Check for updates

Received: 3 May 2024 | Revised: 3 September 2024 | Accepted: 10 September 2024

DOI: 10.1002/alz.14303

Alzheimer's & Dementia®
THE JOURNAL OF THE ALZHEIMER'S ASSOCIATION

RESEARCH ARTICLE

Implementation and validation of face de-identification (de-facing) in ADNI4

Christopher G. Schwarz[1] | **Mark Choe**[2] | **Stephanie Rossi**[3] | **Sandhitsu R. Das**[4] | **Ranjit Ittyerah**[5] | **Evan Fletcher**[6] | **Pauline Maillard**[6] | **Baljeet Singh**[6] | **Danielle J. Harvey**[7] | **Ian B. Malone**[8] | **Lloyd Prosser**[8] | **Matthew L. Senjem**[9] | **Leonard C. Matoush**[9] | **Chadwick P. Ward**[1] | **Carl M. Prakaashana**[1] | **Susan M. Landau**[10] | **Robert A. Koeppe**[11] | **JiaQie Lee**[10] | **Charles DeCarli**[6] | **Michael W. Weiner**[3] | **Clifford R. Jack Jr.**[1] | **William J. Jagust**[10] | **Paul A. Yushkevich**[5] | **Duygu Tosun**[2,3] | **for the Alzheimer's Disease Neuroimaging Initiative**

[1]Department of Radiology, Mayo Clinic, Rochester, Minnesota, USA

[2]Northern California Institute for Research and Education, San Francisco Veterans Affairs Medical Center, San Francisco, California, USA

[3]Department of Radiology, University of California, San Francisco, San Francisco, California, USA

[4]Department of Neurology, University of Pennsylvania, Philadelphia, Pennsylvania, USA

[5]Department of Radiology, University of Pennsylvania, Philadelphia, Pennsylvania, USA

[6]Department of Neurology, University of California, Davis, Davis, California, USA

[7]Division of Biostatistics Department of Public Health Sciences, , University of California, Davis, Davis, California, USA

[8]Dementia Research Centre, Dementia Research Centre, UCL Institute of Neurology, Queen Square, London, UK

[9]Department of Information Technology, Mayo Clinic, Rochester, Minnesota, USA

[10]Helen Wills Neuroscience Institute, University of California, Berkeley, Berkeley, California, USA

[11]Department of Radiology, University of Michigan, Ann Arbor, Michigan, USA

Correspondence
Christopher G. Schwarz, Department of Radiology, Mayo Clinic, 200 First Street SW, Rochester, Minnesota, 55905, USA.
Email: schwarz.christopher@mayo.edu

Funding information
National Institute on Aging, Grant/Award Numbers: R01AG068206, RF1AG056014, R01AG474069, P30AG072979, R01AG070592

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

Abstract

INTRODUCTION: Recent technological advances have increased the risk that de-identified brain images could be re-identified from face imagery. The Alzheimer's Disease Neuroimaging Initiative (ADNI) is a leading source of publicly available de-identified brain imaging, who quickly acted to protect participants' privacy.

METHODS: An independent expert committee evaluated 11 face-deidentification ("de-facing") methods and selected four for formal testing.

RESULTS: Effects of de-facing on brain measurements were comparable across methods and sufficiently small to recommend de-facing in ADNI. The committee ultimately recommended *mri_reface* for advantages in reliability, and for some practical considerations. ADNI leadership approved the committee's recommendation, beginning in ADNI4.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.
© 2024 The Author(s). *Alzheimer's & Dementia* published by Wiley Periodicals LLC on behalf of Alzheimer's Association.

implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A link to a complete listing of ADNI investigators can be found in the appendix.

DISCUSSION: ADNI4 de-faces all applicable brain images before subsequent pre-processing, analyses, and public release. Trained analysts inspect de-faced images to confirm complete face removal and complete non-modification of brain. This paper details the history of the algorithm selection process and extensive validation, then describes the production workflows for de-facing in ADNI.

KEYWORDS
ADNI, anonymization, de-facing, de-identification, face recognition

Highlights

- ADNI is implementing "de-facing" of MRI and PET beginning in ADNI4.
- "De-facing" alters face imagery in brain images to help protect privacy.
- Four algorithms were extensively compared for ADNI and mri_reface was chosen.
- Validation confirms mri_reface is robust and effective for ADNI sequences.
- Validation confirms mri_reface negligibly affects ADNI brain measurements.

1 | BACKGROUND

Rapid advances in computerized face recognition technology over the past decade have increased the potential that publicly available de-identified research brain images may be re-identified using face imagery in magnetic resonance imaging (MRI), positron emission tomography (PET), and computed tomography (CT), with estimates for MRI as high as 98%, PET as high as 42%, and CT as high as 78%–83%.[1–3] Advances in medical imaging technology have also made newer images increasingly identifiable, but many older sequences and scanners also have highly identifiable faces.[3,4] This has led to an increased interest in software for removing or replacing faces in brain images, also called "face-de-identification" or "de-facing." The Alzheimer's Disease Neuroimaging Initiative (ADNI) was launched in 2003 as a public-private partnership, led by Principal Investigator Michael W. Weiner, MD. The primary goal of ADNI has been to test whether serial MRI, PET, other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment (MCI) and early Alzheimer's disease (AD). ADNI is a leading source of publicly available neuroimaging data, including over 120,000 images of over 2800 unique participants from over 59 sites across the United States and Canada. However, this role comes with the responsibility to maintain the privacy of participants in the study who volunteer their participation with the assumption that their identity and personal medical information will not be revealed publicly.

