bioRxiv preprint doi: https://doi.org/10.1101/170209; this version posted August 16, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

The ANTs Longitudinal Cortical Thickness Pipeline

Nicholas J. Tustison[1,2], Andrew J. Holbrook[3], Brian B. Avants[4†], Jared M. Roberts[2], Philip A. Cook[4], Zachariah M. Reagh[2], Jeffrey T. Duda[4], James R. Stone[1], Daniel L. Gillen[3], and Michael A. Yassa[2] for the Alzheimer's Disease Neuroimaging Initiative*

[1]Department of Radiology and Medical Imaging, University of Virginia, Charlottesville, VA

[2]Department of Neurobiology and Behavior, University of California, Irvine, Irvine, CA

[3]Department of Statistics, University of California, Irvine, Irvine, CA

[4]Department of Radiology, University of Pennsylvania, Philadelphia, PA

Corresponding author:

Nicholas J. Tustison

211 Qureshey Research Lab

Irvine, CA 92697-3800

ntustison@virginia.edu

---

† Currently employed by Biogen (Cambridge, MA).

*Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf

bioRxiv preprint doi: https://doi.org/10.1101/170209; this version posted August 16, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Abstract

Longitudinal studies of development and disease in the human brain have motivated the acquisition of large neuroimaging data sets and the concomitant development of robust methodological and statistical tools for quantifying neurostructural changes. Longitudinal-specific strategies for acquisition and processing have potentially significant benefits including more consistent estimates of intra-subject measurements while retaining predictive power. In this work, we introduce the open-source Advanced Normalization Tools (ANTs) registration-based cortical thickness longitudinal processing pipeline and its application to the first phase of the Alzheimer's Disease Neuroimaging Initiative (ADNI-1) comprising over 600 subjects with multiple time points from baseline to 36 months. We demonstrate in these data that the single-subject template construction and same orientation processing results in a simultaneous minimization of residual variability and maximization of between-subject variability immediately estimable from a longitudinal mixed-effects modeling strategy. It is known from the statistical literature that optimizing these dual criteria leads to greater scientific interpretability in terms of tighter confidence intervals in calculated mean trends, smaller prediction intervals, and narrower confidence intervals for determining cross-sectional effects. This strategy is evaluated over the entire cortex, as defined by the Desikan-Killiany-Tourville labeling protocol, where comparisons are made with the cross-sectional and longitudinal FreeSurfer processing streams. Subsequent linear mixed effects modeling for identifying diagnostic groupings within the ADNI cohort is provided as supporting evidence for the utility of the proposed ANTs longitudinal framework which provides unbiased structural neuroimage processing and competitive to superior power for longitudinal structural change detection.

Keywords: Advanced Normalization Tools, FreeSurfer, linear mixed effects models, longitudinal processing

bioRxiv preprint doi: https://doi.org/10.1101/170209; this version posted August 16, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

1 Introduction

Quantification of brain morphology facilitates the investigation of a wide range of neurological conditions with structural correlates, including neurodegenerative conditions such as Alzheimer's disease [1, 2]. Essential for thickness quantification are the computational techniques which were developed to provide accurate measurements of the cerebral cortex. These include various mesh-based (e.g., [3–5]) and volumetric techniques (e.g., [6–9]).

In inferring developmental processes, many studies employ cross-sectional population sampling strategies despite the potential for confounding effects [10]. Large-scale studies involving longitudinal image acquisition of a targeted subject population, such as the Alzheimer's Disease Neuroimaging Initiative (ADNI) [11], are designed to mitigate some of the relevant statistical issues. Analogously, much research has been devoted to exploring methodologies for properly exploiting such studies and avoiding various forms of processing bias [12]. For example, FSL's SIENA (Structural Image Evaluation, using Normalization, of Atrophy) framework [13] for detecting atrophy between longitudinal image pairs avoids a specific type of processing bias by transforming the images to a midspace position between the two time points. As the authors point out "[i]n this way both images are subjected to a similar degree of interpolation-related blurring." Consequences of this "interpolation-related blurring" were formally analyzed in [14] in the context of hippocampal volumetric change where it was shown that interpolation-induced artifacts can artificially create and/or inflate effect size [15]. These insights and others have since been used for making specific recommendations with respect to longitudinal image data processing [12, 16–18].

In [12, 19], the authors motivated the design and implementation of the longitudinal FreeSurfer variant inspired by these earlier insights and the overarching general principle of "treat[ing] all time points exactly the same." It has since been augmented by integrated linear mixed effects modeling capabilities [20] and has been used in a variety of studies including pediatric cortical development [21], differential development in Alzheimer's disease and fronto-temporal dementia [22], and fatigue in the context of multiple sclerosis [23]. Although the FreeSurfer longitudinal processing stream is perhaps one of the most well-known, other important longitudinal-specific methodologies have been proposed for characterizing cortical morphological change. Similar to FreeSurfer, cortical surfaces

bioRxiv preprint doi: https://doi.org/10.1101/170209; this version posted August 16, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

are generated in [24, 25] permitting vertex-wise quantitation of thickness and thickness change. Application to early infants in [24] further demonstrate the utility of targeted longitudinal considerations.

We introduced the Advanced Normalization Tools (ANTs) cortical thickness pipeline in [26] which leverages various pre-processing, registration, segmentation, and other image analysis tools that members of the ANTs and Insight Toolkit (ITK) open-source communities have developed over the years and disseminated publicly [27]. This proposed ANTs-based pipeline has since been directed at a variety of neuroimaging research topics including mild cognitive impairment and depression [28], short term memory in mild cognitive impairment [29], and aphasia [30]. Other authors have extended the general framework to non-human studies [31, 32].

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 guides 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 [33], and extension to non-human data.

2 Methods and materials

2.1 ADNI-1 imaging data

The strict protocol design, large-scale recruitment, and public availability of the Alzheimer's Disease Neuroimaging Initiative (ADNI) makes it an ideal data set for evaluating the ANTs longitudinal cor-

bioRxiv preprint doi: https://doi.org/10.1101/170209; this version posted August 16, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

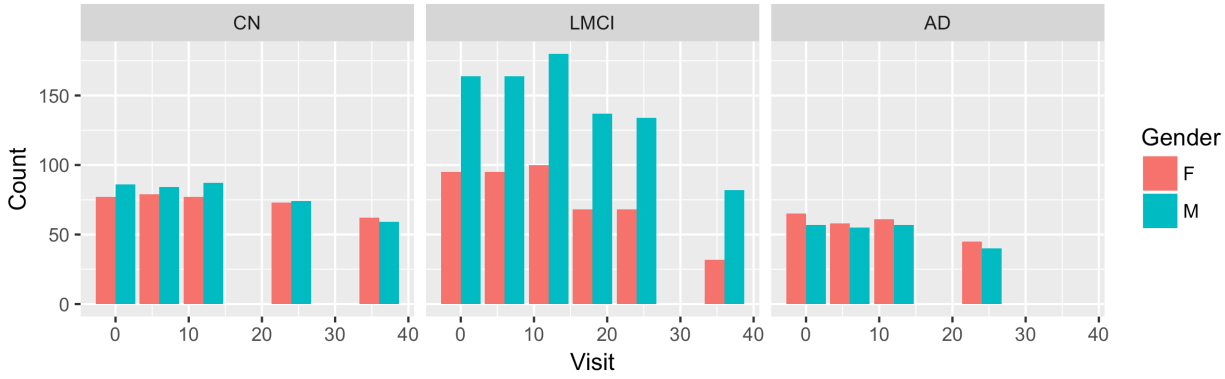

Figure 1: Demographic breakdown of the number of ADNI-1 subjects by diagnosis i.e., normal, mild cognitive impairment (MCI), late mild cognitive impairment (LMCI), and Alzheimer's disease (AD). Within each panel we plot the number of subjects (by gender) per visit—baseline ("bl") and n months ("mn").

tical thickness pipeline. An MP-RAGE [34] sequence for 1.5 and 3.0 T was used to collect the data at the scan sites. Specific acquisition parameters for 1.5 T and 3.0 T magnets are given in Table 1 of [35]. As proposed, collection goals were 200 elderly cognitively normal subjects collected at 0, 6, 12, 24, and 36 months; 400 MCI subjects at risk for AD conversion at 0, 6, 12, 18, 24, and 36 months; and 200 AD subjects at 0, 6, 12, and 24 months.

The ADNI-1 data were downloaded in May of 2014 and first processed using the ANTs cross-sectional cortical thickness pipeline [26] (4399 total images). Data was then processed using two variants of the ANTs longitudinal stream (described in the next section). In the final set of csv files (which we have made publicly available in the GitHub repository associated with this work [36]), we only included time points for which clinical scores (e.g., MMSE) were available. In total, we included 197 cognitive normals, 324 LMCI subjects, and 142 AD subjects with one or more follow-up image acquisition appointments. Further breakdown of demographic information is given in Figures 1 and 2 to provide additional perspective on the data used for this work.

bioRxiv preprint doi: https://doi.org/10.1101/170209; this version posted August 16, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

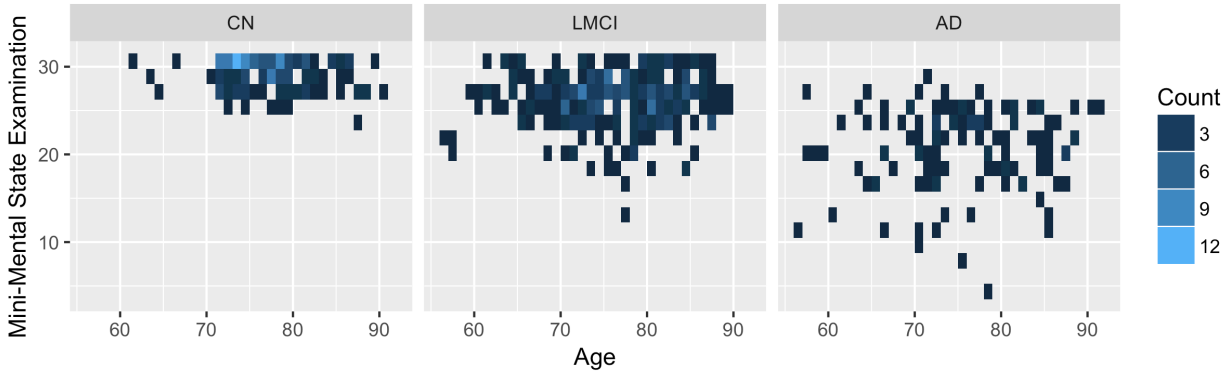

Figure 2: Age vs. Mini-mental examination (MMSE) scores for the ADNI-1 subjects by diagnosis providing additional demographic characterization for the subjects processed for this study.

2.2 ANTs cortical thickness

2.2.1 Cross-sectional processing

A thorough discussion of the ANTs cross-sectional thickness estimation framework was previously provided in [26]. As a brief review, given a T1-weighted brain MR image, processing comprises the following major steps (cf Figure 1 of [26]):

1. preprocessing (e.g., N4 bias correction [37]),
2. brain extraction [38],
3. Atropos n-tissue segmentation [39], and
4. registration-based cortical thickness estimation [8].

ROI-based quantification is achieved through joint label fusion [40] of the cortex coupled with the MindBoggle-101 data. These data use the Desikan–Killiany–Tourville (DKT) labeling protocol [41] to parcellate each cortical hemisphere into 31 anatomical regions (cf Table 1). This pipeline has since been enhanced by the implementation [42] of a patch-based denoising algorithm [43] as an optional preprocessing step and multi-modal integration capabilities (e.g., joint T1- and T2-weighted image processing). All spatial normalizations are generated using the well-known Symmetric Normalization (SyN) image registration algorithm [44, 45] which forms the core of the ANTs toolkit and constitutes the principal component of ANTs-based processing and analysis.

bioRxiv preprint doi: https://doi.org/10.1101/170209; this version posted August 16, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Table 1: The 31 cortical labels (per hemisphere) of the Desikan-Killiany-Tourville atlas. The ROI abbreviations from the R `brainGraph` package are given in parentheses and used in later figures.

1) caudal anterior cingulate (`cACC`)	17) pars orbitalis (`pORB`)
2) caudal middle frontal (`cMFG`)	18) pars triangularis (`pTRI`)
3) cuneus (`CUN`)	19) pericalcarine (`periCAL`)
4) entorhinal (`ENT`)	20) postcentral (`postC`)
5) fusiform (`FUS`)	21) posterior cingulate (`PCC`)
6) inferior parietal (`IPL`)	22) precentral (`preC`)
7) inferior temporal (`ITG`)	23) precuneus (`PCUN`)
8) isthmus cingulate (`iCC`)	24) rosterior anterior cingulate (`rACC`)
9) lateral occipital (`LOG`)	25) rostral middle frontal (`rMFG`)
10) lateral orbitofrontal (`LOF`)	26) superior frontal (`SFG`)
11) lingual (`LING`)	27) superior parietal (`SPL`)
12) medial orbitofrontal (`MOF`)	28) superior temporal (`STG`)
13) middle temporal (`MTG`)	29) supramarginal (`SMAR`)
14) parahippocampal (`PARH`)	30) transverse temporal (`TT`)
15) paracentral (`paraC`)	31) insula (`INS`)
16) pars opercularis (`pOPER`)	

For evaluation, voxelwise regional thickness statistics were summarized based on the DKT parcellation scheme. Test-retest error measurements were presented from a 20-cohort subset of both the OASIS [46] and MMRR [47] data sets and compared with the corresponding FreeSurfer thickness values. Further evaluation employed a training/prediction paradigm where regional cortical thickness values generated from 1205 images taken from four publicly available data sets (i.e., IXI [48], MMRR, NKI [49], and OASIS) were used to predict age and gender using linear and random forest [50] models. The resulting regional statistics (including cortical thickness, surface area [51], volumes, and Jacobian determinant values) were made available online [52]. These include the corresponding FreeSurfer measurements which are also publicly available for research inquiries (e.g., [53]). Since publication, this framework has been used in a number of studies (e.g., [54–56]).

bioRxiv preprint doi: https://doi.org/10.1101/170209; this version posted August 16, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

2.2.2 Unbiased longitudinal processing

Given certain practical limitations (e.g., subject recruitment and retainment), as mentioned earlier, many researchers employ cross-sectional acquisition and processing strategies for studying developmental phenomena. Longitudinal studies, on the other hand, can significantly reduce inter-subject measurement variability. The ANTs longitudinal cortical thickness pipeline extends the ANTs cortical thickness pipeline for longitudinal studies which takes into account various bias issues previously discussed in the literature [12, 14, 19].

Given N time-point T1-weighted MR images (and, possibly, other modalities) and representative images to create a population-specific template and related images, the longitudinal pipeline consists of the following steps:

1. (Offline): Creation of the group template and corresponding prior probability images.
2. Creation of the unbiased single-subject template (SST).
3. Application of the ANTs cross-sectional cortical thickness pipeline [26] to the SST with the group template and priors as input.
4. Creation of the SST prior probability maps.
5. (Optional): Rigid transformation of each individual time point to the SST.
6. Application of the ANTs cross-sectional cortical thickness pipeline [26], with the SST as the reference template, to each individual time-point image. Input includes the SST and the corresponding spatial priors made in Step 3.
7. Joint label fusion to determine the cortical ROIs for analysis.

An overview of these steps is provided in Figure 3 which we describe in greater detail below.

ADNI group template, brain mask, and tissue priors. Prior to any individual subject processing, the group template is constructed from representative population data [57]. For the ADNI-1 processing described in this work, we created a population-specific template from 52 cognitively normal ADNI-1 subjects. Corresponding brain and tissue prior probability maps for the CSF, gray matter, white matter, deep gray matter, brain stem, and cerebellum were created as described in [26]. A brief overview of this process is also provided in the section concerning creation of the single-subject template. Canonical views of the ADNI-1 template and corresponding auxiliary images are given in

bioRxiv preprint doi: https://doi.org/10.1101/170209; this version posted August 16, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

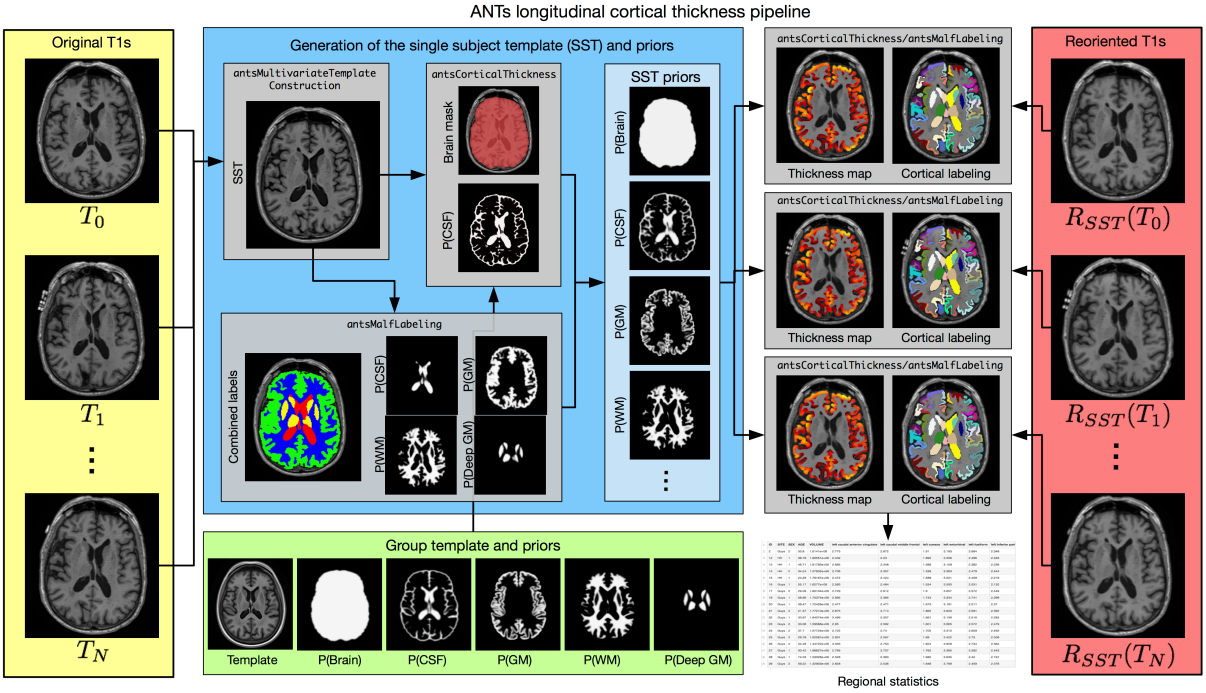

Figure 3: Diagrammatic illustration of the ANTs longitudinal cortical thickness pipeline for a single subject with N time points. From the N original T1-weighted images (left column, yellow panel) and the group template and priors (bottom row, green panel), the single-subject template (SST) and auxiliary prior images are created (center, blue panel). These subject-specific template and other auxiliary images are used to generate the individual time-point cortical thickness maps, in the individual time point's native space (denoted as "ANTs Native" in the text). Optionally, one can rigidly transform the time-point images prior to segmentation and cortical thickness estimation (right column, red panel). This alternative processing scheme is referred to as "ANTs SST". For regional thickness values, regional labels are propagated to each image using a given atlas set (with cortical labels) and joint label fusion.

bioRxiv preprint doi: https://doi.org/10.1101/170209; this version posted August 16, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

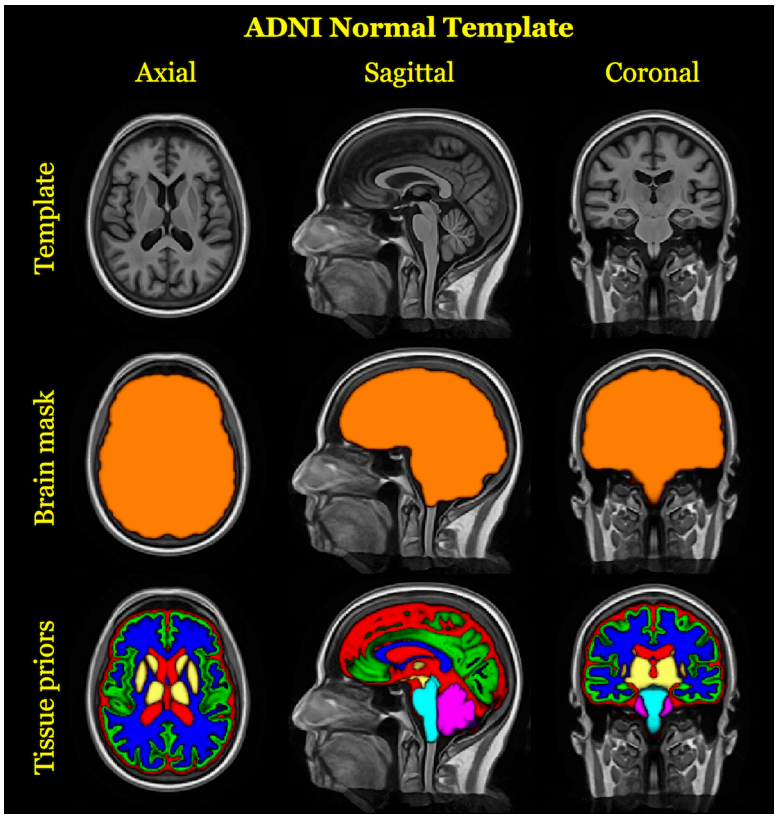

Figure 4: Top row: Canonical views of the template created from 52 randomly selected cognitively normal subjects of the ADNI-1 database. The prior probability mask for the whole brain (middle row) and the six tissue priors (bottom row) are used to "seed" each single-subject template for creation of a probabilistic brain mask and probabilistic tissues priors during longitudinal processing.

bioRxiv preprint doi: https://doi.org/10.1101/170209; this version posted August 16, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Figure 4.

Single-subject template, brain mask, and tissue priors. With the ADNI-1 group template and prior probability images, each subject undergoes identical processing. First, an average shape and intensity single subject template (SST) is created from all time-point images using the same protocol [57] used to produce the ADNI-1 group template. Next, six probabilistic tissue maps (cerebrospinal fluid (CSF), gray matter (GM), white matter (WM), deep gray matter (striatum + thalamus), brain stem, and cerebellum) are generated in the space of the SST. This requires processing the SST through two parallel workflows. First, the SST proceeds through the standard cross-sectional ANTs cortical thickness pipeline which generates a brain extraction mask and the CSF tissue probability map, $P_{Seg}(CSF)$. Second, using a data set of 20 atlases from the OASIS data set that have been expertly annotated and made publicly available [41], a multi-atlas joint label fusion step (JLF) [40] is performed to create individualized probability maps for all six tissue types. Five of the JLF probabilistic tissue estimates (GM, WM, deep GM, brain stem, and cerebellum) and the JLF CSF estimate, $P_{JLF}(CSF)$, are used as the SST prior probabilities after smoothing with a Gaussian kernel (isotropic, $\sigma = 1mm$) whereas the CSF SST tissue probability is derived as a combination of the JLF and segmentation CSF estimates, i.e., $P(CSF) = \max\left(P_{Seg}(CSF), P_{JLF}(CSF)\right)$, also smoothed with the same Gaussian kernel. Finally, $P(CSF)$ is subtracted out from the other five tissue probability maps. Note that the unique treatment of the CSF stems from the fact that the 20 expertly annotated atlases only label the ventricular CSF. Since cortical segmentation accuracy depends on consideration of the external CSF, the above protocol permits such inclusion in the SST CSF prior probability map. The final version of the SST and auxiliary images enable unbiased, non-linear mappings to the group template, subject-specific tissue segmentations, and cortical thickness maps for each time point of the original longitudinal image series.

Individual time point processing. The first step for subject-wise processing involves creating the SST from all the time points for that individual [57]. For the cross-sectional ANTs processing, the group template and auxiliary images are used to perform tasks such as individual brain extraction and n-tissue segmentation prior to cortical thickness estimation [26]. However, in the longitudinal variant, the SST serves this purpose. We thus deformably map the SST and its priors to the native space of each time point where individual-level segmentation and cortical thickness is estimated.

bioRxiv preprint doi: https://doi.org/10.1101/170209; this version posted August 16, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

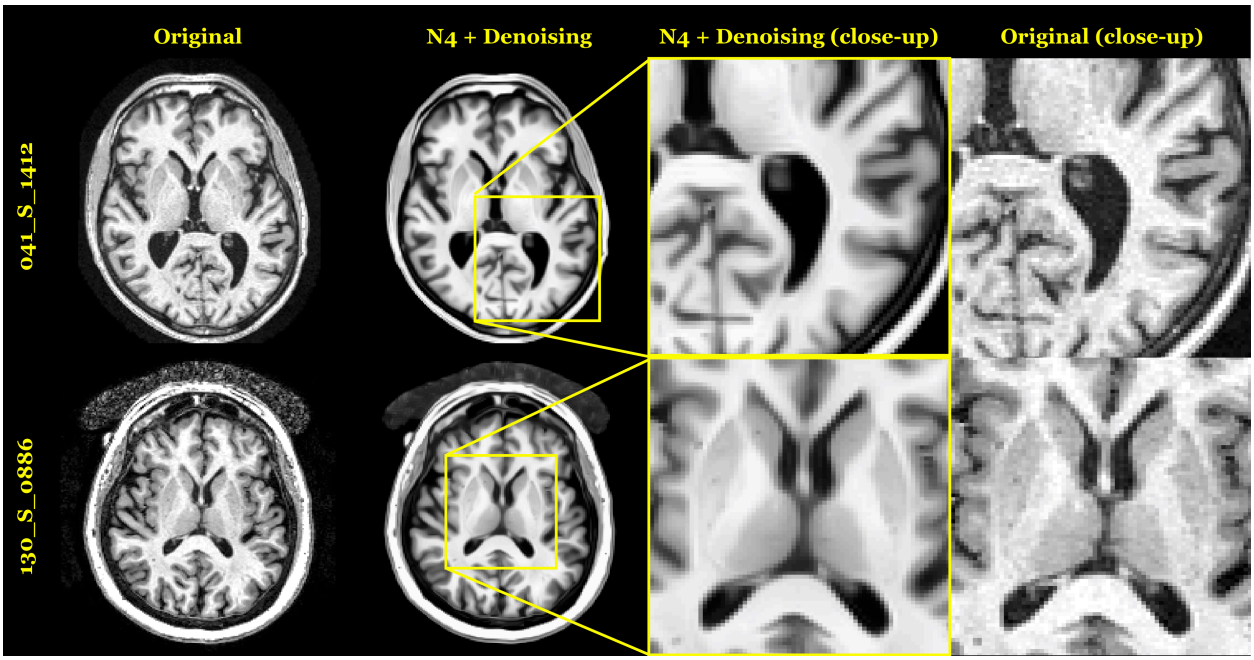

Figure 5: Images from two randomly chosen subjects were chosen to illustrate the effects of bias correction and denoising. The former mitigates artificial low spatial frequency modulation of intensities whereas the latter reduces the high frequency spatially-varying Rician noise characteristic of MRI.

Note that this unbiased longitudinal pipeline is completely agnostic concerning ordering of the input time-point images, i.e., we "treat all time points exactly the same."

An ANTs implementation of the denoising algorithm described in [43] is a recent addition to the toolkit and has been added as an option to both the cross-sectional and longitudinal pipelines. This denoising algorithm employs a non-local means filter [58] to account for the spatial varying noise in MR images in addition to specific consideration of the Rician noise inherent to MRI [59]. This preprocessing step has been used in a variety of imaging studies for enhancing segmentation-based protocols including hippocampal and ventricle segmentation [60], voxel-based morphometry in cannabis users [61], and anterior temporal lobe gray matter volume in bilingual adults [62]. An illustration of the transformed data resulting from image preprocessing (bias correction plus denoising) is provided in Figure 5.

In the FreeSurfer longitudinal stream, each time-point image is processed using the FreeSurfer cross-sectional stream. The resulting processed data from all time points is then used to create a mean, or median, single-subject template. Following template creation, each time-point image is rigidly

bioRxiv preprint doi: https://doi.org/10.1101/170209; this version posted August 16, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

transformed to the template space where it undergoes further processing (e.g., white and pial surface deformation). This reorientation to the template space "further reduce[s] variability" and permits an "implicit vertex correspondence" across all time points [12].

The ANTs framework also permits rotation of the individual time point image data to the SST, similar to FreeSurfer, for reducing variability, minimizing or eliminating possible orientation bias, and permitting a 4-D segmentation given that the Atropos segmentation implementation is dimensionality-agnostic [39]. Regarding the 4-D brain segmentation, any possible benefit is potentially outweighed by the occurrence of "over-regularization" [12] whereby smoothing across time reduces detection ability of large time-point changes. Additionally, it is less than straightforward to accommodate irregular temporal sampling such as the acquisition schedule of the ADNI-1 protocol.

Registration-based cortical thickness. The underlying registration-based estimation of cortical thickness, Diffeomorphic Registration-based Estimation of Cortical Thickness (DiReCT), was introduced in [8]. Given a probabilistic estimate of the cortical gray and white matters, diffeomorphic-based image registration is used to register the white matter probability map to the combined gray/white matters probability map. The resulting mapping defines the diffeomorphic path between a point on the GM/WM interface and the GM/CSF boundary. Cortical thickness values are then assigned at each spatial location within the cortex by integrating along the diffeomorphic path starting at each GM/WM interface point and ending at the GM/CSF boundary. A more detailed explanation is provided in [8] with the actual implementation provided in the class `itk::DiReCTImageFilter` available as part of the ANTs library.

Joint label fusion and pseudo-geodesic for large cohort labeling. Cortical thickness ROI-based analyses are performed using joint label fusion [40] and whatever cortical parcellation scheme is deemed appropriate for the specific study. The brute force application of the joint label fusion algorithm would require N pairwise non-linear registrations for each time-point image where N is the number of atlases used. This would require a significant computational cost for a relatively large study such as ADNI. Instead, we use the "pseudo-geodesic" approach for mapping atlases to individual time point images (e.g., [63]). The transformations between the atlas and the group template are computed offline. With that set of non-linear transforms, we are able to concatenate a set of existing transforms from each atlas through the group template, to the SST, and finally to each individual

bioRxiv preprint doi: https://doi.org/10.1101/170209; this version posted August 16, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

time point for estimating regional labels for each image.

2.3 Statistical evaluation

Based on the above ANTs pipeline descriptions, there are three major variants for cortical thickness processing of longitudinal data. We denote these alternatives as:

- **ANTs Cross-sectional** (or **ANTs Cross**). Process each subject's time point independently using the cross-sectional pipeline originally described in [26].
- **ANTs Longitudinal-SST** (or **ANTs SST**). Rigidly transform each subject to the SST and then segment and estimate cortical thickness in the space of the SST.
- **ANTs Longitudinal-native** (or **ANTs Native**). Segment and estimate cortical thickness in the native space.

For completeness, we also include a comparison with both the cross-section and longitudinal FreeSurfer v5.3 streams respectively denoted as "FreeSurfer Cross-sectional" (or "FS Cross") and "FreeSurfer Longitudinal" (or "FS Long").

2.4 Cross-sectional and longitudinal evaluation strategies

Possible evaluation strategies for cross-sectional methods have employed manual measurements in the histological [64] or virtual [65] domains but would require an inordinate labor effort for collection to be comparable with the size of data sets currently analyzed. Other quantitative measures representing "reliability," "reproducibility," or, more generally, "precision" can also be used to characterize such tools. For example, [66] used FreeSurfer cortical thickness measurements across image acquisition sessions to demonstrate improved reproducibility with the longitudinal stream over the cross-sectional stream. In [67] comparisons for ANTs, FreeSurfer, and the proposed method were made using the range of measurements and their correspondence to values published in the literature. However, none of these precision-type measurements, per se, indicate the utility of a pipeline-specific cortical thickness value as a potential biomarker. For example, Figure 8 in [26] confirms what was found in [67] which is that the range of ANTs cortical thickness values for a particular

bioRxiv preprint doi: https://doi.org/10.1101/170209; this version posted August 16, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

region exceeds those of FreeSurfer. However, for the same data, the demographic predictive capabilities of the former was superior to that of the latter. Thus, better assessment strategies are necessary for determining clinical utility. For example, the intra-class correlation (ICC) coefficient used in [26] demonstrated similarity in both ANTs and FreeSurfer for repeated acquisitions despite the variance discrepancy between both sets of measurements. This is understood with the realization that the ICC takes into account both inter-observer and intra-observer variability.

Similarly, evaluation strategies for longitudinal studies have been proposed with resemblance to those employed for cross-sectional data such as the use of visual assessment [24], scan-rescan data [12, 25], and 2-D comparisons of post mortem images and corresponding MRI [25]. In addition, longitudinal methods offer potential for other types of assessments such as the use of simulated data (e.g., atrophy [12, 25], infant development [24]) where "ground-truth" is known and regression analysis of longitudinal trajectories of cortical thickness [68].

2.5 Cortical residual and between-subject thickness variability

For a longitudinal biomarker to be effective at classifying subpopulations, it should have low residual variation and high between-subject variation. Without these simultaneous conditions, subpopulation distinctions would not be possible (e.g., if measurements within the subject vary more than those between subjects). A summary measure related to the ICC statistic [69] is used to quantify this intuition for assessing relative performance of these cross-sectional and longitudinal ANTs pipeline variants along with the cross-sectional and longitudinal FreeSurfer streams. Specifically, we use longitudinal mixed-effects (LME) modeling to quantify pipeline-specific between-subject and residual variabilities where comparative performance is determined by maximizing the ratio between the former and the latter. Such a quantity implies greater within-subject reproducibility while distinguishing between patient sub-populations. As such this amounts to higher precision when cortical thickness is used as a predictor variable or model covariate in statistical analyses upstream.

LME models comprise a well-established and widely used class of regression models designed to estimate cross-sectional and longitudinal linear associations between quantities while accounting for subject-specific trends. As such, these models are useful for the analysis of longitudinally collected

bioRxiv preprint doi: https://doi.org/10.1101/170209; this version posted August 16, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

cohort data. Indeed, [20] provides an introduction to the mixed-effects methodology in the context of longitudinal neuroimaging data and compare it empirically to competing methods such as repeated measures ANOVA. For more complete treatments of the subject matter, see [69] and [70]. LME models are also useful for estimating and comparing residual and between-subject variability after conditioning out systematic time trends in longitudinally measured data. In the context of the current investigation, by fitting LME models to the data resulting from cross-sectional and longitudinal processing techniques, we are able to quantify the relative performance of each approach with respect to residual, between-subject, and total variability in a way that [71] hint at in their exposition of the longitudinal FreeSurfer stream.

As previously noted we observed a longitudinal sampling of cortical thickness measurements from the 62 parcellated cortical DKT regions. To assess the above variability-based criteria while accounting for changes that may occur through the passage of time, we used a Bayesian LME model for parameter estimation. Let Y_{ij}^k denote the i^{th} individual's cortical thickness measurement corresponding to the k^{th} region of interest at the time point indexed by j. Under the Bayesian paradigm we utilized a model of the form

$$
\begin{aligned}
Y_{ij}^k &\sim N(\alpha_i^k + \beta_i^k t, \sigma_k^2) \qquad (1)\\
\alpha_i^k &\sim N(\alpha_0^k, \tau_k^2) \qquad \beta_i^k \sim N(\beta_0^k, \rho_k^2)\\
\alpha_0^k, \beta_0^k &\sim N(0, 10) \qquad \sigma_k, \tau_k, \rho_k \sim \text{Cauchy}^+(0, 5)
\end{aligned}
$$

where specification of variance priors to half-Cauchy distributions reflects commonly accepted best practice in the context of hierarchical models [72]. These priors concentrate mass near zero but have heavy tails, meaning small variance values are expected but large variance values are not prohibited. Even so, results demonstrated robustness to parameter selection.

In Model (1), τ_k represents the between-subject standard deviation, and σ_k represents the within-subject standard deviation, conditional upon time, and β_k^i denotes the subject-specific slopes of cortical thickness change. For each region k, the quantity of interest is thus the ratio

$$
r^k = \frac{\tau_k}{\sigma_k},\ k = 1, \ldots, 62\,. \qquad (2)
$$

bioRxiv preprint doi: https://doi.org/10.1101/170209; this version posted August 16, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

The posterior distribution of r^k was summarized via the posterior median where the posterior distributions were obtained using the Stan probabilistic programming language [73]. The R interface to Stan was used to calculate the point estimates of Model (1) for cortical thickness across the different pipelines using the default parameters. The csv files containing the regional cortical thickness values for all five pipelines, the Stan model file, and the R script to run the analysis and produce the plots are all located in the GitHub repository created for this work [36].

This ratio is at the heart of classical statistical discrimination methods as it features both in the ANOVA methodology and in Fisher's linear discriminant analysis. These connections are important since the utility of cortical thickness as a biomarker lies in the ability to discriminate between patient sub-populations with respect to clinical outcomes. In particular, [74] (Sections 9.6.2 and 9.6.5) demonstrate the role that randomness and measurement error in explanatory variables play in statistical inference. When the explanatory variable is fixed but measured with error (as is plausible for cortical thickness measurements), the residual variance divided by the between subject variance is proportional to the bias of the estimated linear coefficient when the outcome of interest is regressed over the explanatory variable (Example 9.2). In short, the larger the r^k, the less bias in statistical analyses. When the explanatory variable is considered random and is measured with error (a common assumption in the measurement error literature [75, 76]), this bias is expressed as attenuation of regression coefficient estimates to zero by a multiplicative factor $r^k/(1+r^k)$ (Example 9.3). Thus, larger r^k means less less attenuation bias and hence more discriminative capacity. Note that effect estimator bias is not the only problem—the residual variance is increased by a factor proportional to $r^k/(1+r^k)$ ([74], Chapter 3). The same authors refer to the combination of bias and added variance as a 'double whammy'. Indeed, a worse reliability ratio causes greater bias in multiple linear regression in the presence of collinearity and even biases the estimators for other covariates, progression through time included (cf [76], Section 3.3.1). The same authors state that this bias is typical even in generalized linear models (Section 3.6) and use the ratio as a measure of reliability even in the longitudinal context (Section 11.9).

bioRxiv preprint doi: https://doi.org/10.1101/170209; this version posted August 16, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

2.6 Regional diagnostic contrasts based on cortical atrophy

The variance ratio explored in the previous section is a generic desideratum for statistical assessment of performance over the set of possible use cases. In this section, we narrow the focus to the unique demographical characteristics of the ADNI-1 study data and look at performance of the various pipelines in distinguishing between diagnostic groups on a region-by-region basis. Previous work has explored various aspects of Alzheimer's disease with respect to its spatial distribution and the regional onset of cerebral atrophy. For example, although much work points to the entorhinal cortex as the site for initial deposition of amyloid and tau [77], other evidence points to the basal forebrain as preceding cortical spread [78]. Other considerations include the use of hippocampal atrophy rates as an image-based biomarker of cognitive decline [79], differentiation from other dementia manifestations (e.g., posterior cortical atrophy [80]), and the use of FreeSurfer for monitoring disease progression [81]. Thus, longitudinal measurements have immediate application in Alzheimer's disease research. To showcase the utility of the ANTs framework, we compare the generated longitudinal measurements and their ability to differentiate diagnostic groups (i.e., CN vs. LMCI vs. AD).

Pipeline-specific LME models were constructed for each DKT region relating the change in cortical thickness to diagnosis and other pertinent covariates. In the notation of [82], these regional LME models are defined as:

$$\begin{aligned}\Delta Y^k \sim\ & Y^k_{bl} + AGE_{bl} + ICV_{bl} + APOE + GENDER \\ & + DIAGNOSIS + VISIT : DIAGNOSIS + (1|ID) + (1|SITE)\end{aligned} \tag{3}$$

where ΔY^k is the change in thickness of the k^{th} DKT region from baseline (bl) thickness measurement Y^k_{bl}. We also include random intercepts for both the individual subject (ID) and the acquisition site. Modeling was performed in R using the `lme4` package [83] followed by Tukey post-hoc analyses with false discovery rate (FDR) adjustment using the `multcomp` package in R to test the significance of the "LMCI−CN", "AD−LMCI", and "AD−CN" diagnostic contrasts.

bioRxiv preprint doi: https://doi.org/10.1101/170209; this version posted August 16, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

3 Results

All imaging data were processed through the five processing streams (i.e., FS Cross, FS Long, ANTs Cross, ANTs SST, and ANTs Native) on the high performance computing cluster at the University of California, Irvine (UCI). Based on the evaluation design described in the previous section, we compare pipeline performance when applied to the ADNI-1 data. Specifically, we calculate the variance ratios, described in Section 2.5, for each of the 62 DKT regions for each of the five pipelines. These are compared and discussed. We then explore how this general criterion for evaluating measurement quality applies specifically to a longitudinal analysis of the ADNI-1 data in discriminating the diagnostic stages of Alzheimer's disease.

After processing the image data through the various pipelines, we tabulated the regional thickness values and made them available as .csv files online in the corresonding GitHub repository [36]. We also provide the Perl scripts used to run the pipelines on the UCI cluster and the R scripts used to do the analysis below. Additional figures and plots have also been created which were not included in this work. For example, spaghetti plots showing absolute thickness and longitudinal thickness changes are contained in the subdirectory CrossLong/Data/RegionalThicknessSpaghettiPlots/.

3.1 Cortical residual and between-subject thickness variability

The LME model defined in (1) was used to quantify the between-subject and residual variance with the expectation that maximizing the former while minimizing the latter optimizes measurement quality in terms of prediction and confidence intervals. Figure 6 provides the resulting 95% credible intervals for the distributions of region-specific variance ratios $r^k = \tau_k/\sigma_k$ for each of the five pipelines. Based on the discussion in the previous section, superior methododologies are designated by larger variance ratios.

ANTs SST has the highest ratio variance across most of the 62 regions over the other methods. It rarely overlaps with ANTs Native and never with ANTs Cross. In contrast to the majority of FreeSurfer regional ratio variances (from both FS Cross and FS Long) which are smaller than those of the ANTs pipelines, FS Long has larger ratio values for the EC region with the only overlap in the

bioRxiv preprint doi: https://doi.org/10.1101/170209; this version posted August 16, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

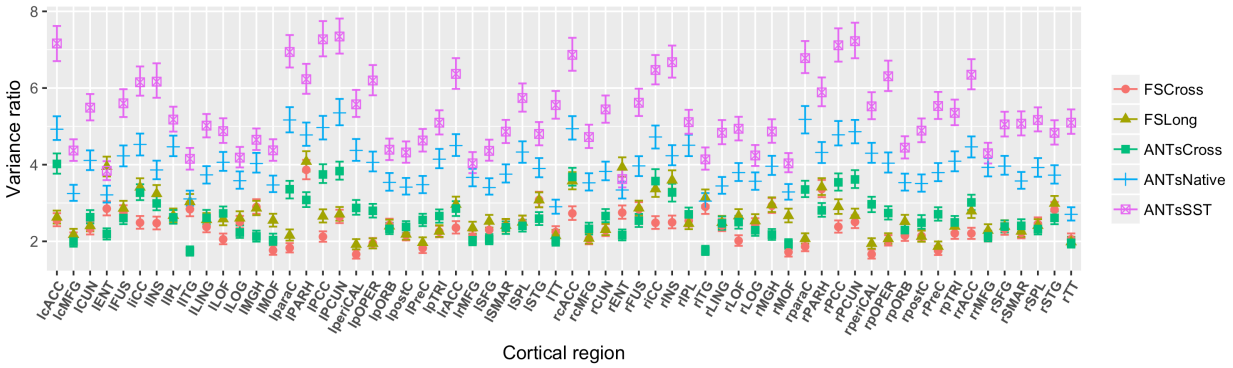

Figure 6: 95% credible intervals of the region-specific variance ratios $r^k = \tau_k/\sigma_k$ are presented for each processing method. The ANTs SST method dominates the others across the majority of regions—its point estimates (posterior medians) are greater than those of the other processing methods except for the left and right EC values in FreeSurfer Long (although there is significant overlap in the credible intervals in those regions). These results also suggest that longitudinal processing is to be preferred for both packages.

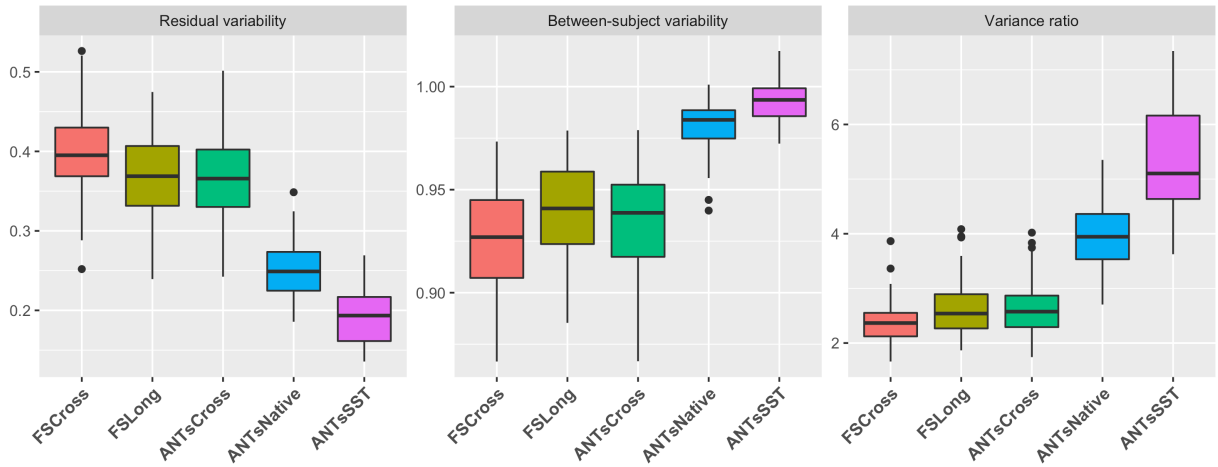

Figure 7: Box plots showing the distribution of the residual variability, between subject variability, and ratio of the between-subject variability and residual variability for each of the 62 DKT regions. Note that the "better" measurement maximizes this latter ratio.

credible intervals with ANTs SST.

The plot in Figure 7 shows a relative summary of all the regional quantities for all three variance measurements (residual, between-subject, and variance ratio) via box plots. These relative distributions show that both between-subject and residual quantities contribute to the disparities in the ratio evaluation metric. Finally, we overlay the variance ratio values on the corresponding regions of a 3-D

bioRxiv preprint doi: https://doi.org/10.1101/170209; this version posted August 16, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

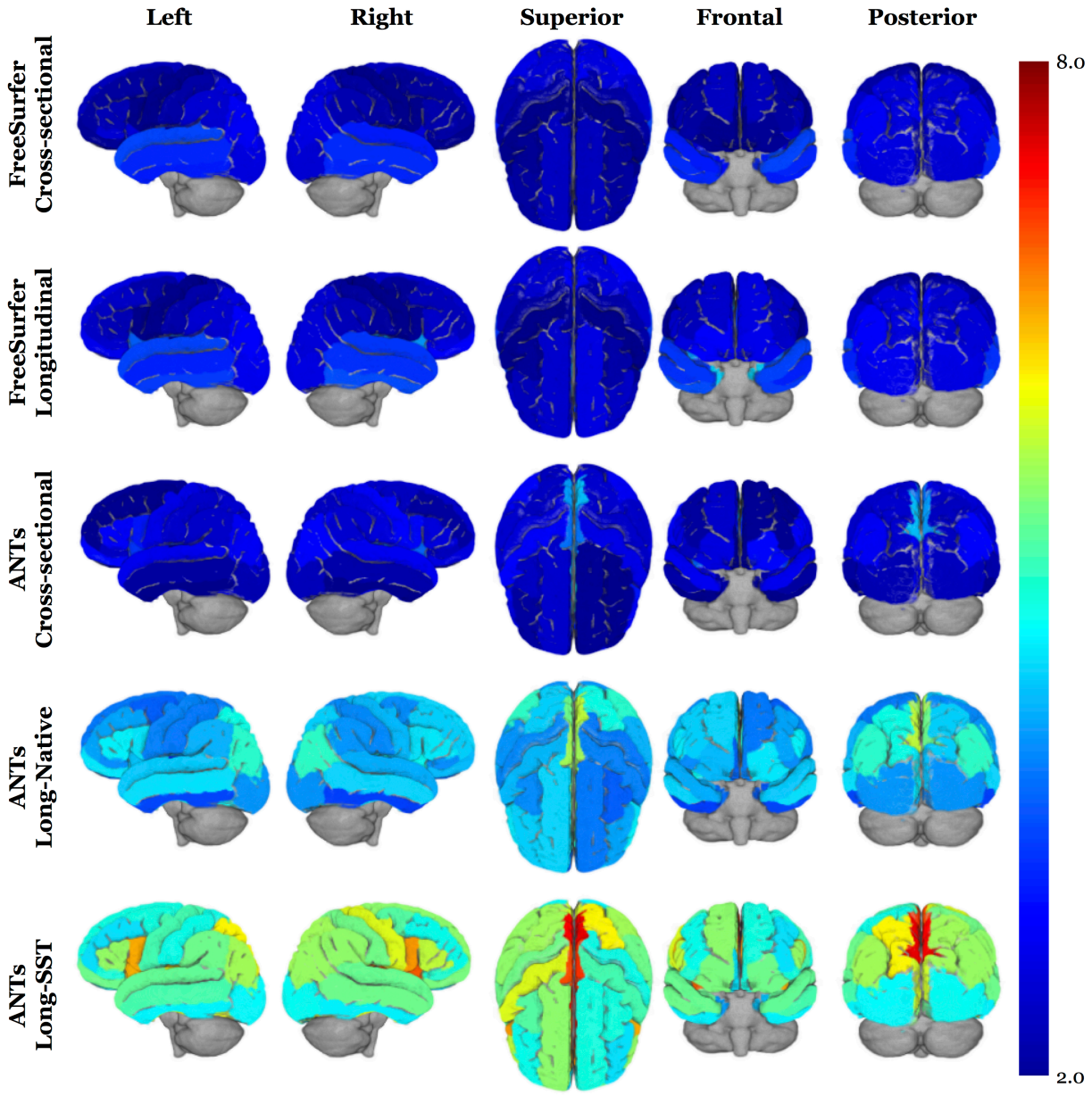

Figure 8: 3-D volumetric rendering of the regional variance ratio values on the generated ADNI template. The higher variance ratios indicate greater between-subject to residual variability.

rendering of the ADNI template (Figure 8) to provide an additional visual comparison between the methods.

bioRxiv preprint doi: https://doi.org/10.1101/170209; this version posted August 16, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Table 2: 95% confidence intervals for the difference in slope values for the three diagnoses (CN, LMCI, AD) of the ADNI-1 data set for each DKT region of the left hemisphere. Each cell is color-coded based on the adjusted log-scaled p-value significance from dark orange (p < 1e-10) to yellow (p = 0.1). Absence of color denotes nonsignificance.

	LMCI−CN					AD−LMCI					AD−CN				
DKT	FSCross	FSLong	ANTsCross	ANTsNative	ANTsSST	FSCross	FSLong	ANTsCross	ANTsNative	ANTsSST	FSCross	FSLong	ANTsCross	ANTsNative	ANTsSST
lcACC	-0.03,0.068	-0.038,0.058	-0.132,0.024	-0.235,-0.065	-0.238,-0.065	-0.075,0.033	-0.08,0.024	-0.103,0.068	-0.187,-0.001	-0.211,-0.022	-0.063,0.058	-0.077,0.041	-0.168,0.024	-0.349,-0.14	-0.375,-0.162
lcMFG	-0.111,-0.043	-0.11,-0.04	-0.197,-0.067	-0.188,-0.063	-0.188,-0.061	-0.109,-0.034	-0.13,-0.053	-0.201,-0.058	-0.21,-0.073	-0.201,-0.062	-0.19,-0.106	-0.21,-0.124	-0.342,-0.182	-0.345,-0.19	-0.334,-0.178
lCUN	-0.042,0.004	-0.046,0.002	-0.049,0.049	-0.073,0.038	-0.08,0.03	-0.03,0.021	-0.034,0.019	-0.108,0	-0.138,-0.016	-0.145,-0.024	-0.053,0.005	-0.059,0	-0.115,0.006	-0.164,-0.026	-0.177,-0.041
lENT	-0.404,-0.24	-0.405,-0.24	-0.285,-0.054	-0.479,-0.219	-0.486,-0.223	-0.359,-0.179	-0.385,-0.204	-0.354,-0.1	-0.462,-0.177	-0.514,-0.226	-0.692,-0.489	-0.719,-0.515	-0.539,-0.254	-0.83,-0.509	-0.887,-0.562
lFUS	-0.129,-0.059	-0.126,-0.054	-0.196,-0.029	-0.308,-0.116	-0.321,-0.128	-0.149,-0.073	-0.161,-0.082	-0.302,-0.117	-0.398,-0.187	-0.418,-0.207	-0.248,-0.162	-0.255,-0.167	-0.425,-0.218	-0.623,-0.385	-0.656,-0.418
lIPL	-0.099,-0.037	-0.099,-0.035	-0.245,-0.069	-0.245,-0.07	-0.246,-0.067	-0.138,-0.07	-0.144,-0.074	-0.357,-0.163	-0.349,-0.157	-0.344,-0.148	-0.21,-0.134	-0.215,-0.136	-0.526,-0.308	-0.52,-0.302	-0.514,-0.292
lITG	-0.146,-0.07	-0.148,-0.072	-0.225,-0.012	-0.359,-0.132	-0.406,-0.165	-0.165,-0.082	-0.17,-0.086	-0.454,-0.218	-0.535,-0.286	-0.574,-0.31	-0.279,-0.185	-0.285,-0.19	-0.586,-0.322	-0.797,-0.516	-0.876,-0.579
liCC	-0.091,-0.02	-0.096,-0.02	-0.17,-0.025	-0.252,-0.089	-0.255,-0.089	-0.117,-0.039	-0.137,-0.054	-0.231,-0.071	-0.307,-0.129	-0.328,-0.147	-0.177,-0.09	-0.2,-0.107	-0.338,-0.159	-0.489,-0.288	-0.512,-0.307
lLOG	-0.065,-0.013	-0.062,-0.01	-0.172,-0.017	-0.18,-0.027	-0.176,-0.021	-0.07,-0.013	-0.074,-0.017	-0.262,-0.091	-0.285,-0.119	-0.304,-0.134	-0.112,-0.049	-0.113,-0.049	-0.367,-0.175	-0.399,-0.211	-0.413,-0.222
lLOF	-0.062,0	-0.071,-0.011	-0.153,0	-0.23,-0.068	-0.236,-0.069	-0.078,-0.01	-0.086,-0.02	-0.217,-0.048	-0.279,-0.101	-0.31,-0.127	-0.114,-0.037	-0.131,-0.057	-0.303,-0.114	-0.439,-0.239	-0.474,-0.269
lLING	-0.041,0.005	-0.041,0.004	-0.097,0.033	-0.148,-0.003	-0.157,-0.015	-0.065,-0.015	-0.07,-0.02	-0.141,0.003	-0.184,-0.024	-0.19,-0.035	-0.086,-0.03	-0.091,-0.035	-0.181,-0.02	-0.269,-0.09	-0.285,-0.111
lMOF	-0.079,-0.011	-0.095,-0.03	-0.181,-0.004	-0.294,-0.114	-0.303,-0.114	-0.08,-0.005	-0.086,-0.015	-0.255,-0.059	-0.326,-0.128	-0.378,-0.171	-0.13,-0.046	-0.153,-0.073	-0.359,-0.139	-0.542,-0.32	-0.599,-0.367
lMGH	-0.147,-0.076	-0.137,-0.066	-0.239,-0.06	-0.32,-0.142	-0.35,-0.161	-0.166,-0.088	-0.167,-0.089	-0.361,-0.162	-0.442,-0.247	-0.471,-0.265	-0.283,-0.194	-0.274,-0.186	-0.523,-0.3	-0.686,-0.465	-0.74,-0.506
lPARH	-0.151,-0.044	-0.137,-0.036	-0.189,-0.022	-0.26,-0.074	-0.265,-0.077	-0.189,-0.071	-0.182,-0.07	-0.301,-0.116	-0.355,-0.152	-0.37,-0.164	-0.294,-0.161	-0.275,-0.15	-0.417,-0.211	-0.536,-0.306	-0.554,-0.322
lparaC	-0.075,-0.011	-0.088,-0.021	-0.075,0.024	-0.087,0.027	-0.087,0.027	-0.037,0.034	-0.068,0.006	-0.111,-0.002	-0.114,0.011	-0.122,0.003	-0.084,-0.005	-0.127,-0.044	-0.143,-0.021	-0.152,-0.011	-0.16,-0.019
lpOPER	-0.07,-0.007	-0.074,-0.011	-0.188,-0.065	-0.222,-0.099	-0.211,-0.091	-0.072,-0.002	-0.08,-0.01	-0.212,-0.076	-0.218,-0.084	-0.212,-0.08	-0.115,-0.037	-0.127,-0.048	-0.346,-0.194	-0.387,-0.236	-0.371,-0.223
lpORB	-0.105,-0.031	-0.099,-0.026	-0.163,0.007	-0.203,-0.041	-0.205,-0.042	-0.053,0.028	-0.067,0.014	-0.225,-0.037	-0.244,-0.066	-0.267,-0.088	-0.126,-0.034	-0.134,-0.044	-0.314,-0.104	-0.377,-0.176	-0.402,-0.2
lpTRI	-0.094,-0.035	-0.099,-0.038	-0.183,-0.029	-0.209,-0.066	-0.191,-0.046	-0.073,-0.008	-0.083,-0.016	-0.217,-0.047	-0.229,-0.072	-0.24,-0.081	-0.142,-0.068	-0.155,-0.08	-0.333,-0.143	-0.377,-0.2	-0.369,-0.189
lperiCAL	-0.019,0.021	-0.031,0.016	-0.06,0.065	-0.079,0.058	-0.095,0.042	-0.028,0.017	-0.044,0.008	-0.168,-0.03	-0.19,-0.04	-0.207,-0.057	-0.029,0.021	-0.055,0.003	-0.174,-0.019	-0.211,-0.041	-0.243,-0.074
lpostC	-0.05,-0.001	-0.052,-0.001	-0.12,-0.03	-0.109,-0.012	-0.11,-0.009	-0.051,0.004	-0.062,-0.005	-0.13,-0.03	-0.14,-0.033	-0.143,-0.031	-0.08,-0.019	-0.092,-0.028	-0.211,-0.1	-0.207,-0.087	-0.209,-0.084
lPCC	-0.058,0.004	-0.068,-0.003	-0.157,-0.026	-0.258,-0.103	-0.267,-0.104	-0.075,-0.006	-0.084,-0.013	-0.177,-0.033	-0.244,-0.074	-0.272,-0.093	-0.106,-0.029	-0.124,-0.044	-0.278,-0.116	-0.436,-0.244	-0.469,-0.267
lPreC	-0.092,-0.023	-0.101,-0.03	-0.145,-0.055	-0.138,-0.036	-0.135,-0.027	-0.077,-0.001	-0.093,-0.013	-0.141,-0.042	-0.145,-0.032	-0.139,-0.021	-0.139,-0.054	-0.162,-0.074	-0.247,-0.136	-0.239,-0.112	-0.228,-0.095
lPCUN	-0.099,-0.042	-0.105,-0.045	-0.17,-0.041	-0.199,-0.05	-0.197,-0.045	-0.091,-0.028	-0.112,-0.046	-0.208,-0.067	-0.241,-0.078	-0.245,-0.08	-0.165,-0.095	-0.191,-0.117	-0.323,-0.164	-0.376,-0.193	-0.378,-0.19
lrACC	-0.088,0.001	-0.095,-0.011	-0.133,0.05	-0.253,-0.05	-0.249,-0.042	-0.054,0.044	-0.054,0.039	-0.151,0.05	-0.266,-0.044	-0.308,-0.083	-0.103,0.007	-0.113,-0.008	-0.205,0.021	-0.431,-0.182	-0.468,-0.214
lrMFG	-0.087,-0.032	-0.095,-0.039	-0.247,-0.054	-0.293,-0.115	-0.297,-0.108	-0.087,-0.027	-0.089,-0.027	-0.292,-0.078	-0.302,-0.106	-0.332,-0.125	-0.151,-0.083	-0.16,-0.091	-0.455,-0.216	-0.519,-0.298	-0.548,-0.314
lSFG	-0.107,-0.049	-0.112,-0.053	-0.197,-0.076	-0.218,-0.093	-0.215,-0.088	-0.086,-0.023	-0.099,-0.033	-0.202,-0.069	-0.234,-0.098	-0.242,-0.103	-0.168,-0.097	-0.186,-0.112	-0.347,-0.197	-0.398,-0.244	-0.403,-0.246
lSPL	-0.086,-0.026	-0.087,-0.023	-0.143,-0.024	-0.105,0.016	-0.098,0.024	-0.08,-0.014	-0.097,-0.026	-0.164,-0.033	-0.146,-0.014	-0.141,-0.008	-0.14,-0.066	-0.156,-0.077	-0.256,-0.109	-0.199,-0.05	-0.187,-0.037
lSTG	-0.137,-0.069	-0.132,-0.066	-0.201,-0.074	-0.228,-0.105	-0.23,-0.105	-0.132,-0.057	-0.133,-0.06	-0.244,-0.104	-0.289,-0.153	-0.297,-0.161	-0.239,-0.155	-0.237,-0.155	-0.39,-0.233	-0.464,-0.312	-0.474,-0.319
lSMAR	-0.09,-0.026	-0.091,-0.027	-0.209,-0.074	-0.194,-0.064	-0.195,-0.058	-0.109,-0.038	-0.122,-0.051	-0.234,-0.086	-0.236,-0.094	-0.24,-0.09	-0.171,-0.092	-0.185,-0.106	-0.385,-0.219	-0.374,-0.214	-0.376,-0.208
lTT	-0.09,0.003	-0.088,0.001	-0.112,-0.002	-0.118,-0.023	-0.103,-0.014	-0.082,0.02	-0.091,0.007	-0.16,-0.038	-0.142,-0.037	-0.137,-0.04	-0.132,-0.016	-0.141,-0.031	-0.224,-0.088	-0.218,-0.101	-0.201,-0.092
lINS	-0.097,-0.023	-0.098,-0.023	-0.208,-0.045	-0.275,-0.108	-0.274,-0.109	-0.089,-0.007	-0.102,-0.02	-0.248,-0.069	-0.349,-0.165	-0.351,-0.171	-0.153,-0.062	-0.167,-0.075	-0.386,-0.185	-0.552,-0.345	-0.554,-0.351

bioRxiv preprint doi: https://doi.org/10.1101/170209; this version posted August 16, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Table 3: 95% confidence intervals for the difference in slope values for the three diagnoses (CN, LMCI, AD) of the ADNI-1 data set for each DKT region of the right hemisphere. Each cell is color-coded based on the adjusted log-scaled p-value significance from dark orange (p < 1e-10) to yellow (p = 0.1). Absence of color denotes nonsignificance.

	LMCI−CN					AD−LMCI					AD−CN				
DKT	FSCross	FSLong	ANTsCross	ANTsNative	ANTsSST	FSCross	FSLong	ANTsCross	ANTsNative	ANTsSST	FSCross	FSLong	ANTsCross	ANTsNative	ANTsSST
rcACC	-0.058,0.024	-0.048,0.028	-0.138,0.004	-0.214,-0.066	-0.222,-0.072	-0.088,0.002	-0.076,0.008	-0.141,0.015	-0.203,-0.041	-0.214,-0.05	-0.11,-0.009	-0.091,0.003	-0.217,-0.043	-0.353,-0.171	-0.371,-0.187
rcMFG	-0.117,-0.048	-0.114,-0.042	-0.184,-0.059	-0.178,-0.053	-0.178,-0.05	-0.118,-0.041	-0.14,-0.061	-0.248,-0.11	-0.251,-0.114	-0.241,-0.101	-0.205,-0.119	-0.223,-0.134	-0.377,-0.223	-0.376,-0.221	-0.364,-0.206
rCUN	-0.025,0.022	-0.03,0.019	-0.059,0.042	-0.085,0.032	-0.094,0.021	-0.057,-0.006	-0.062,-0.008	-0.12,-0.008	-0.133,-0.004	-0.14,-0.015	-0.062,-0.004	-0.071,-0.01	-0.135,-0.01	-0.167,-0.023	-0.184,-0.044
rENT	-0.408,-0.254	-0.403,-0.247	-0.271,-0.026	-0.446,-0.201	-0.438,-0.184	-0.331,-0.162	-0.356,-0.187	-0.407,-0.136	-0.511,-0.242	-0.561,-0.282	-0.672,-0.482	-0.692,-0.5	-0.571,-0.269	-0.852,-0.548	-0.89,-0.576
rFUS	-0.114,-0.045	-0.114,-0.043	-0.221,-0.046	-0.337,-0.144	-0.343,-0.15	-0.149,-0.073	-0.165,-0.086	-0.347,-0.154	-0.44,-0.228	-0.462,-0.252	-0.233,-0.148	-0.249,-0.16	-0.492,-0.276	-0.694,-0.455	-0.722,-0.485
rIPL	-0.107,-0.042	-0.113,-0.045	-0.201,-0.025	-0.228,-0.057	-0.229,-0.051	-0.15,-0.079	-0.16,-0.086	-0.366,-0.172	-0.356,-0.169	-0.358,-0.164	-0.229,-0.149	-0.243,-0.16	-0.491,-0.274	-0.51,-0.299	-0.512,-0.291
rITG	-0.162,-0.086	-0.14,-0.065	-0.281,-0.064	-0.408,-0.19	-0.44,-0.208	-0.179,-0.096	-0.193,-0.111	-0.432,-0.191	-0.539,-0.299	-0.591,-0.337	-0.308,-0.215	-0.301,-0.208	-0.618,-0.35	-0.853,-0.583	-0.931,-0.645
riCC	-0.08,-0.007	-0.084,-0.005	-0.182,-0.044	-0.256,-0.097	-0.26,-0.099	-0.116,-0.035	-0.139,-0.053	-0.239,-0.087	-0.319,-0.145	-0.335,-0.16	-0.164,-0.073	-0.189,-0.092	-0.361,-0.19	-0.507,-0.31	-0.526,-0.328
rLOG	-0.055,-0.004	-0.055,-0.004	-0.152,0.016	-0.179,-0.018	-0.184,-0.017	-0.073,-0.018	-0.08,-0.024	-0.31,-0.124	-0.325,-0.148	-0.337,-0.155	-0.106,-0.044	-0.113,-0.05	-0.389,-0.181	-0.434,-0.235	-0.449,-0.244
rLOF	-0.081,-0.018	-0.08,-0.022	-0.11,0.036	-0.192,-0.031	-0.193,-0.026	-0.077,-0.008	-0.09,-0.025	-0.195,-0.035	-0.255,-0.078	-0.286,-0.102	-0.13,-0.053	-0.145,-0.072	-0.241,-0.063	-0.377,-0.179	-0.406,-0.201
rLING	-0.05,-0.006	-0.055,-0.01	-0.087,0.05	-0.153,0.001	-0.16,-0.011	-0.063,-0.014	-0.067,-0.017	-0.14,0.012	-0.18,-0.011	-0.196,-0.033	-0.094,-0.039	-0.103,-0.046	-0.168,0.002	-0.266,-0.077	-0.292,-0.109
rMOF	-0.077,-0.012	-0.08,-0.017	-0.158,0.019	-0.258,-0.077	-0.268,-0.076	-0.088,-0.016	-0.096,-0.027	-0.25,-0.054	-0.318,-0.118	-0.37,-0.16	-0.137,-0.057	-0.149,-0.071	-0.331,-0.112	-0.498,-0.273	-0.555,-0.318
rMGH	-0.162,-0.093	-0.153,-0.084	-0.242,-0.044	-0.331,-0.139	-0.36,-0.149	-0.167,-0.09	-0.17,-0.094	-0.399,-0.18	-0.496,-0.286	-0.526,-0.296	-0.299,-0.213	-0.293,-0.208	-0.555,-0.31	-0.745,-0.507	-0.796,-0.535
rPARH	-0.193,-0.068	-0.184,-0.066	-0.178,-0.005	-0.277,-0.09	-0.267,-0.082	-0.242,-0.106	-0.23,-0.102	-0.324,-0.134	-0.386,-0.182	-0.401,-0.199	-0.382,-0.228	-0.364,-0.219	-0.427,-0.214	-0.583,-0.352	-0.589,-0.36
rparaC	-0.073,-0.006	-0.091,-0.02	-0.091,0.01	-0.093,0.022	-0.094,0.023	-0.051,0.023	-0.075,0.005	-0.101,0.009	-0.111,0.016	-0.117,0.011	-0.095,-0.012	-0.135,-0.046	-0.149,-0.024	-0.154,-0.011	-0.16,-0.016
rpOPER	-0.078,-0.015	-0.078,-0.016	-0.175,-0.05	-0.195,-0.071	-0.192,-0.07	-0.07,0	-0.084,-0.015	-0.188,-0.05	-0.219,-0.083	-0.217,-0.084	-0.12,-0.042	-0.135,-0.058	-0.309,-0.155	-0.36,-0.208	-0.356,-0.207
rpORB	-0.068,0.008	-0.071,0.001	-0.159,-0.006	-0.174,-0.028	-0.149,-0.002	-0.097,-0.013	-0.105,-0.026	-0.237,-0.067	-0.263,-0.103	-0.274,-0.113	-0.132,-0.038	-0.145,-0.056	-0.33,-0.14	-0.373,-0.193	-0.359,-0.179
rpTRI	-0.085,-0.025	-0.083,-0.022	-0.19,-0.044	-0.191,-0.057	-0.184,-0.049	-0.069,-0.002	-0.083,-0.016	-0.232,-0.072	-0.238,-0.092	-0.25,-0.103	-0.128,-0.053	-0.14,-0.064	-0.359,-0.179	-0.372,-0.207	-0.376,-0.21
rperiCAL	-0.022,0.02	-0.035,0.011	-0.052,0.065	-0.09,0.048	-0.11,0.027	-0.032,0.014	-0.046,0.005	-0.124,0.005	-0.16,-0.009	-0.18,-0.03	-0.036,0.016	-0.061,-0.004	-0.125,0.019	-0.19,-0.02	-0.231,-0.062
rpostC	-0.058,-0.009	-0.06,-0.01	-0.118,-0.028	-0.106,-0.007	-0.106,0.002	-0.072,-0.018	-0.079,-0.024	-0.131,-0.032	-0.15,-0.042	-0.165,-0.047	-0.109,-0.048	-0.117,-0.055	-0.21,-0.099	-0.214,-0.092	-0.224,-0.092
rPCC	-0.055,0.005	-0.068,-0.007	-0.153,-0.03	-0.233,-0.085	-0.243,-0.09	-0.068,-0.002	-0.081,-0.013	-0.135,0	-0.202,-0.04	-0.225,-0.058	-0.097,-0.022	-0.123,-0.047	-0.235,-0.083	-0.372,-0.189	-0.402,-0.213
rPreC	-0.094,-0.025	-0.1,-0.029	-0.12,-0.032	-0.113,-0.011	-0.11,0	-0.08,-0.003	-0.1,-0.021	-0.165,-0.068	-0.186,-0.073	-0.196,-0.075	-0.144,-0.057	-0.169,-0.081	-0.247,-0.138	-0.254,-0.128	-0.258,-0.122
rPCUN	-0.102,-0.038	-0.111,-0.044	-0.184,-0.042	-0.21,-0.048	-0.206,-0.042	-0.096,-0.026	-0.119,-0.045	-0.239,-0.084	-0.273,-0.097	-0.285,-0.107	-0.17,-0.092	-0.201,-0.118	-0.362,-0.187	-0.414,-0.214	-0.421,-0.219
rrACC	-0.062,0.024	-0.069,0.012	-0.21,-0.035	-0.336,-0.137	-0.352,-0.149	-0.096,-0.001	-0.096,-0.006	-0.231,-0.039	-0.324,-0.106	-0.362,-0.14	-0.121,-0.014	-0.13,-0.029	-0.365,-0.149	-0.574,-0.329	-0.627,-0.376
rrMFG	-0.096,-0.038	-0.097,-0.037	-0.203,-0.005	-0.238,-0.059	-0.237,-0.049	-0.09,-0.025	-0.1,-0.033	-0.373,-0.154	-0.359,-0.163	-0.381,-0.176	-0.161,-0.089	-0.171,-0.096	-0.49,-0.245	-0.52,-0.299	-0.538,-0.306
rSFG	-0.102,-0.041	-0.107,-0.043	-0.185,-0.062	-0.203,-0.079	-0.198,-0.073	-0.101,-0.033	-0.115,-0.044	-0.217,-0.081	-0.231,-0.095	-0.242,-0.104	-0.177,-0.101	-0.195,-0.115	-0.348,-0.196	-0.38,-0.227	-0.386,-0.231
rSPL	-0.093,-0.031	-0.098,-0.032	-0.141,-0.011	-0.116,0.02	-0.105,0.034	-0.093,-0.024	-0.106,-0.034	-0.207,-0.064	-0.18,-0.03	-0.171,-0.019	-0.158,-0.082	-0.175,-0.094	-0.291,-0.131	-0.237,-0.069	-0.216,-0.045
rSTG	-0.139,-0.071	-0.131,-0.065	-0.184,-0.04	-0.22,-0.087	-0.232,-0.091	-0.132,-0.056	-0.137,-0.064	-0.269,-0.111	-0.295,-0.15	-0.31,-0.155	-0.241,-0.157	-0.24,-0.158	-0.391,-0.213	-0.458,-0.294	-0.482,-0.307
rSMAR	-0.111,-0.049	-0.11,-0.048	-0.214,-0.069	-0.195,-0.061	-0.186,-0.041	-0.117,-0.049	-0.136,-0.067	-0.248,-0.089	-0.27,-0.123	-0.289,-0.131	-0.201,-0.124	-0.22,-0.142	-0.399,-0.221	-0.407,-0.241	-0.412,-0.234
rTT	-0.105,-0.017	-0.1,-0.015	-0.107,0.013	-0.11,-0.006	-0.108,-0.008	-0.105,-0.007	-0.104,-0.01	-0.225,-0.09	-0.192,-0.077	-0.169,-0.06	-0.171,-0.062	-0.167,-0.062	-0.28,-0.129	-0.258,-0.128	-0.235,-0.111
rINS	-0.079,-0.006	-0.083,-0.014	-0.212,-0.051	-0.271,-0.106	-0.272,-0.109	-0.094,-0.014	-0.098,-0.022	-0.22,-0.043	-0.32,-0.139	-0.323,-0.145	-0.141,-0.052	-0.152,-0.066	-0.362,-0.164	-0.52,-0.316	-0.525,-0.324

bioRxiv preprint doi: https://doi.org/10.1101/170209; this version posted August 16, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

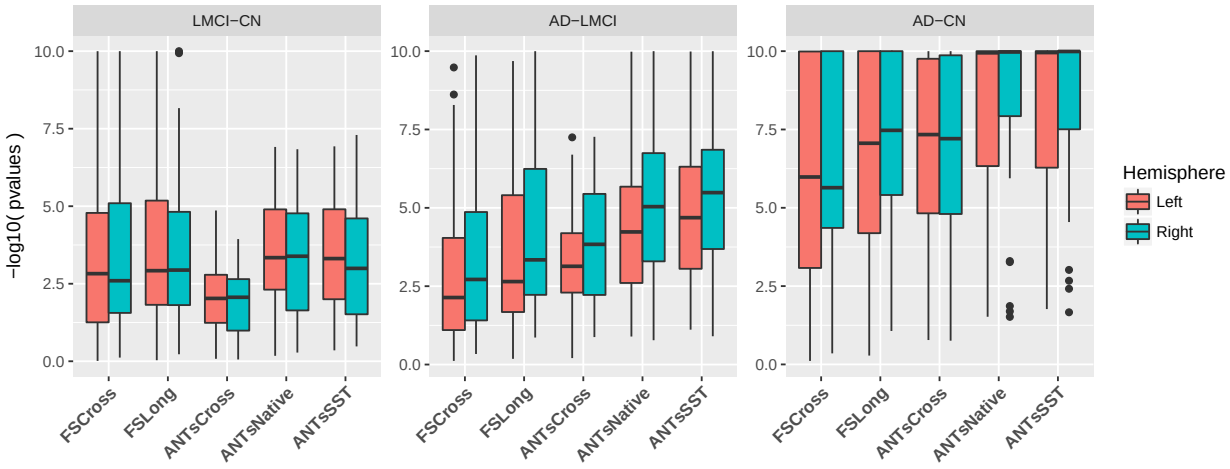

Figure 9: Log-scaled p-values summarizing Tables 2 and 3 demonstrating performance differences across cross-sectional and longitudinal pipelines for the three diagnostic contrasts.

3.2 Regional diagnostic contrasts based on cortical atrophy

The LME model described in Equation (3) was used to determine region-by-region contrasts for each pairing "LMCI−CN", "AD−LMCI", and "AD−CN" using post-hoc Tukey significance testing. It should be noted that no subjects were included that switched diagnostic groups during the acquired study schedule. These findings are provided in Tables 2 and 3. The adjusted p-values were log-scaled for use in specifying the individual color cell for facilitating visual differentiation. Each cell contains the corresponding 95% confidence intervals. Figure 9 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.

Consideration of performance over all three diagnostic pairings illustrates the superiority of the longitudinal ANTs methodologies over their ANTs cross-sectional counterpart. Several regions demonstrating statistically significant non-zero atrophy in ANTs Native and ANTs SST do not manifest similar trends in ANTs Cross (e.g., LMCI-CN: lateral occipital gyri). Pronounced differences between the ANTs longitudinal vs. cross-sectional methodologies can be seen in both the LMCI−CN and AD−CN contrasts. Although ANTs Cross demonstrates discriminative capabilities throughout the cortex and, specifically, in certain AD salient regions, such as the entorhinal and parahippocampal cortices, the contrast is not nearly as strong as the other methods including FS Cross and FS Long thus motivating

bioRxiv preprint doi: https://doi.org/10.1101/170209; this version posted August 16, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

the use of longitudinal considerations for processing of AD data.

Differentiation between the longitudinal methods is not as obvious although trends certainly exist. In general, for differentiating CN vs. LMCI, all methods are comparable except for ANTs Cross. However, for the other two diagnostic contrasts "AD−LMCI" and "AD−CN", the trend is similar to what we found in the evaluation via the variance ratio, viz., the longitudinal ANTs methods tend towards greater contrast means versus ANTs Cross and the two FreeSurfer methods. Looking at specific cortical areas, though, we see that comparable regions ("comparable" in terms of variance ratio) are consistent with previous findings. For example, we noted in the last section that FSLong has a relatively large variance ratio in the entorhinal regions which is consistent with the results seen in Tables 2 and 3.

4 Discussion

Herein we detailed the ANTs registration-based longitudinal cortical thickness framework which is designed to take advantage of longitudinal data acquisition protocols. This framework has been publicly available as open-source in the ANTs GitHub repository for some time. It has been employed in various neuroimaging studies and this work constitutes a formalized exploration of performance for future reference. It inherits the performance capabilities of the ANTs cross-sectional pipeline providing high reliability for large studies, robust registration and segmentation in human lifespan data, and accurate processing in data (human and animal) which exhibit large shape variation. In addition, the ANTs longitudinal pipeline accounts for the various bias issues that have been associated with processing such data. For example, denoising and N4 bias correction mitigate the effects of noise and intensity artifacts across scanners and visits. The use of the single-subject template provides an unbiased subject-specific reference space and a consistent intermediate space for normalization between the group template and individual time points. Undergirding all normalization components is the well-performing SyN registration algorithm which has demonstrated superior performance for a variety of neuroimaging applications (e.g., [63, 84]) and provides accurate correspondence estimation even in the presence of large anatomical variation. Also, given that the entire pipeline is image-based, conversion issues between surface- and voxel-based representations [85] are non-existent which en-

bioRxiv preprint doi: https://doi.org/10.1101/170209; this version posted August 16, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

hances inclusion of other imaging data and employment of other image-specific tools for multi-modal studies (e.g., tensor-based morphometry, longitudinal cortical labeling using joint label fusion and the composition of transformations). All ANTs components are built from the Insight Toolkit which leverages the open-source developer community from academic and industrial institutions leading to a robust (e.g., low failure rate) software platform which can run on a variety of platforms.

With respect to these data and AD in general, the ANTs longitudinal cortical thickness pipelines uses unbiased diffeomorphic registration to provide robust mapping of individual brains to group template space and, simultaneously, high-resolution sensitivity to subtle longitudinal changes over time. Both advantages are relevant to AD. High baseline atrophy levels in AD lead to the need for robustness to large deformations. Sensitivity to subtle longitudinal change over time is particularly relevant to early or preclinical AD studies due to the relatively reduced atrophy rates and smaller difference from control populations. We demonstrate that our approach leads to competitive or superior estimates of annualized atrophy that are biologically plausible in AD populations and that may, in the future, support the use of T1 neuroimaging to detect treatment effects in clinical trials. Furthermore, in ADNI-1, we report a zero percent failure rate with no subject-specific tuning required.

Over 600 subjects from the well-known longitudinal ADNI-1 data set with diagnoses distributed between cognitively normal, LMCI, and AD were processed through the original ANTs cross-sectional framework [26] and two longitudinal variants. One of the variants, ANTs SST, is similar to the FreeSurfer longitudinal stream in that each time-point image is reoriented to an unbiased single-subject template for subsequent processing. ANTs Native, in contrast, estimates cortical thickness in the native space while also using tissue prior probabilities generated from the SST.

Comparative assessment utilized LME models to determine the between-subject to residual variance ratios over the 62 regions of the brain defined by the DKT parcellation scheme where higher values indicate greater generic statistical salience. In these terms, ANTs SST outperformed all other pipeline variants including both the FreeSurfer longitudinal and cross-sectional streams. Regional disparities between the ANTs SST and Native pipelines point to increases in both between-subject and residual variances which might be due to reorientation to a common space similar to other longitudinal strategies.

Further evidence motivating the longitudinal strategies proposed in this work and elsewhere stems

bioRxiv preprint doi: https://doi.org/10.1101/170209; this version posted August 16, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

from the subsequent exploration of differentiating between diagnostic groups using LMEs with the change in cortical thickness as an outcome variable. Almost across the entire cortex, longitudinal strategies (both ANTs and FreeSurfer) outperformed their cross-sectional counterparts in pairwise differentiation of diagnostic groups although these trends varied based on region and diagnosis. In the context of AD, where certain regions have increased saliency in terms of neuroscientific research, and practical considerations might give more weight to certain diagnostic results over others, further exploration is required to tease out these subtle differences and their implications for future research.

One interesting finding was the performance of FS Long in the EC regions where the variance ratios was slightly larger than those of ANTs Long/Native where the credible intervals have significant overlap. Given the small volume and indistinguishability from surrounding structures, segmentation of the EC can be relatively difficult [86]. This segmentation complexity has led to EC-specific [87] and related [88] strategies for targeted regional processing. For this work, we wanted to avoid such tuning and simply employ off-the-shelf input parameters and data. Future work will explore refining input template priors in these problematic regions for ANTs-based estimation of cortical thickness.

These findings promote longitudinal analysis considerations and motivates such techniques over cross-sectional methods for longitudinal data despite the increase in computational costs. While we focus on cortical thickness in this work, there are obvious limitations with the ANTs volume-based framework. Without a direct reconstruction of the cortical surfaces, many important cortical properties (e.g., surface area, cortical folding, sulcal depth, and gyrification) [89] cannot be generated in a straightforward manner. Additional work will want to examine these features more closely working towards a more comprehensive idea of how structure changes. This will help determine the relative importance of such cortical features and will undoubtedly guide future methodological development.

However despite these deficiencies, being inherently voxel-based, the ANTs framework does have advantages not explored in this work but certainly to be utilized in future research. Specifically, the voxel-based input/output processing is conducive to voxel-based analysis strategies (e.g., Eigenanatomy [33]) and straightforward application to non-human research domains. Also, tensor-based morphometric data are directly extracted from the output of the longitudinal processing. And while mesh-based geometric measures are unavailable, digital analogs (e.g., surface area from the digitized Crofton formula [51] and surface curvature [90]) provide a convenient data format

bioRxiv preprint doi: https://doi.org/10.1101/170209; this version posted August 16, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

for integrated data analysis. Finally, given the importance of structural data, such as T1-weighted images, for other types of neuroimaging studies (e.g., resting state fMRI and diffusion tensor imaging), the longitudinal processing stream provides convenient output for facilitating these other types of analyses.

The ANTs longitudinal pipeline provides several additional features that may be worth investigation in future studies. The segmentation approach provides tissue probability maps that may be used in identifying abnormalities of white matter or in voxel-wise studies of gray matter density. The longitudinal formulation of the pipeline is also likely to improve the variance ratio for other transformation-based measurements such as the log-jacobian, often employed in tensor-based morphometry [91]. Local folding and other curvature-based metrics are available, as well, through ANTsR [92]. These quantification tools, individually or jointly, may provide insight into aging and neurodegeneration and will be the subject of future evaluation efforts.

The longitudinal thickness framework is available in script form within the ANTs software library along with the requisite processing components (cf Appendix). All generated data used for input, such as the ADNI template and tissue priors, are available upon request. As previously mentioned, we also make available the csv files containing the regional thickness values for all three pipelines.

bioRxiv preprint doi: https://doi.org/10.1101/170209; this version posted August 16, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

5 Appendix

5.1 Implementation overview

The script `antsLongitudinalCorticalThickness.sh` performs cortical thickness estimation for a longitudinal image series from a single subject. The following principal steps are performed:

1. A single-subject template (SST) is created from all the time point images.
2. The tissue prior probability images are generated for the SST. These six tissues are label 1: CSF, label 2: cortical gray matter, label 3: white matter, label 4: deep gray matter, label 5: brain stem, and label 6: cerebellum. Prior probability creation involves the following steps:
    1. The SST is passed through `antsCorticalThicknes.sh`.
    2. The brain extraction posterior for the SST is created by smoothing the brain extraction mask created during 2a.
    3. If labeled atlases are not provided, we smooth the posteriors from 2.1 to create the SST segmentation priors, otherwise we use the `antsJointFusion` program to create a set of posteriors using the script `antsCookTemplatePriors.sh`.
3. Using the SST + priors, each subject is processed through the `antsCorticalThickness.sh` script.

A typical command line call is:

```
antsLongitudinalCorticalThickness.sh \
              -d ${imageDimension} \
              -e ${brainTemplate} \
              -m ${brainExtractionProbabilityMask} \
              -p ${brainSegmentationPriors}
              -o ${outputPrefix}
              ${anatomicalImages[@]}
```

bioRxiv preprint doi: https://doi.org/10.1101/170209; this version posted August 16, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

5.2 Input parameters

- `imageDimension`: dimensionality of the input images. Can handle 2 or 3 dimensions.
- `brainTemplate`: the group template. We have made several publicly available along with the prior tissue and brain extraction images (https://figshare.com/articles/ANTs_ANTsR_Brain_Templates/915436).
- `brainExtractionProbabilityMask`: prior probability image for the whole brain corresponding to the `brainTemplate`.
- `brainSegmentationPriors`: prior probability images for the six brain tissues mentioned above. These files are specified with the relevant labels, e.g., `prior1.nii.gz`, `prior2.nii.gz`, `prior3.nii.gz`, `prior4.nii.gz`, `prior5.nii.gz`, and `prior6.nii.gz`. The command line argument is specified in C-style formatting, e.g., `prior%d.nii.gz`.
- `anatomicalImages`: the time point images for a single subjects.
- other optional input parameters are available. `antsLongitudinalCorticalThickness -h` provides a listing of the full set of parameters, their descriptions, and other help information.

5.3 Output

In the specified output directory, the following subdirectories are created:

- `${outputPrefix}SST`
- `${outputPrefix}${anatomicalImagesPrefix[0]}`
- `${outputPrefix}${anatomicalImagesPrefix[1]}`
- `${outputPrefix}${anatomicalImagesPrefix[2]}`
- ...

Each subdirectory contains the output of `antsCorticalThickness.sh` applied to the corresponding image. Output consists of the following files:

- `BrainExtractionMask`: Brain extraction mask in subject space.
- `BrainNormalizedToTemplate`: Extracted brain image normalized to the template space.
- `BrainSegmentation0N4`: Input to the segmentation algorithm. It is not brain ex-

bioRxiv preprint doi: https://doi.org/10.1101/170209; this version posted August 16, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

tracted, but is bias-corrected. If multiple images are used for segmentation, there will be `BrainSegmentation1N4` and so on. The brain extracted version of this is `ExtractedBrain0N4`.

- `BrainSegmentation`: Segmentation image, one label per tissue class. The number of classes is determined by the input priors.
- `BrainSegmentationPosteriors1`: Posterior probability of class 1. A similar image is produced for all classes. The numbering scheme matches the input priors.
- `CorticalThickness`: Cortical thickness image in subject space.
- `CorticalThicknessNormalizedToTemplate`: Cortical thickness image in template space.
- `ExtractedBrain0N4`: Brain-extracted version of `BrainSegmentation0N4`.
- `SubjectToTemplate1Warp`, `SubjectToTemplate0GenericAffine.mat`: Transforms to be used when warping images from the subject space to the template space.
- `SubjectToTemplateLogJacobian`: Log of the determinant of the Jacobian, quantifies volume changes in the subject to template warp.
- `TemplateToSubject0Warp`, `TemplateToSubject1GenericAffine.mat`: Transforms to be used when warping images from the template to the subject space.

In addition to these files, the SST subdirectory contains additional warps, suffixed `SubjectToGroupTemplateWarp` and `SubjectToTemplate0GenericAffine.mat`, that can be used to warp each time point image to the group template. These are a combination of the subject to SST warp, and the SST to group template warp. Also included are the SST brain and tissue prior probability images.

bioRxiv preprint doi: https://doi.org/10.1101/170209; this version posted August 16, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Acknowledgments

Additional support to N.T. and M.Y. provided by NIMH R01 MH102392 and NIA R21 AG049220, P50 AG16573.

Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

bioRxiv preprint doi: https://doi.org/10.1101/170209; this version posted August 16, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

References

1. Du, A.-T., Schuff, N., Kramer, J. H., Rosen, H. J., Gorno-Tempini, M. L., Rankin, K., Miller, B. L., and Weiner, M. W. "**Different Regional Patterns of Cortical Thinning in Alzheimer's Disease and Frontotemporal Dementia**" *Brain* 130, no. Pt 4 (2007): 1159–66. doi:10.1093/brain/awm016

2. Dickerson, B. C., Bakkour, A., Salat, D. H., Feczko, E., Pacheco, J., Greve, D. N., Grodstein, F., Wright, C. I., Blacker, D., Rosas, H. D., Sperling, R. A., Atri, A., Growdon, J. H., Hyman, B. T., Morris, J. C., Fischl, B., and Buckner, R. L. "**The Cortical Signature of Alzheimer's Disease: Regionally Specific Cortical Thinning Relates to Symptom Severity in Very Mild to Mild AD Dementia and Is Detectable in Asymptomatic Amyloid-Positive Individuals**" *Cereb Cortex* 19, no. 3 (2009): 497–510. doi:10.1093/cercor/bhn113

3. MacDonald, D., Kabani, N., Avis, D., and Evans, A. C. "**Automated 3-D Extraction of Inner and Outer Surfaces of Cerebral Cortex from MRI**" *Neuroimage* 12, no. 3 (2000): 340–56. doi:10.1006/nimg.1999.0534

4. Magnotta, V. A., Andreasen, N. C., Schultz, S. K., Harris, G., Cizadlo, T., Heckel, D., Nopoulos, P., and Flaum, M. "**Quantitative in Vivo Measurement of Gyrification in the Human Brain: Changes Associated with Aging**" *Cereb Cortex* 9, no. 2 (1999): 151–60.

5. Kim, J. S., Singh, V., Lee, J. K., Lerch, J., Ad-Dab'bagh, Y., MacDonald, D., Lee, J. M., Kim, S. I., and Evans, A. C. "**Automated 3-D Extraction and Evaluation of the Inner and Outer Cortical Surfaces Using a Laplacian Map and Partial Volume Effect Classification**" *Neuroimage* 27, no. 1 (2005): 210–21. doi:10.1016/j.neuroimage.2005.03.036

6. Zeng, X., Staib, L. H., Schultz, R. T., and Duncan, J. S. "**Segmentation and Measurement of the Cortex from 3-D MR Images Using Coupled-Surfaces Propagation**" *IEEE Trans Med Imaging* 18, no. 10 (1999): 927–37. doi:10.1109/42.811276

7. Jones, S. E., Buchbinder, B. R., and Aharon, I. "**Three-Dimensional Mapping of Cortical Thickness Using Laplace's Equation**" *Hum Brain Mapp* 11, no. 1 (2000): 12–32.

8. Das, S. R., Avants, B. B., Grossman, M., and Gee, J. C. "**Registration Based Cortical Thick-**

bioRxiv preprint doi: https://doi.org/10.1101/170209; this version posted August 16, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

ness Measurement" *Neuroimage* 45, no. 3 (2009): 867–79. doi:10.1016/j.neuroimage.2008.12.016

9. Vachet, C., Hazlett, H. C., Niethammer, M., Oguz, I., Cates, J., Whitaker, R., Piven, J., and Styner, M. "**Group-Wise Automatic Mesh-Based Analysis of Cortical Thickness**" *SPIE medical imaging: Image processing* (2011):

10. Kraemer, H. C., Yesavage, J. A., Taylor, J. L., and Kupfer, D. "**How Can We Learn About Developmental Processes from Cross-Sectional Studies, or Can We?**" *Am J Psychiatry* 157, no. 2 (2000): 163–71. doi:10.1176/appi.ajp.157.2.163

11. Weiner, M. W., Veitch, D. P., Aisen, P. S., Beckett, L. A., Cairns, N. J., Green, R. C., Harvey, D., Jack, C. R., Jagust, W., Liu, E., Morris, J. C., Petersen, R. C., Saykin, A. J., Schmidt, M. E., Shaw, L., Siuciak, J. A., Soares, H., Toga, A. W., Trojanowski, J. Q., and, Alzheimer's Disease Neuroimaging Initiative. "**The Alzheimer's Disease Neuroimaging Initiative: A Review of Papers Published Since Its Inception.**" *Alzheimers Dement* 8, no. 1 Suppl (2012): S1–68.

12. Reuter, M., Schmansky, N. J., Rosas, H. D., and Fischl, B. "**Within-Subject Template Estimation for Unbiased Longitudinal Image Analysis**" *Neuroimage* 61, no. 4 (2012): 1402–18. doi:10.1016/j.neuroimage.2012.02.084

13. Smith, S. M., Zhang, Y., Jenkinson, M., Chen, J., Matthews, P. M., Federico, A., and De Stefano, N. "**Accurate, Robust, and Automated Longitudinal and Cross-Sectional Brain Change Analysis**" *Neuroimage* 17, no. 1 (2002): 479–89.

14. Yushkevich, P. A., Avants, B. B., Das, S. R., Pluta, J., Altinay, M., Craige, C., and Alzheimer's Disease Neuroimaging Initiative. "**Bias in Estimation of Hippocampal Atrophy Using Deformation-Based Morphometry Arises from Asymmetric Global Normalization: An Illustration in ADNI 3 T MRI Data**" *Neuroimage* 50, no. 2 (2010): 434–45. doi:10.1016/j.neuroimage.2009.12.007

15. Thompson, W. K., Holland, D., and Alzheimer's Disease Neuroimaging Initiative. "**Bias in Tensor Based Morphometry Stat-ROI Measures May Result in Unrealistic Power Estimates**" *Neuroimage* 57, no. 1 (2011): 1–4. doi:10.1016/j.neuroimage.2010.11.092

16. Avants, B., Cook, P. A., McMillan, C., Grossman, M., Tustison, N. J., Zheng, Y., and Gee, J. C.

bioRxiv preprint doi: https://doi.org/10.1101/170209; this version posted August 16, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

"**Sparse Unbiased Analysis of Anatomical Variance in Longitudinal Imaging**" *Med Image Comput Comput Assist Interv* 13, no. Pt 1 (2010): 324–31.

17. Fox, N. C., Ridgway, G. R., and Schott, J. M. "**Algorithms, Atrophy and Alzheimer's Disease: Cautionary Tales for Clinical Trials**" *Neuroimage* 57, no. 1 (2011): 15–8. doi:10.1016/j.neuroimage.2011.01.077

18. Hua, X., Hibar, D. P., Ching, C. R. K., Boyle, C. P., Rajagopalan, P., Gutman, B. A., Leow, A. D., Toga, A. W., Jack, C. R., Jr, Harvey, D., Weiner, M. W., Thompson, P. M., and Alzheimer's Disease Neuroimaging Initiative. "**Unbiased Tensor-Based Morphometry: Improved Robustness and Sample Size Estimates for Alzheimer's Disease Clinical Trials**" *Neuroimage* 66, (2013): 648–61. doi:10.1016/j.neuroimage.2012.10.086

19. Reuter, M. and Fischl, B. "**Avoiding Asymmetry-Induced Bias in Longitudinal Image Processing**" *Neuroimage* 57, no. 1 (2011): 19–21. doi:10.1016/j.neuroimage.2011.02.076

20. Bernal-Rusiel, J. L., Greve, D. N., Reuter, M., Fischl, B., Sabuncu, M. R., and Alzheimer's Disease Neuroimaging Initiative. "**Statistical Analysis of Longitudinal Neuroimage Data with Linear Mixed Effects Models**" *Neuroimage* 66, (2013): 249–60. doi:10.1016/j.neuroimage.2012.10.065

21. Wierenga, L. M., Langen, M., Oranje, B., and Durston, S. "**Unique Developmental Trajectories of Cortical Thickness and Surface Area**" *Neuroimage* 87, (2014): 120–6. doi:10.1016/j.neuroimage.2013.11.010

22. Landin-Romero, R., Kumfor, F., Leyton, C. E., Irish, M., Hodges, J. R., and Piguet, O. "**Disease-Specific Patterns of Cortical and Subcortical Degeneration in a Longitudinal Study of Alzheimer's Disease and Behavioural-Variant Frontotemporal Dementia**" *Neuroimage* (2016): doi:10.1016/j.neuroimage.2016.03.032

23. Nourbakhsh, B., Azevedo, C., Nunan-Saah, J., Maghzi, A.-H., Spain, R., Pelletier, D., and Waubant, E. "**Longitudinal Associations Between Brain Structural Changes and Fatigue in Early Ms**" *Mult Scler Relat Disord* 5, (2016): 29–33. doi:10.1016/j.msard.2015.10.006

24. Li, G., Nie, J., Wang, L., Shi, F., Gilmore, J. H., Lin, W., and Shen, D. "**Measuring the Dynamic Longitudinal Cortex Development in Infants by Reconstruction of Temporally Consis-**

bioRxiv preprint doi: https://doi.org/10.1101/170209; this version posted August 16, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

tent Cortical Surfaces" *Neuroimage* 90, (2014): 266–79. doi:10.1016/j.neuroimage.2013.12.038

25. Nakamura, K., Fox, R., and Fisher, E. "**CLADA: Cortical Longitudinal Atrophy Detection Algorithm**" *Neuroimage* 54, no. 1 (2011): 278–89. doi:10.1016/j.neuroimage.2010.07.052

26. Tustison, N. J., Cook, P. A., Klein, A., Song, G., Das, S. R., Duda, J. T., Kandel, B. M., Strien, N. van, Stone, J. R., Gee, J. C., and Avants, B. B. "**Large-Scale Evaluation of ANTs and FreeSurfer Cortical Thickness Measurements**" *Neuroimage* 99, (2014): 166–79. doi:10.1016/j.neuroimage.2014.05.044

27. Available at https://github.com/stnava/ANTs

28. Fujishima, M., Maikusa, N., Nakamura, K., Nakatsuka, M., Matsuda, H., and Meguro, K. "**Mild Cognitive Impairment, Poor Episodic Memory, and Late-Life Depression Are Associated with Cerebral Cortical Thinning and Increased White Matter Hyperintensities**" *Front Aging Neurosci* 6, (2014): 306. doi:10.3389/fnagi.2014.00306

29. Das, S. R., Mancuso, L., Olson, I. R., Arnold, S. E., and Wolk, D. A. "**Short-Term Memory Depends on Dissociable Medial Temporal Lobe Regions in Amnestic Mild Cognitive Impairment**" *Cereb Cortex* 26, no. 5 (2016): 2006–17. doi:10.1093/cercor/bhv022

30. Olm, C. A., Kandel, B. M., Avants, B. B., Detre, J. A., Gee, J. C., Grossman, M., and McMillan, C. T. "**Arterial Spin Labeling Perfusion Predicts Longitudinal Decline in Semantic Variant Primary Progressive Aphasia**" *J Neurol* 263, no. 10 (2016): 1927–38. doi:10.1007/s00415-016-8221-1

31. Pagani, M., Damiano, M., Galbusera, A., Tsaftaris, S. A., and Gozzi, A. "**Semi-Automated Registration-Based Anatomical Labelling, Voxel Based Morphometry and Cortical Thickness Mapping of the Mouse Brain.**" *Journal of neuroscience methods* 267, (2016): 62–73. doi:10.1016/j.jneumeth.2016.04.007

32. Majka, P., Chaplin, T. A., Yu, H.-H., Tolpygo, A., Mitra, P. P., Wójcik, D. K., and Rosa, M. G. P. "**Towards a Comprehensive Atlas of Cortical Connections in a Primate Brain: Mapping Tracer Injection Studies of the Common Marmoset into a Reference Digital Template.**" *The Journal of comparative neurology* 524, no. 11 (2016): 2161–2181. doi:10.1002/cne.24023

bioRxiv preprint doi: https://doi.org/10.1101/170209; this version posted August 16, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

33. Kandel, B. M., Wang, D. J. J., Gee, J. C., and Avants, B. B. "**Eigenanatomy: Sparse Dimensionality Reduction for Multi-Modal Medical Image Analysis**" *Methods* 73, (2015): 43–53. doi:10.1016/j.ymeth.2014.10.016

34. Mugler, J. P., 3rd and Brookeman, J. R. "**Three-Dimensional Magnetization-Prepared Rapid Gradient-Echo Imaging (3D MP RAGE)**" *Magn Reson Med* 15, no. 1 (1990): 152–7.

35. Jack, C. R., Jr, Bernstein, M. A., Fox, N. C., Thompson, P., Alexander, G., Harvey, D., Borowski, B., Britson, P. J., L Whitwell, J., Ward, C., Dale, A. M., Felmlee, J. P., Gunter, J. L., Hill, D. L. G., Killiany, R., Schuff, N., Fox-Bosetti, S., Lin, C., Studholme, C., DeCarli, C. S., Krueger, G., Ward, H. A., Metzger, G. J., Scott, K. T., Mallozzi, R., Blezek, D., Levy, J., Debbins, J. P., Fleisher, A. S., Albert, M., Green, R., Bartzokis, G., Glover, G., Mugler, J., and Weiner, M. W. "**The Alzheimer's Disease Neuroimaging Initiative (ADNI): MRI Methods**" *J Magn Reson Imaging* 27, no. 4 (2008): 685–91. doi:10.1002/jmri.21049

36. Available at https://github.com/ntustison/CrossLong

37. Tustison, N. J., Avants, B. B., Cook, P. A., Zheng, Y., Egan, A., Yushkevich, P. A., and Gee, J. C. "**N4ITK: Improved N3 Bias Correction**" *IEEE Trans Med Imaging* 29, no. 6 (2010): 1310–20. doi:10.1109/TMI.2010.2046908

38. Avants, B. B., Klein, A., Tustison, N. J., Woo, J., and Gee, J. C. "**Evaluation of an Open-Access, Automated Brain Extraction Method on Multi-Site Multi-Disorder Data**" (2010):

39. Avants, B. B., Tustison, N. J., Wu, J., Cook, P. A., and Gee, J. C. "**An Open Source Multivariate Framework for n-Tissue Segmentation with Evaluation on Public Data**" *Neuroinformatics* 9, no. 4 (2011): 381–400. doi:10.1007/s12021-011-9109-y

40. Wang, H., Suh, J. W., Das, S. R., Pluta, J. B., Craige, C., and Yushkevich, P. A. "**Multi-Atlas Segmentation with Joint Label Fusion**" *IEEE Trans Pattern Anal Mach Intell* 35, no. 3 (2013): 611–23. doi:10.1109/TPAMI.2012.143

41. Klein, A. and Tourville, J. "**101 Labeled Brain Images and a Consistent Human Cortical Labeling Protocol**" *Front Neurosci* 6, (2012): 171. doi:10.3389/fnins.2012.00171

42. Tustison, N. J. and Herrera, J. M. "**Two Luis Miguel Fans Walk into a Bar in Nagoya —**

bioRxiv preprint doi: https://doi.org/10.1101/170209; this version posted August 16, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

> (Yada, Yada, Yada) —> an ITK-Implementation of a Popular Patch-Based Denoising Filter" *Insight Journal* (2016):

43. Manjón, J. V., Coupé, P., Martí-Bonmatí, L., Collins, D. L., and Robles, M. "**Adaptive Non-Local Means Denoising of Mr Images with Spatially Varying Noise Levels**" *J Magn Reson Imaging* 31, no. 1 (2010): 192–203. doi:10.1002/jmri.22003

44. Avants, B. B., Tustison, N. J., Song, G., Cook, P. A., Klein, A., and Gee, J. C. "**A Reproducible Evaluation of ANTs Similarity Metric Performance in Brain Image Registration**" *Neuroimage* 54, no. 3 (2011): 2033–44. doi:10.1016/j.neuroimage.2010.09.025

45. Avants, B. B., Tustison, N. J., Stauffer, M., Song, G., Wu, B., and Gee, J. C. "**The Insight ToolKit Image Registration Framework**" *Front Neuroinform* 8, (2014): 44. doi:10.3389/fninf.2014.00044

46. Available at http://www.oasis-brains.org

47. Landman, B. A., Huang, A. J., Gifford, A., Vikram, D. S., Lim, I. A. L., Farrell, J. A. D., Bogovic, J. A., Hua, J., Chen, M., Jarso, S., Smith, S. A., Joel, S., Mori, S., Pekar, J. J., Barker, P. B., Prince, J. L., and Zijl, P. C. M. van. "**Multi-Parametric Neuroimaging Reproducibility: A 3-T Resource Study**" *Neuroimage* 54, no. 4 (2011): 2854–66. doi:10.1016/j.neuroimage.2010.11.047

48. Available at http://brain-development.org/ixi-dataset/

49. Available at http://fcon_1000.projects.nitrc.org/indi/pro/nki.html

50. Breiman, L. "**Random Forests**" *Machine learning* (2001): 5–32.

51. Lehmann, G. and Legland, D. "**Efficient N-Dimensional Surface Estimation Using Crofton Formula and Run-Length Encoding**" *Insight Journal* (2012):

52. Available at https://github.com/ntustison/KapowskiChronicles

53. Hasan, K. M., Mwangi, B., Cao, B., Keser, Z., Tustison, N. J., Kochunov, P., Frye, R. E., Savatic, M., and Soares, J. "**Entorhinal Cortex Thickness Across the Human Lifespan**" *J Neuroimaging* 26, no. 3 (2016): 278–82. doi:10.1111/jon.12297

54. Price, A. R., Bonner, M. F., Peelle, J. E., and Grossman, M. "**Converging Evidence for the**

bioRxiv preprint doi: https://doi.org/10.1101/170209; this version posted August 16, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Neuroanatomic Basis of Combinatorial Semantics in the Angular Gyrus" *J Neurosci* 35, no. 7 (2015): 3276–84. doi:10.1523/JNEUROSCI.3446-14.2015

55. Wisse, L. E. M., Butala, N., Das, S. R., Davatzikos, C., Dickerson, B. C., Vaishnavi, S. N., Yushkevich, P. A., Wolk, D. A., and Alzheimer's Disease Neuroimaging Initiative. "**Suspected Non-Ad Pathology in Mild Cognitive Impairment**" *Neurobiol Aging* 36, no. 12 (2015): 3152–62. doi:10.1016/j.neurobiolaging.2015.08.029

56. Betancourt, L. M., Avants, B., Farah, M. J., Brodsky, N. L., Wu, J., Ashtari, M., and Hurt, H. "**Effect of Socioeconomic Status (Ses) Disparity on Neural Development in Female African-American Infants at Age 1 Month**" *Dev Sci* (2015): doi:10.1111/desc.12344

57. Avants, B. B., Yushkevich, P., Pluta, J., Minkoff, D., Korczykowski, M., Detre, J., and Gee, J. C. "**The Optimal Template Effect in Hippocampus Studies of Diseased Populations**" *Neuroimage* 49, no. 3 (2010): 2457–66. doi:10.1016/j.neuroimage.2009.09.062

58. Buades, A. "**A Non-Local Algorithm for Image Denoising**" *Computer Vision and Pattern Recognition* 2, (2005): 60–65.

59. Gudbjartsson, H. and Patz, S. "**The Rician Distribution of Noisy Mri Data**" *Magn Reson Med* 34, no. 6 (1995): 910–4.

60. Coupé, P., Manjón, J. V., Fonov, V., Pruessner, J., Robles, M., and Collins, D. L. "**Patch-Based Segmentation Using Expert Priors: Application to Hippocampus and Ventricle Segmentation**" *Neuroimage* 54, no. 2 (2011): 940–54. doi:10.1016/j.neuroimage.2010.09.018

61. Cousijn, J., Wiers, R. W., Ridderinkhof, K. R., Brink, W. van den, Veltman, D. J., and Goudriaan, A. E. "**Grey Matter Alterations Associated with Cannabis Use: Results of a Vbm Study in Heavy Cannabis Users and Healthy Controls**" *Neuroimage* 59, no. 4 (2012): 3845–51. doi:10.1016/j.neuroimage.2011.09.046

62. Abutalebi, J., Canini, M., Della Rosa, P. A., Sheung, L. P., Green, D. W., and Weekes, B. S. "**Bilingualism Protects Anterior Temporal Lobe Integrity in Aging**" *Neurobiol Aging* 35, no. 9 (2014): 2126–33. doi:10.1016/j.neurobiolaging.2014.03.010

63. Tustison, N. J., Shrinidhi, K. L., Wintermark, M., Durst, C. R., Kandel, B. M., Gee, J. C., Gross-

bioRxiv preprint doi: https://doi.org/10.1101/170209; this version posted August 16, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

man, M. C., and Avants, B. B. "**Optimal Symmetric Multimodal Templates and Concatenated Random Forests for Supervised Brain Tumor Segmentation (Simplified) with Antsr**" *Neuroinformatics* 13, no. 2 (2015): 209–25. doi:10.1007/s12021-014-9245-2

64. Rosas, H. D., Liu, A. K., Hersch, S., Glessner, M., Ferrante, R. J., Salat, D. H., Kouwe, A. van der, Jenkins, B. G., Dale, A. M., and Fischl, B. "**Regional and Progressive Thinning of the Cortical Ribbon in Huntington's Disease**" *Neurology* 58, no. 5 (2002): 695–701.

65. Kuperberg, G. R., Broome, M. R., McGuire, P. K., David, A. S., Eddy, M., Ozawa, F., Goff, D., West, W. C., Williams, S. C. R., Kouwe, A. J. W. van der, Salat, D. H., Dale, A. M., and Fischl, B. "**Regionally Localized Thinning of the Cerebral Cortex in Schizophrenia**" *Arch Gen Psychiatry* 60, no. 9 (2003): 878–88. doi:10.1001/archpsyc.60.9.878

66. Jovicich, J., Marizzoni, M., Sala-Llonch, R., Bosch, B., Bartrés-Faz, D., Arnold, J., Benninghoff, J., Wiltfang, J., Roccatagliata, L., Nobili, F., Hensch, T., Tränkner, A., Schönknecht, P., Leroy, M., Lopes, R., Bordet, R., Chanoine, V., Ranjeva, J.-P., Didic, M., Gros-Dagnac, H., Payoux, P., Zoccatelli, G., Alessandrini, F., Beltramello, A., Bargalló, N., Blin, O., Frisoni, G. B., and The Pharma-Cog Consortium. "**Brain Morphometry Reproducibility in Multi-Center 3T MRI Studies: A Comparison of Cross-Sectional and Longitudinal Segmentations**" *Neuroimage* (2013): doi:10.1016/j.neuroimage.2013.05.007

67. Klein, A., Ghosh, S. S., Bao, F. S., Giard, J., Häme, Y., Stavsky, E., Lee, N., Rossa, B., Reuter, M., Chaibub Neto, E., and Keshavan, A. "**Mindboggling Morphometry of Human Brains**" *PLoS Comput Biol* 13, no. 2 (2017): e1005350. doi:10.1371/journal.pcbi.1005350

68. Li, G., Nie, J., Wu, G., Wang, Y., Shen, D., and Alzheimer's Disease Neuroimaging Initiative. "**Consistent Reconstruction of Cortical Surfaces from Longitudinal Brain Mr Images**" *Neuroimage* 59, no. 4 (2012): 3805–20. doi:10.1016/j.neuroimage.2011.11.012

69. Verbeke, G. and Molenberghs, G. "**Linear Mixed Models for Longitudinal Data**" (2009):

70. Fitzmaurice, G. M., Laird, N. M., and Ware, J. H. "**Applied Longitudinal Analysis**" 998, (2012):

71. Reuter, M., Schmansky, N. J., Rosas, H. D., and Fischl, B. "**Within-Subject Template Esti-**

bioRxiv preprint doi: https://doi.org/10.1101/170209; this version posted August 16, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

mation for Unbiased Longitudinal Image Analysis" *Neuroimage* 61, no. 4 (2012): 1402–18. doi:10.1016/j.neuroimage.2012.02.084

72. Gelman, A. and others. "**Prior Distributions for Variance Parameters in Hierarchical Models (Comment on Article by Browne and Draper)**" *Bayesian analysis* 1, no. 3 (2006): 515–534.

73. Carpenter, B., Gelman, A., Hoffman, M., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M. A., Guo, J., Li, P., and Riddell, A. "**Stan: A Probabilistic Programming Language**" *J Stat Softw* (2016):

74. Seber, G. A. and Lee, A. J. "**Linear Regression Analysis**" 329, (2012):

75. Fuller, W. A. "**Measurement Error Models**" 305, (2009):

76. Carroll, R. J., Ruppert, D., Stefanski, L. A., and Crainiceanu, C. M. "**Measurement Error in Nonlinear Models: A Modern Perspective**" (2006):

77. Yassa, M. A. "**Ground Zero in Alzheimer's Disease**" *Nat Neurosci* 17, no. 2 (2014): 146–7. doi:10.1038/nn.3631

78. Schmitz, T. W., Nathan Spreng, R., and Alzheimer's Disease Neuroimaging Initiative. "**Basal Forebrain Degeneration Precedes and Predicts the Cortical Spread of Alzheimer's Pathology**" *Nat Commun* 7, (2016): 13249. doi:10.1038/ncomms13249

79. Andrews, K. A., Frost, C., Modat, M., Cardoso, M. J., AIBL, Rowe, C. C., Villemagne, V., Fox, N. C., Ourselin, S., and Schott, J. M. "**Acceleration of Hippocampal Atrophy Rates in Asymptomatic Amyloidosis**" *Neurobiol Aging* 39, (2016): 99–107. doi:10.1016/j.neurobiolaging.2015.10.013

80. Crutch, S. J., Schott, J. M., Rabinovici, G. D., Murray, M., Snowden, J. S., Flier, W. M. van der, Dickerson, B. C., Vandenberghe, R., Ahmed, S., Bak, T. H., Boeve, B. F., Butler, C., Cappa, S. F., Ceccaldi, M., Souza, L. C. de, Dubois, B., Felician, O., Galasko, D., Graff-Radford, J., Graff-Radford, N. R., Hof, P. R., Krolak-Salmon, P., Lehmann, M., Magnin, E., Mendez, M. F., Nestor, P. J., Onyike, C. U., Pelak, V. S., Pijnenburg, Y., Primativo, S., Rossor, M. N., Ryan, N. S., Scheltens, P., Shakespeare, T. J., Suárez González, A., Tang-Wai, D. F., Yong, K. X. X., Carrillo, M., Fox, N. C., and Alzheimer's Association ISTAART Atypical Alzheimer's Disease and Associated Syndromes Professional Interest

bioRxiv preprint doi: https://doi.org/10.1101/170209; this version posted August 16, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Area. "**Consensus Classification of Posterior Cortical Atrophy**" *Alzheimers Dement* 13, no. 8 (2017): 870–884. doi:10.1016/j.jalz.2017.01.014

81. Falahati, F., Ferreira, D., Muehlboeck, J.-S., Eriksdotter, M., Simmons, A., Wahlund, L.-O., and Westman, E. "**Monitoring Disease Progression in Mild Cognitive Impairment: Associations Between Atrophy Patterns, Cognition, Apoe and Amyloid**" *Neuroimage Clin* 16, (2017): 418–428. doi:10.1016/j.nicl.2017.08.014

82. Wilkinson, G. N. and Rogers, C. E. "**Symbolic Description of Factorial Models for Analysis of Variance**" *Journal of the Royal Statistical Society. Series C (Applied Statistics)* 22, no. 3 (1973): 392–399.

83. Bates, D., Mächler, M., Bolker, B., and Walker, S. "**Fitting Linear Mixed-Effects Models Using Lme4**" *Journal of Statistical Software, Articles* 67, no. 1 (2015): 1–48. doi:10.18637/jss.v067.i01, Available at https://www.jstatsoft.org/v067/i01

84. Klein, A., Andersson, J., Ardekani, B. A., Ashburner, J., Avants, B., Chiang, M.-C., Christensen, G. E., Collins, D. L., Gee, J., Hellier, P., Song, J. H., Jenkinson, M., Lepage, C., Rueckert, D., Thompson, P., Vercauteren, T., Woods, R. P., Mann, J. J., and Parsey, R. V. "**Evaluation of 14 Nonlinear Deformation Algorithms Applied to Human Brain Mri Registration**" *Neuroimage* 46, no. 3 (2009): 786–802. doi:10.1016/j.neuroimage.2008.12.037

85. Klein, A., Ghosh, S. S., Avants, B., Yeo, B. T. T., Fischl, B., Ardekani, B., Gee, J. C., Mann, J. J., and Parsey, R. V. "**Evaluation of Volume-Based and Surface-Based Brain Image Registration Methods**" *Neuroimage* 51, no. 1 (2010): 214–20. doi:10.1016/j.neuroimage.2010.01.091

86. Price, C. C., Wood, M. F., Leonard, C. M., Towler, S., Ward, J., Montijo, H., Kellison, I., Bowers, D., Monk, T., Newcomer, J. C., and Schmalfuss, I. "**Entorhinal Cortex Volume in Older Adults: Reliability and Validity Considerations for Three Published Measurement Protocols**" *J Int Neuropsychol Soc* 16, no. 5 (2010): 846–55. doi:10.1017/S135561771000072X

87. Fischl, B., Stevens, A. A., Rajendran, N., Yeo, B. T. T., Greve, D. N., Van Leemput, K., Polimeni, J. R., Kakunoori, S., Buckner, R. L., Pacheco, J., Salat, D. H., Melcher, J., Frosch, M. P., Hyman, B. T., Grant, P. E., Rosen, B. R., Kouwe, A. J. W. van der, Wiggins, G. C., Wald, L. L., and Augustinack, J. C. "**Predicting the Location of Entorhinal Cortex from MRI**" *Neuroimage* 47, no. 1 (2009):

bioRxiv preprint doi: https://doi.org/10.1101/170209; this version posted August 16, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

8–17. doi:10.1016/j.neuroimage.2009.04.033

88. Augustinack, J. C., Huber, K. E., Stevens, A. A., Roy, M., Frosch, M. P., Kouwe, A. J. W. van der, Wald, L. L., Van Leemput, K., McKee, A. C., Fischl, B., and Alzheimer's Disease Neuroimaging Initiative. "**Predicting the Location of Human Perirhinal Cortex, Brodmann's Area 35, from MRI**" *Neuroimage* 64, (2013): 32–42. doi:10.1016/j.neuroimage.2012.08.071

89. Shimony, J. S., Smyser, C. D., Wideman, G., Alexopoulos, D., Hill, J., Harwell, J., Dierker, D., Van Essen, D. C., Inder, T. E., and Neil, J. J. "**Comparison of Cortical Folding Measures for Evaluation of Developing Human Brain**" *Neuroimage* 125, (2016): 780–790. doi:10.1016/j.neuroimage.2015.11.001

90. Avants, B. and Gee, J. "**The Shape Operator for Differential Analysis of Images**" *Inf Process Med Imaging* 18, (2003): 101–13.

91. Vemuri, P., Senjem, M. L., Gunter, J. L., Lundt, E. S., Tosakulwong, N., Weigand, S. D., Borowski, B. J., Bernstein, M. A., Zuk, S. M., Lowe, V. J., Knopman, D. S., Petersen, R. C., Fox, N. C., Thompson, P. M., Weiner, M. W., Jack, C. R., Jr, and Alzheimer's Disease Neuroimaging Initiative. "**Accelerated Vs. Unaccelerated Serial Mri Based Tbm-Syn Measurements for Clinical Trials in Alzheimer's Disease**" *Neuroimage* 113, (2015): 61–9. doi:10.1016/j.neuroimage.2015.03.026

92. Muschelli, J., Gherman, A., Fortin, J.-P., Avants, B., Whitcher, B., Clayden, J. D., Caffo, B. S., and Crainiceanu, C. M. "**Neuroconductor: An R Platform for Medical Imaging Analysis**" *Biostatistics* (2018): doi:10.1093/biostatistics/kxx068