**Response to the reviewers**

**Decision letter**

I am pleased to inform you that your manuscript is acceptable for publication in the Journal of Alzheimer's Disease pending minor revision.

Below is a link to the decision and reviewers' comments regarding your submission. Please click it to confirm receipt of this letter.

Please revise your manuscript according to the reviewers' suggestions and provide a point-by-point response to the reviews at the beginning of your revised manuscript. Please track changes in your manuscript and submit it to our online submission system (<http://mstracker.com/submit1.php?jc=jad>). Be sure the manuscript is formatted per our instructions to authors (<http://www.j-alz.com/prep>). When resubmitting please mention the reference number in the cover letter. Your revision should be returned within two months.

Sincerely,

George Perry

Editor-in-Chief, JAD

*We appreciate the time spent by the editors and reviewers in assessing our manuscript. Below is a point-by-point response. Please note that textual changes are highlighted in blue font in the current version of the manuscript.*

**Reviewer 1**

The paper titled "Longitudinal mapping of cortical thickness measurements: an ADNI-based evaluation study" compared 5 different cortical thickness measurement methods: ANTS Cross, ANTS SST, ANTS Native, FreeSurfer-Cross and FreeSurfer Long in terms of the variance ratio, and statistical power for distinguishing LMCI-CN, AD-LMCI, and AD-CN through the linear mixed effect (LME) model. While freesurfer is the most popular tool for cortical thickness analysis, ANTs is now quite popular and their comparisons have done before [26]. I believe that this manuscript's main contribution is introducing the new ANTs longitudinal pipeline first time (though it was released through BioXriv in 2018) Also, its thorough data analysis including LME, a popular analysis approach for the longitudinal data, raised its value by guiding others to use the longitudinal pipeline and statistical analysis. While the main scientific conclusion of the paper is the importance of longitudinal pipeline, as the author indicates in page 15, Section 2.4 line 8. Its real value is not just that. So, I strongly recommend this paper.

*We thank the Reviewer for their willingness to review our manuscript and the thoughtfulness of the suggestions provided below.*

I have two major issues.

1. Organization of paper.

1) too much details in the method.

The method has too much details for the journal paper. I think some of them should move to the appendix: Table 1, Figure 4 and Figure 5. Figure 1 and 2 look quite interesting but they are not essential for the results. So, it is recommended to move to the appendix also. I do not think the figure 9 shows important findings. If it is please clarify in the text. If not, it is also recommended to move to the appendix.

*We agree that the current manuscript contains unnecessary figures. As we write below, this work was originally submitted elsewhere, went through a few rounds of reviews and, as such, accumulated material requested by the Reviewers (which we included even if we did not necessarily agree with the request). Specifically, we removed Figures 1, 2, and 5. We do not think they add to the specific focus of the work and therefore do not include them in the Appendix, either. However, they are still available in our GitHub repository.*

*We prefer to keep Figure 4 in the manuscript as template choice plays an important role in the processing pipeline and we want the readers to be able to easily visualize the template and corresponding prior spatial maps (which are also available in our GitHub repository). Table 1 is important for inspecting the results. As mentioned in the Table caption, the ROI abbreviations in Table 1 are used in later figures illustrating the results (current Tables 2 & 3 and Figure 3). Former Figure 9 (current Figure 6) is also important in terms of the findings as it essentially summarizes Tables 2 and 3. As we write in* **Section 3.2***:*

Figure 6 provides a side-by-side comparison of the distribution of log-scaled *p*-values separated into left and right hemispherical components and grouped according to contrast.

2) Section 2.3~2.6

I think Section 2.3~2.6 are all about the statistical analysis. Also Section 2.5 is relatively long since it has two contents in it: review of LME and variance ratio. I think breaking it into two section will make easier to read. How about the following organization?

Section 2.1. ADNI

Section 2.2. ANTs

Section 2.3. Statistical analysis

Section 2.3.1. Evaluation between cross-sectional and longitudinal

pipelines

Section 2.3.2. Review of LME

Section 2.3.3. variance ratio for between-subject thickness

variability evaluation

Section 2.3.4. Regional diagnostic contrasts based on cortical

atrophy

*Done.* **Section 2 Methods and materials** *is now organized as follows:*

**2.1 ADNI-1 imaging data**

**2.2 ANTs cortical thickness**

**2.2.1 Cross-sectional processing**

**2.2.2 Unbiased longitudinal processing**

**2.3 Statistical analysis**

**2.3.1 Evaluation of cross-sectional and longitudinal pipelines**

**2.3.2 Longitudinal biomarkers and the use of linear mixed effects modeling**

**2.3.3 A variance ratio for assessing residual and between-subject cortical**

**thickness variability**

**2.3.4 Regional diagnostic contrasts based on cortical atrophy**

3) Better introduction...

I think section 2.4's first paragraph is more clear to understand the real intention of the paper, compared to the last paragraph of the introduction. Why don't the authors try to use the wording there for the introduction?

*Done. We changed the last paragraph of the* **Introduction** *to reflect more closely the first paragraph of the previous* **Section 2.4** *(now* **Section 2.3.1***). Specifically, it was changed from*

In this work, we introduce the longitudinal version of the ANTs registration-based cortical thickness pipeline and demonstrate its utility on the publicly available ADNI-1 data set. In addition, we demonstrate that certain longitudinal processing choices have significant impact on measurement quality in terms of residual and between-subject variances which is known to impact the scientific interpretability of results, produce tighter confidence intervals in calculated mean trends and smaller prediction intervals, as well as less varied confidence/credible intervals for discerning cross-sectional effects. ~~Similar to previously outlined research, we show that reorienting individual time point images to a single-subject template has favorable performance effects which guide processing choices for the proposed ANTs longitudinal pipeline. Although we limit exploration in this work to ROI-based analysis for simplifying comparison with FreeSurfer, there are several additional applications permitted by the ANTs framework such as longitudinal tensor-based morphometry, Eigenanatomy, and extension to non-human data.~~

*Done. We changed the last paragraph of the* **Introduction** *to reflect more closely the*

In this work, we introduce the longitudinal version of the ANTs registration-based cortical thickness pipeline and demonstrate its utility on the publicly available ADNI-1 data set. In addition, we demonstrate that certain longitudinal processing choices have significant impact on measurement quality in terms of residual and between-subject variances which is known to impact the scientific interpretability of results, produce tighter confidence intervals in calculated mean trends and smaller prediction intervals, as well as less varied confidence/credible intervals for discerning cross-sectional effects. *This evaluation strategy goes beyond previously used precision-style assessment quantities which are limited in determining the actual clinical utility of cortical thickness as a longitudinal biomarker.*

2. The journal scope?

I wonder that it is matched with the journal's scope. This manuscript is rather methodological paper. Though it uses the ADNI dataset, its main goal is to introducing ANTs longitudinal process and its superiority in diagnostic performance. So, I believe that it is more fit to Neuroimage or other methodology- oriented journals. I just let the editor and other reviewer to decide this.

*We originally submitted this work to Neuroimage but despite the constructive first couple rounds, we were ultimately unsuccessful with the selected reviewers. We also submitted to Human Brain Mapping on two different occasions with the pre-screening rejections being explained by variants of the ambiguous “Sorry but this is not what we are looking for.” In our subsequent search for alternative publication venues, we believed JAD would also be appropriate based, in part, on recent publications. As we wrote in our letter to the editor accompanying the original submission:*

Cortical morphology is a widespread biomarker for scientific inquiry into Alzheimer’s disease which is why we believe that this manuscript is appropriate for your journal. A cursory topical search yields similar type articles that have been recently published in JAD:

* J. Lee, et al., Sex-Related Reserve Hypothesis in Alzheimer's Disease: Changes in Cortical Thickness with a Five-Year Longitudinal Follow-Up, 2019.
* A. Pink, et al., Cortical Thickness and Depressive Symptoms in Cognitively Normal Individuals: The Mayo Clinic Study of Aging, 2018.
* L. Pasquini, et al., Increased Intrinsic Activity of Medial-Temporal Lobe Subregions is Associated with Decreased Cortical Thickness of Medial-Parietal Areas in Patients with Alzheimer's Disease Dementia, 2016.

Minor points.

1. Section 2.5, page 16, line 6: longitudinal mixed effects (LME)

I think generally LME stands for linear mixed effects, and the references are also used the wording "linear mixed effects". If the author did not used the non-linear mixed effects, please use the traditional nomenclature. (well, from the equations, the authors did not use non-linear models.)

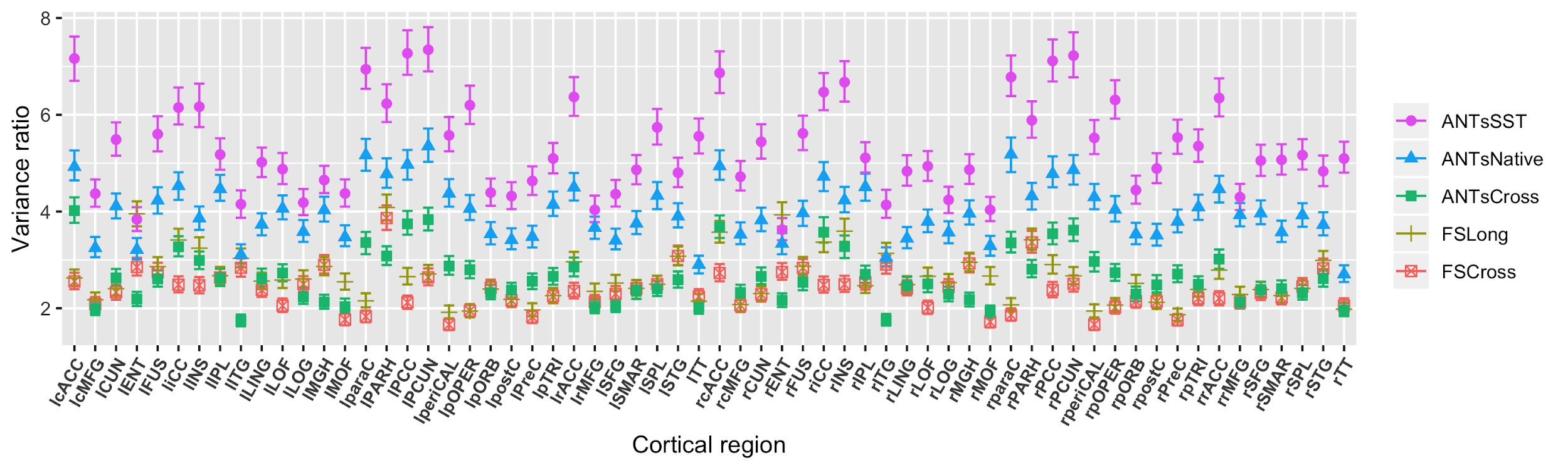
*Done. We changed (“longitudinal mixed effects” to “linear mixed effects”) the instance pointed out by the Reviewer in Section 2.5 as well as the one found in the Abstract.*

2. Section 2.5, page 19, line 4: less less -> less

*Done.*

3. Figure 6 legend on the right, the topmost markers are ANTs SST and bottommosts are FSCross in the figure, however, the legend on the right is in the reversed order. If the order of showing is reversed (matched with the Section 2.4's order) it is easier to read the figure!

*Done.*



4. Reference 68. Brain Mr Images -> Brain MR images

*Done.*

5. I am not sure JAD allows the reference like 27, 36, 46, 48, 49.

*Done. These references have been removed and replaced by footnotes within the text (except in the case of [49] where we now cite the formal reference provided by the Reviewer).*

6. In case of the reference 49, there is a formal reference. Nooner et al, (2012). The NKI-Rockland Sample: A model for accelerating the pace of discovery science in psychiatry. Frontiers in neuroscience 6, 152.

*Thanks. As mentioned in the previous item, we have replaced reference [49] with the formal reference provided by the Reviewer.*

**Reviewer 2**

The manuscript deals with a very important issue of processing longitudinal structural MR images in a robust way. The authors have compared two popular methods in the literature and analyzed the longitudinal data using sound statistical methods. I believe the proposed methods help many researchers in the neuroimaging field.

*We thank the Reviewer for the time spent reviewing our submission and appreciate the positive assessment of our work.*

**Reviewer 3**

In this paper, Tustison et al. evaluated the utility of longitudinal FreeSurfer and ANTs surrogate thickness values in the context of a longitudinal mixed-effects (LME) modeling strategy using the first phase of the Alzheimer's Disease Neuroimaging Initiative (ADNI-1) data. This paper is well written and deals with an important methodological issue regarding longitudinal cortical thickness measurement. As this ANTs framework is an open source, it may be useful for future researches.

*Again, we appreciate the time spent by the Reviewer in reading the manuscript as well as providing positive suggestions for improvement.*

Major Comments

The advantages of ANTs longitudinal pipeline over other methods were evaluated in a single large cohort. Please acknowledge in the discussion that this has to be replicated in other cohorts as well.

*Done. We have added the following to the* ***Discussion*** *section:*

Finally, it should be noted that while the current findings certainly have utility, they are limited to a specific population and the community would benefit from replication and exploration in other populations.

Minor Comments

1. Please write the full term of ANTs in the abstract.

*Done.*

2. Table 2 and 3 have identical title. Did the authors intend to demonstrate right and left hemisphere for each table?

*The Reviewer is correct and we have changed the title accordingly.*