Shortly before ADNI began in 2004, leadership discussed the concern of privacy risks from faces in brain images, and they decided to release images without de-facing because (1) there were concerns that de-facing would adversely affect brain measurements, and (2) the concern was largely theoretical; no research had measured the potential for face recognition from brain images. In a much later, 2009 study, only 40% of human visual raters exceeded chance in their rates of successfully matching MRI-based face reconstructions with participant photographs.[5] The first study using automated face recognition was not until 2012, and the face recognition software of that generation, Google Picasa, correctly matched only 27.5% of photographs with CT-based face reconstructions.[6] Although these studies were available before ADNI3 began in 2016, their findings did not cause a re-evaluation of ADNI's strategy. ADNI's decision not to implement de-facing in its earlier phases was also consistent with most contemporary neuroimaging studies of aging in the United States.

In late 2018, Dr Schwarz (also the first author of this paper and the primary author of the mri_reface de-facing software) presented data to ADNI leadership that demonstrated a newly increased risk posed by recent advances in face recognition.[1] In response, ADNI convened a committee of neuroimaging experts led by Drs Duygu Tosun and Paul Yushkevich. To avoid any potential conflict of interest, or the appearance of one, the committee was comprised of imaging and image quantification experts who were not otherwise involved with de-facing research or associated with specific de-facing software. ADNI leadership was interested in implementing de-facing if available software could remove face imagery accurately and reliably without substantially affecting brain imagery or brain measurements derived from the images. The committee was charged with evaluating available de-facing software for potential use in ADNI and making a specific recommendation to ADNI leadership. To maintain the committee's independence while they compared various de-facing software, Dr Schwarz and the Mayo team were not involved in the design, analyses, or decisions at that time. His contributions were limited to providing software for face reconstruction from MRI, for potential use in validation, and to running mri_reface on the committee's designated ADNI datasets (in the same equal role as the other software's authors).

This paper intends firstly provide a history of the analyses and rationale that drove ADNI's decisions regarding de-facing and, ultimately, to use mri_reface specifically. These initial analyses occurred in early 2021 and we recommend that current readers interpret them within

15525279, 2024, 11, Downloaded from https://alz-journals.onlinelibrary.wiley.com/doi/10.1002/alz.14303, Wiley Online Library on [17/06/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

their historical context, as we present them with only minor updates for more consistent plotting. Secondly, this paper returns to the present context and details which image types are de-faced in ADNI, its workflows for de-facing and quality control of the de-facing process, and the rationales for these decisions.

2 | METHODS

2.1 | Preliminary comparison of de-facing algorithms using ADNI MRI

Selection of four methods: The committee considered organizing "grand challenge" style competition to select the ADNI de-facing strategy but, given the time-sensitive nature of addressing the perceived risk of participant re-identification posed by artificial intelligence (AI), opted for a faster and narrower approach. The committee conducted an initial literature survey of 11 published de-facing programs and narrowed their consideration to four candidate programs, based on three criteria: (1) that they should already be in use by major imaging consortia, with exception made for new approaches for which there was evidence of effectiveness; (2) that they would not only remove facial features (de-facing), but also replace the missing regions with synthetic image content (re-facing), although exceptions would be made for methods with widespread current use; (3) and that they would be practical to implement within the ADNI software ecosystem, with developers willing to assist in this implementation. The programs selected by the committee were: fsl_deface (https://git.fmrib.ox.ac.uk/fsl/fsl_deface), mri_reface (https://www.nitrc.org/projects/mri_reface/), BIC Defacing algorithm (https://github.com/BIC-MNI/bic-pipelines), and the HCP/XNAT defacing pipeline (https://wiki.xnat.org/xnat-tools/face-masking). The fsl_deface is an open-source method that was included primarily because of its successful large-scale deployment in the UK Biobank.[7] The mri_reface was included because of its extensive published validations including testing with automated face recognition.[8] The BIC Defacing Algorithm[9] was included as another open source option, which had been used in PREVENT-AD.[10] The HCP/XNAT pipeline[11] was also included for being open source, available in XNAT, and because it had been previously used in some releases from the Human Connectome Project.[12] An example of images de-faced with all four programs is shown in Figure 1. A dataset was constructed, and authors of each of the four candidate programs were asked to process this dataset in their own environment or to provide the committee with instructions for doing so. The committee was charged with conducting an unbiased evaluation of the four methods and presenting the findings to the ADNI MRI Core and ADNI leadership, who would make the final selection of the de-facing method to implement in ADNI.

Dataset: A dataset was constructed by sampling a subset of 61 participants from those previously selected for the TADPOLE Challenge dataset.[13] The sample was approximately equally distributed across the clinically defined (biomarker-independent) ADNI subgroups: control (cognitively unimpaired, $n = 23$), mild cognitive impairment (MCI, $n = 20$), and AD dementia ($n = 18$), with equal distribution across the sexes (30 female, 31 male) and significant representation of Black and Hispanic Americans (approximately one-quarter ($n = 15$) of the participants self-identified as Hispanic; of the non-Hispanic participants, one-third ($n = 16$) self-identified as Black, one-third ($n = 16$) as White, and one-third ($n = 14$) as other racial/ethnical category or multiple racial/ethnical categories ($n = 7$ Asian, $n = 1$ American Indian/Alaskan Native, $n = 6$ more than one race). The average age was 73.8 years, with a range of 56.3–89. The significantly higher

RESEARCH IN CONTEXT

1. **Systematic review**: This manuscript describes the rationale and internal validations that were used by Alzheimer's Disease Neuroimaging Initiative (ADNI) regarding face de-identification or "de-facing". It adds to existing prior literature about de-facing and its effects. The authors reviewed the literature using traditional (e.g., PubMed) sources and meeting abstracts and presentations.
2. **Interpretation**: This study's findings on the efficacy of de-facing and its effects on brain measurements are consistent with previously published literature, which had only partial overlap in the de-facing software examined and in the analyses tested.
3. **Future directions**: ADNI is committed to providing robust, leading privacy protections for its generous ADNI participants. The current, described approach is being implemented for ADNI4, but ADNI will continue to evaluate new developments in this area as technology continues to advance.

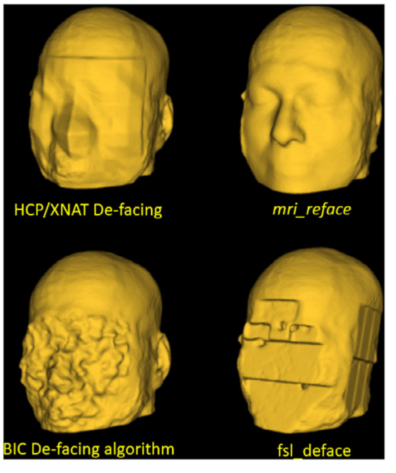

FIGURE 1 Face reconstructions (surface renderings) of example outputs from the four de-facing programs tested in the preliminary algorithm comparison phase of the analyses. Due to privacy concerns, we do not show reconstructions of unmodified images in this paper.

15525279, 2024, 11, Downloaded from https://alz-journals.onlinelibrary.wiley.com/doi/10.1002/alz.14303, Wiley Online Library on [17/06/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

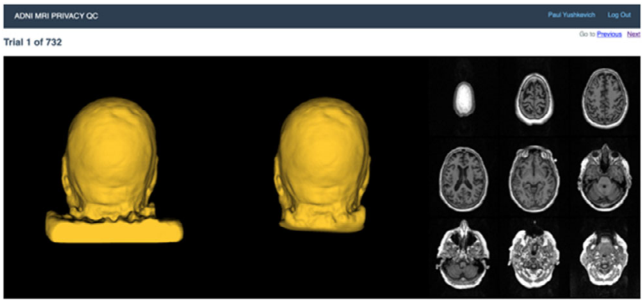

FIGURE 2 Interface developed at University of Pennsylvania for visual evaluation of de-facing in this work. Face reconstructions from unmodified images were shown on the left, and from de-faced images in the center, and axial slices from the de-faced images on the right. Users could freely rotate the face reconstructions.

representation of Hispanic and Black Americans in this dataset than in ADNI overall was driven by the objective to avoid the potential for bias with respect to race and ethnicity in algorithm selection, that is to avoid choosing a method that could have performed worse with images from under-represented groups. All participants were required to have three longitudinal imaging time points, for a total of 183 longitudinal scans.

Submission of de-faced images: Authors of each of the four candidate programs were contacted and sent instructions to download 183 specific T1-weighted, preprocessed "N3m" ADNI images (specified using their unique image IDs) in NIfTI format and run their programs on each image. They were instructed to use only one set of algorithm parameters for all 183 scans and avoid any manual editing or "tweaking" of the settings or outputs, that is, any "failure" images should be left as-is, replicating a fully automated workflow. Outputs were to be named in a standard format, and authors would upload each de-faced NIfTI image, corresponding to each input NIfTI image, to a standard location with a pre-specified naming format. Each author ran their methods themselves as instructed, except the author of one method (HCP/XNAT) was not sufficiently available, so this one was run by a member of the ADNI evaluation committee instead, while coordinating with the original authors.

Visual inspection of results: The University of Pennsylvania team developed a web-based interface to display and allow free rotation of face reconstructions from the unmodified and de-faced images side-by-side. This display also included six axial slices of the de-faced MRI output (Figure 2). For each de-facing method on each image (61 participants $\times$ 4 methods $\times$ 3 timepoints = 732 comparisons total), three independent raters (M.C., R.I., S.R.) were each asked to evaluate (1) recognizability of six different facial features in the surface rendering (cheekbones, chin/jaw, ears, eyebrows, mouth, nose), and (2) preservation of brain voxels in the MRI scan. All ratings were performed using a Likert scale. These ratings used only the display shown in Figure 2, so "recognizability" refers to whether each face-part in the de-faced image on the right subjectively resembled the corresponding face-part in the unmodified image on the left, and brain preservation was based on the slices from the de-faced image. When viewing de-faced images, raters were blinded to which de-face method was used on that image, although the methods are visually distinct enough that complete blinding would be impossible. It should be noted that these experiments were not designed to absolutely quantify the risk of participant re-identification via face recognition, but instead to compare available software relative to each other in terms of how much of the face was obscured, based on the intuitive assumption that there is a correlation relationship between face preservation and re-identification risk.

Effects on ADNI standard brain measurements: The University of California, San Francisco team quantitatively compared brain measurements from unmodified images with those from de-faced images from each method. Four brain measurement pipelines routinely applied by different analyses groups within the ADNI MRI core were each tested using outputs from each de-facing software: (1) FreeSurfer version 5.1 (longitudinal stream)[14] at the University of California, San Francisco; (2) Boundary Shift Integral (BSI), at University College, London[15]; (3) Tensor-Based Morphometry with Symmetric Normalization (TBM-SyN) at Mayo Clinic[16]; and (4) Automatic Segmentation of Hippocampal Subfields (ASHS) at University of Pennsylvania.[17] Each analysis group was blinded to the method used for de-facing as well as demographics and diagnostic information. FreeSurfer performance was measured using total subcortical volume and average cortical thickness measurements. BSI measures included were brain BSI (BBSI), ventricle BSI (VBSI), brain volume, and ventricle volume. Analyses from TBM-SyN included longitudinal change in parahippocampal cortex volume, and analyses from ASHS included volumes of the entorhinal cortex (ERC) and total intracranial volume (TIV), also known as intracranial volume (ICV). All measurements were compared using two metrics: intraclass correlation coefficient (ICC); and mean bias (100*(defaced-unmodified))/unmodified).

2.2 | Secondary validation of *mri_reface* in ADNI

From the results of the preliminary analyses (see Results, below) and deliberations among the ADNI MRI Core members, the MRI Core selected mri_reface as the de-facing solution for ADNI. Before the decision to adopt mri_reface in ADNI4 was finalized, an additional round of validations was performed using mri_reface on a new dataset, for several reasons: (1) there were substantial updates in newer versions of mri_reface since the initial evaluation; (2) white matter hyperintensity (WMH) measurements from T2-FLAIR and PET SUVR analyses were not included in the initial round conducted using only T1-weighted MRI, and (3) an abundance of caution and thoroughness. Unlike the previous validations that were intended to compare de-facing software, this secondary validation of mri_reface was designed with the help of its primary author (also the first author of this manuscript). Changes between mri_reface 0.2 and 0.3 included adding support for PET and CT images, and adding noise in the replacement face for each image to match the faces in the input image, among smaller convenience and troubleshooting additions that did not change the primary outputs.

15525279, 2024, 11, Downloaded from https://alz-journals.onlinelibrary.wiley.com/doi/10.1002/alz.14303, Wiley Online Library on [17/06/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Dataset: A second, cross-sectional, multi-modal dataset was constructed by the ADNI Biostatistics Core, consisting of 100 ADNI3 participants with T1-weighted MRI, T2-FLAIR-weighted MRI, amyloid PET, and tau PET. Its demographics were: age: mean 74 years (range 56–89); cognitive status: 52 cognitively unimpaired, 36 mild cognitive impairment, 12 with dementia; sex: 55 female, 45 male; 50% self-identified as a member of at least one under-represented racial or ethnic group. The distribution of MRI scanner manufacturers was: 62 Siemens, 21 GE, 17 Philips. The distribution of PET scanners for amyloid PET was: 48 Siemens, 29 GE, 16 Philips, after seven scans were rejected for image quality issues. The distribution of PET scanners for tau PET was: 40 Siemens, 35 GE, 16 Philips, after nine scans were rejected for image quality issues.

Methods: The team at Mayo Clinic ran *mri_reface* version 0.3 (the latest at the time, vs. 0.2 at the initial phase of validation) on all 100 imaging studies, each including T1-weighted MRI, T2-FLAIR MRI, amyloid PET, and tau PET. All de-facing was conducted on the most "raw" form of each series prior to any standard preprocessing, unlike the previous phase that used "processed" images, because the mri_reface team recommended that this would be the most appropriate workflow for two reasons: (1) to minimize distribution of the more identifiable "raw" data to analyses sites performing the preprocessing, and (2) to allow subsequent preprocessing steps like inhomogeneity correction to help correct any potential inhomogeneity that might be introduced during the de-facing. This first phase did not consult the mri_reface team in its design (to avoid potential conflicts of interest [COIs]), and did not consider these factors. For PET, this was the individual, dynamic PET image frames, *before* summed images are produced later during preprocessing. Despite the "mri_reface" name, the software already also supported PET (and CT) by that time, and no ADNI-specific modifications were needed. By this time, the Mayo group had developed DICOM support for mri_reface and provided all de-faced images in DICOM format for analyses by the various groups, differing from the previous round that used all NIfTI images.

BSI: The team at University College London tested the new dataset with their same BSI pipeline. However, because BSI is an intrinsically longitudinal measurement (produces direct measures of change), its usage was modified in this second analyses of cross-sectional data. Instead of measuring longitudinal data with versus without de-facing, as in the previous longitudinal analysis with a longitudinal dataset, these analyses with cross-sectional data used paired data of original versus de-faced images from the same time-point. Under this scenario, zero change would be expected, and any deviation would be attributable to an unknown combination of noise and the effects of de-facing. This type of analyses is less representative of "real world" usage than the previous variant that compared longitudinal pairs of original versus de-faced images, but it acts as a simulation that provides a more straightforward measurement with a known ground truth, and the alternative was already performed in the previous phase, so they would complement each other.

WMH: The UC Davis IDeA laboratory tested the de-faced T1-w and T2-FLAIR images using their in-house pipeline for WMH analyses, which uses a multi-step approach: (1) Removal of non-brain tissue using a convolutional neural net architecture[18]; (2) inhomogeneity correction[19]; (3) adaptive image segmentation of gray, white and CSF tissues[20] and Bayesian estimation of WMH.[21]

PET: Finally, all de-faced, dynamic PET images were sent to the University of Michigan for visual QA and standard ADNI PET preprocessing including motion correction, standardizing image resolution and voxel size, and creation of summed images.[22] These de-faced summed images were then sent for quantitative analyses at UC Berkeley, also following the standard ADNI pipelines.[22] These pipelines use MRI for quantification of PET images, and de-faced PET were measured using corresponding de-faced MRI, as would be done in practice.

3 | RESULTS

3.1 | Preliminary comparison of de-facing algorithms in T1-weighted MRI

Visual face recognizability analyses: Results of the visual face recognition analyses are shown in Figure 3, top and center. The major findings were: (1) raters favored the fsl_deface and mri_reface methods over HCP/XNAT and BIC because the former remove ears and the latter do not; (2) raters found the results from fsl_deface to be the least visually recognizable when it performed correctly, but it occasionally failed and left significant portions of the face intact (32/1626, ≈2.0%, of ratings were very recognizable, 4 or 5); (3) raters found the results from mri_reface to be "recognizable," but they noted that this was likely a product of mri_reface being the only method where the face was replaced with an average face rather than removed, so it inherently looked more like a face than other methods, and some participants are bound to look more like an average face. They also noted that mri_reface had the fewest complete failures, that is, it *never* completely retained the face for any of the tested images (evidenced by a lack of ratings of "5" for mri_reface in Figure 3, top and center, although there were two images where one rater marked the ears as "most recognizable").

Visual checks for brain alteration by de-facing: Results of the visual checks for brain alteration by the de-facing algorithms are shown in Figure 3, bottom. Raters found all four methods were consistent in preserving the brain, except for *fsl_deface*, which "occasionally" (according to two raters) or "frequently" (according to one rater) removed some brain voxels in superior frontal regions.

FreeSurfer: Unmodified and de-faced images from each program were analyzed with FreeSurfer 5.1, using the Longitudinal stream, to measure total subcortical volume and whole-brain average cortical thickness. These results are shown in Figure 4, top. ICC values from global cortical thickness were 0.97 for all de-facing methods and no differences between them were significant. Similarly, ICC values from total subcortical volumes were 0.98 for all de-facing methods and no differences between them were significant. Biases were significantly different for only one of the six pairings across the four de-facing

15525279, 2024, 11, Downloaded from https://alz-journals.onlinelibrary.wiley.com/doi/10.1002/alz.14303, Wiley Online Library on [17/06/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Alzheimer's & Dementia®
THE JOURNAL OF THE ALZHEIMER'S ASSOCIATION

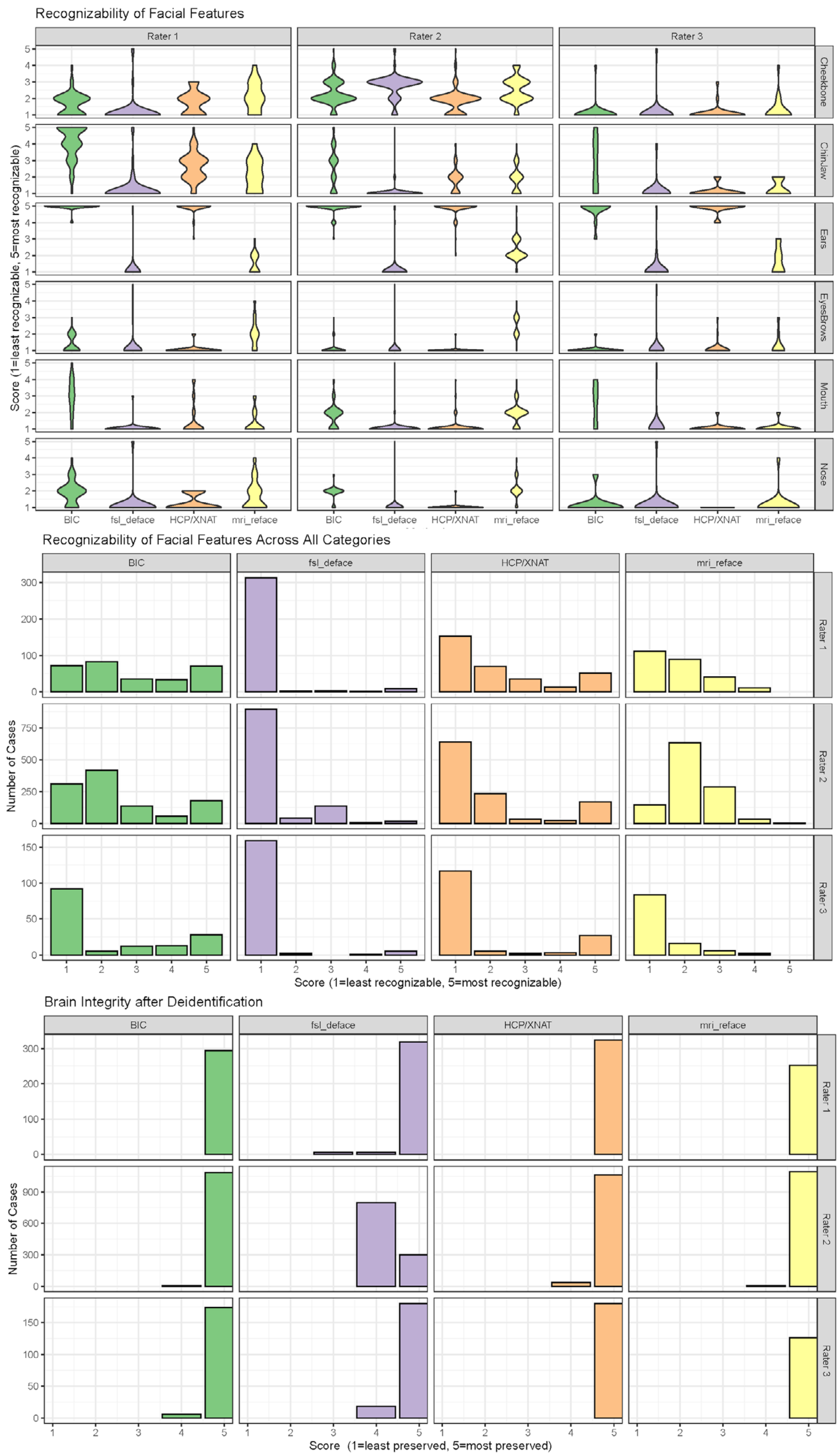

FIGURE 3 Three independent raters scored their subjective assessment of recognizability of six facial features in images de-faced with each candidate program, using a Likert scale (1 = least recognizable, 5 = most recognizable). Top: Their ratings for each of the facial features individually; Center: Their ratings summed across all six facial features. mri_reface and fsl_deface were favored because the other two did not remove ears. fsl_deface was considered least recognizable for most cases because the results look less like faces at all, versus the other methods, but it occasionally failed and left the face partly intact. mri_reface had the fewest failures at removing the complete face. Bottom: Three independent raters scored their subjective assessment of brain integrity after de-identification in images de-faced with each candidate program, using a Likert scale (1 = least preserved; 5 = most preserved). All methods mostly preserved the brain, but mri_reface had the fewest accidental brain modifications, and fsl_deface commonly removed some superior brain voxels.

15525279, 2024, 11, Downloaded from https://alz-journals.onlinelibrary.wiley.com/doi/10.1002/alz.14303, Wiley Online Library on [17/06/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

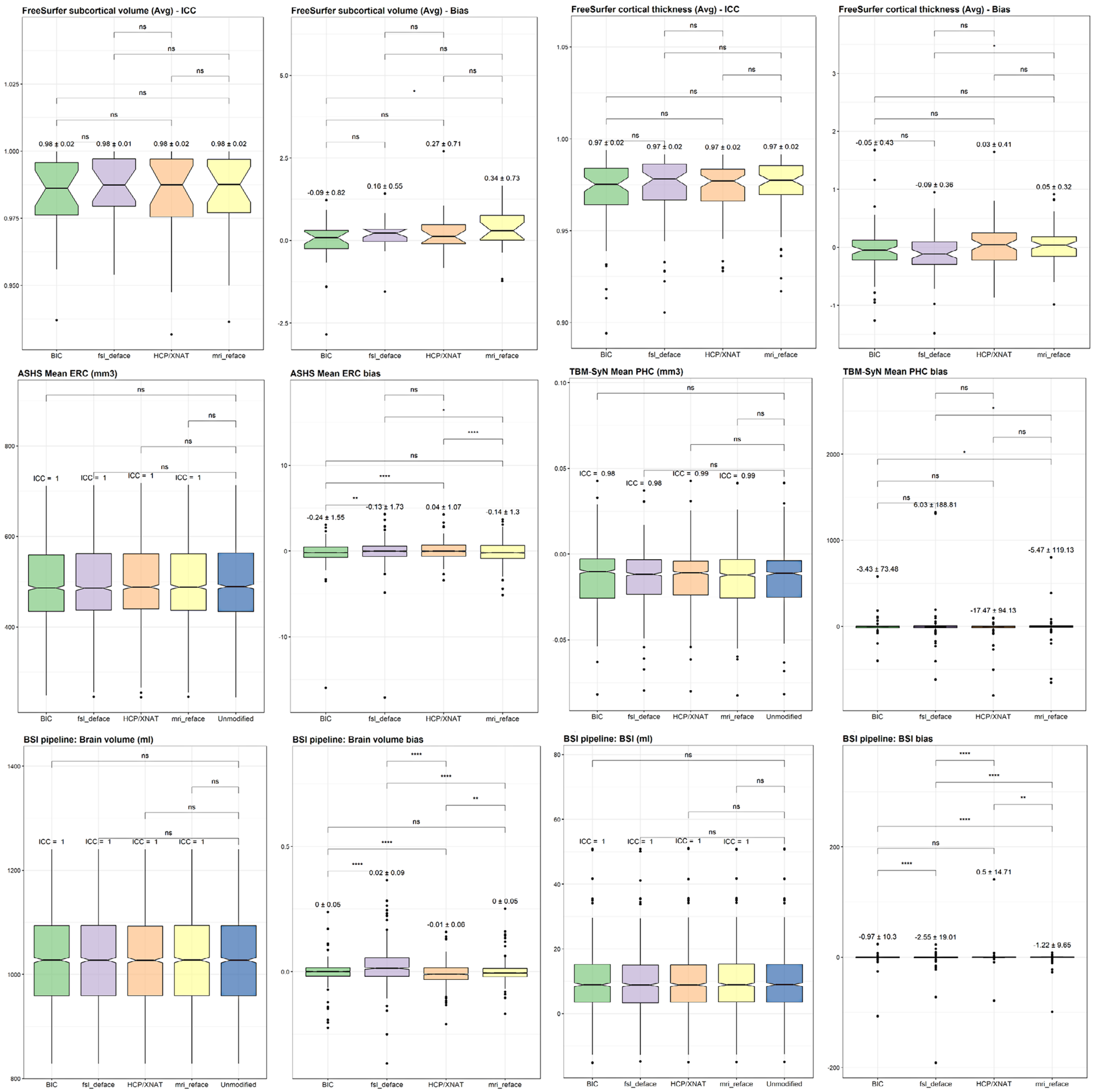

FIGURE 4 Intraclass correlation coefficient (ICC) and bias (subtractive difference %) of measurements from unmodified vs. de-faced images, from each of the four candidate programs. Top: Left two: effects of de-facing on subcortical volume measurements; right two: effects on cortical thickness measurements, both from longitudinal FreeSurfer 5.1. . Center: Left two: effects on entorhinal cortex (ERC) volumes from the Automated Segmentation of Hippocampal Subfields (ASHS) pipeline at University of Pennsylvania; right two: effects on longitudinal measurements of atrophy in the parahippocampal cortex from the tensor-based morphometry with symmetric normalization (TBM-SyN) pipeline at Mayo Clinic. Bottom: Effects on brain volume (left two) and boundary shift integral (BSI), both as measured by the BSI pipeline at UCL. For all of these plots, "ns" denotes non-significance at $p > 0.05$, "*" denotes $p \leq 0.05$, "**" denotes $p \leq 0.01$, "***" denotes $p \leq 0.001$, and "****" denotes $p \leq 0.0001$.

methods for each of these two measurements, and the significant pairings were not consistent.

ASHS: To measure effects of de-facing on ASHS, the team analyzed entorhinal cortex volumes (mm^3) across the dataset, and also total intracranial volume measurements, and the effects are shown in Figure 4, center left. For the ERC volumes, there were no significant pair-wise differences between the raw images and those from any de-facing software. ICC values between each de-face method and unmodified images were all 1.0. Pair-wise differences between biases were significant for four of the six pairings, but their magnitudes were not considered large enough to be practically relevant.

15525279, 2024, 11, Downloaded from https://alz-journals.onlinelibrary.wiley.com/doi/10.1002/alz.14303, Wiley Online Library on [17/06/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TBM-SyN: To measure effects of de-facing on TBM-SyN, the team analyzed annualized rates of atrophy (mm^3) in the parahippocampal cortex, and these are shown in Figure 4, center right. There were no significant pair-wise differences in this measurement between the raw images and those from any de-facing software. ICC values between each de-face method and unmodified images were between 0.98 and 0.99. Pair-wise differences between biases were significant for two of the six pairings, but their magnitudes were not considered large enough to be practically relevant.

BSI: To measure effects of de-facing on the BSI pipeline, the team analyzed its measurements of brain BSI and brain volume with unmodified images and with images from each de-facing software, and these are shown in Figure 4, bottom. There were no significant pair-wise differences in BSI between the raw images and those from any de-facing software, and also not between any of the de-facing methods. ICC values between original and de-faced measurements were all 1.00. Differences in biases in this measurement were significant for five of the six pair-wise comparisons between the de-facing software, but the magnitudes were all very small. For the brain volumes from the BSI pipeline, there were also no significant pair-wise differences, and all ICC values were 1.00. Significant pair-wise differences in biases in this measurement were observed in five of the six comparisons, but their magnitudes were not considered large enough to be practically relevant.

Summary of preliminary comparison: Effects of all de-facing methods on downstream pipelines were relatively small and mostly not significant, but there were some aspects that differentiated them. fsl_deface was excluded primarily because it modified brain more often than the other methods, despite its output faces being most consistently being rated as "least recognizable." BIC and HCP/XNAT do not remove the ears, and the committee felt that removing the ears would be preferrable, even though these methods performed relatively well in some measurements of downstream effects. mri_reface was considered the most reliable choice to remove the face (fewest "most recognizable" ratings) and not modify the brain (fewest brain integrity ratings worse than "most preserved"), and its effects on downstream measurements were comparable with the others, but there was some concern that the replaced face sometimes resembled the original more than other methods (i.e., some participants look a lot like the replacement average face, but none ever resemble a blurry or missing face). ADNI ultimately needed to select only one method, so additional, practical factors were also considered to help make a decision. The close involvement of the developer of mri_reface in the ADNI community, his ongoing research on de-facing effectiveness in the context of AI-based face matching, and his availability to continue developing the method to further suit the needs of ADNI, were considered as additional strengths in favor of mri_reface. These results were presented to ADNI leadership, and ADNI leadership decided to follow the committee's recommendations. ADNI leadership also favored mri_reface for the additional, practical reason that de-facing would be implemented by the MRI core at Mayo Clinic, and the Mayo team would be better able to implement and support their own method than any others. After extensive discussion and analyses, ADNI leadership decided that de-facing of all applicable images would use mri_reface and begin with ADNI phase 4 (ADNI4). Images from previous ADNI phases (ADNI 1, 2, GO, and 3) would not be de-faced retrospectively because the time and cost required would have relatively limited benefit when these images from previous cycles had already been downloaded by so many researchers around the world. Before this de-facing would begin for ADNI4, a secondary analysis was conducted using a since-released later version of mri_reface (0.3), including additional analyses and image types (described previously in Methods). Its results are described below.

3.2 | Secondary validation of mri_reface in ADNI multi-modal imaging

The secondary validation of an updated mri_reface in ADNI examined its effects on measurements from the ADNI standard BSI, TBM-SyN, WMH, and PET pipelines.

BSI: The team from UCL ran their BSI pipeline on pairs of original versus de-faced imaging data. Since these both came from the same participant images at the time point, zero change would be expected. For each volume measurement, the percent difference between the values, relative to the mean of the two values, was mean (standard deviation [SD]): brain volume: –0.041% (0.51%); ventricle volume: 0.033% (0.59%); The two BSI measurements are directly calculated as a percent change, and they were mean (SD): brain BSI (BBSI): –0.014% (0.10%); ventricle BSI (VBSI): –0.013% (0.173%). Brain BSI values were not significantly different from zero (t-test, $p = 0.33$), and neither were ventricle BSI ($p = 0.52$). In synopsis, the UCL group concluded that there was "little evidence of non-zero change" in BSI measurements.

WMH: Effects of running mri_reface on the T1-weighted and T2-FLAIR-weighted images on the UC Davis WMH pipeline are shown in Figure 5. The effects were considered minimal ($R^2 > 0.97$, ICC > 0.97), and the UC Davis group approved the plan to de-face these images.

PET: The team at University of Michigan performed visual inspection of all the de-faced PET images from *mri_reface*, performed their standard preprocessing, and visually inspected these outputs as well. They did not report any concerns from these visual inspections, and they passed the preprocessed images to the UC Berkeley team for quantitative analyses. Results are presented in Figure 6. Across all included measurements, the mean of absolute percent differences between values from unmodified and de-faced images ranged from 0.26%–2.94%. R^2 values were examined for the summary measures for each PET tracer separately, and for both pipeline variants, and all R^2 values were ≥ 0.99. The UC Berkeley team reported that they were satisfied with the plan to use mri_reface in ADNI4. Later, mri_reface version 0.3.4 was proposed, which had significant improvements to the algorithms for adding noise in replacement faces to match the noise in PET images (but no changes for MRI or other image types). Sixty images were subsampled from the previous 100, which included 30 of the images that previously had relatively larger deviations in measured SUVR from their unmodified versions, and 30 that previously had

15525279, 2024, 11, Downloaded from https://alz-journals.onlinelibrary.wiley.com/doi/10.1002/alz.14303, Wiley Online Library on [17/06/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

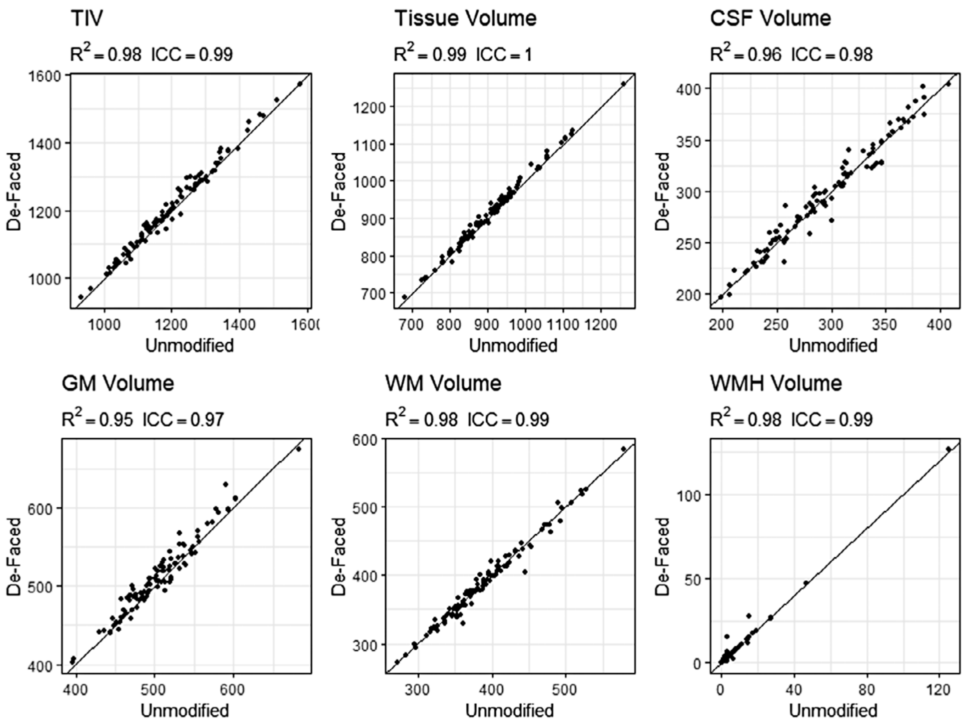

FIGURE 5 Effects of de-facing T1-w and T2-FLAIR-w images with *mri_reface* on the UC Davis WMH pipeline.

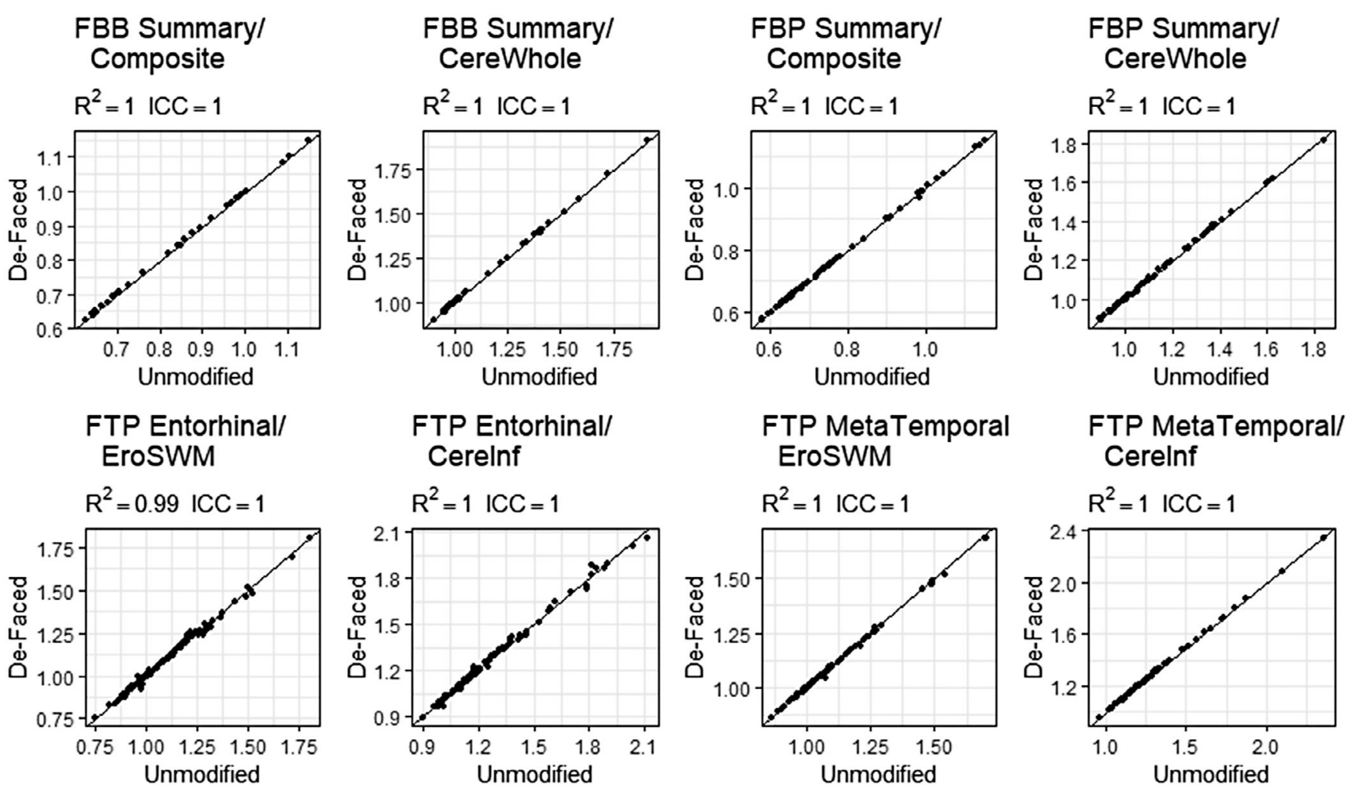

FIGURE 6 Effects of de-facing T1-w MRI and "raw" PET images with *mri_reface* on PET quantification pipelines at UC Berkeley, after their ADNI standard preprocessing at U. