed by Group 06

General Requirements:
The group met the length requirements (between 1-12 pages) not counting figures and tables and the requirements on the number of figures and tables.

Summary:
It may be useful to note what the potentially confounding health risk factors are.

Background:
It was not clear whether atrophy was as important in predicting mortality as cerebral infarction. They appeared to have been given equal weights in the Background. In addition, the authors chose to investigate atrophy as a biomarker rather than infarct-like lesions, which seems to be more predictive than atrophy, which may not be the case. It's recommended that the authors remove the term infarct-like lesions (or clarify why atrophy was the focus on their investigation) as this tends to mislead the readers into thinking that the study was going to investigate infarct-like lesions as a potential biomarker.

It is recommended that the term "marker" be changed to "biomarker" to remain consistent with your first use of term.

Questions of Interest:
These seem appropriate. The authors established the scientific questions that they believed a priori they could address with the dataset. However, it would have been interesting to see whether effect modification was present across different age groups. The Background alluded that elderly patients may be selected out due to survivorship. It would be reasonable to see how 90 year olds fared against 80 years old. Perhaps, the author should tackle this scientific question in a secondary analysis.

Description of Data:
It appears that the time of MRI to measure cerebral atrophy was performed at approximately 3 years. Please indicate if this was 3 years exactly or if there was some variability on what this point estimate is. For instance, is this a mean of 3.0 years with a range of 2.5 to 5 years?

The authors write "at the time of the MRI, investigators also collected information on thousands of variables pertaining to each subject." Please clarify if there actually were "thousands of variables" collected. I find this unlikely, especially since the dataset that was presented to them was most likely trimmed down to its most important components and unlikely to have "thousands of variables" as suggested.

Table 1--General Health category "Poor" for the column Atrophy score > 60 is empty. Since this was not due to missing data, it's likely a "0" count with 0%. Please indicate this rather than leaving the cell empty. In addition, please include a footnote explaining what the acronyms mean. For instance, what does DSST mean and what is the range?

Statistical Methods:
5.1 It would be helpful to note what the acronyms stand for since they were never defined. Presumably the collaborator would be familiar with these terms, but it would probably be safe to confirm.

5.2 The authors mention dividing atrophy into the levels in 5.1 and also in 5.2. It might be best to combine this with the description in 5.1.

5.3 Did the authors use classic linear regression or robust? What is the rationale for the selection? In the last paragraph the authors mention using robust methods, is this true for the linear regression? The first paragraph is a little confusing-- was this to assess association after adjustment or just association between variables and atrophy score? From the first sentence it sounds like the latter, but in the second sentence the authors mention "controlling" for variables. It is also unclear from the last sentence whether the authors opted to use a multiplicative model or not. What is "rescaling"? The collaborator may be unfamiliar with the term heteroscedasticity, so it may be useful to define the term.

Overall, it would be useful to note if 95% CIs will be reported and whether the p-values are one or two-sided. Are the authors using Wald Statistics to calculate CIs or another method?
It would be nice to state the assumptions we make for using a PH regression and whether it suits our data.

Results:
5.1 The authors repeat needing the information about MCAR/MAR. It might be best to include this just in the methods section. Since the authors mention behavioral characteristics and measures of organ system functioning it could be a nice touch to incorporate these "subheadings" into the table for ease of understanding.

Unclear why KM curves cannot be used with MCAR. Please expand. The authors write that "For purposes of analysis with censored data it is important to know whether the data is MCAR or MAR." However, there was no mention of data missing patterns that are not missing at random (NMAR). I believe that MCAR and MAR are find for KM curves; but not NMAR. Please verify this.

5.3 Are the p-values two-sided? How were the other adjusted variables chosen and were they transformed for the model? What are your confounders?
Why are variables categorized into demographics and risk factors as we should adjust for all of them anyways.

You should choose which variables to adjust for based on scientific reasons. Table 4 might be a process you do but not present in the scientific paper? I'm not too sure the order and purpose of it.

Discussion:
Table 1 presents descriptive statistics rather than inferential statistics-- is this a typo (second sentence of first paragraph)?

In Table 4 the authors present adjustments, but it is unclear why the adjustments are added in. Was there a rationale for adding pack-years before cardiovascular disease into the model?

No discussion about the potential correlation between age and atrophy score as an issue in their adjusted model was presented. In Table 1 there is a clear trend. The authors state this in the Descriptive Analysis, but did not pursue this further.

A conclusion section would nicely summarize your findings. Please include.

Some minor comments on the format of tables and figure

Tables/Figures:
Figure 1:
“distribution of time from study enrollment to death” is not clean, should change to something like survival function?

Tables 3:
95% CI low and 95% CI high maybe combined into one column if space limits is an issue

Table 4:
Z score is normally not of our interest and we could remove this column.
Since there are only two adjustment lead to non-significant p-value, it may just report the last two rows since others are not of our interests.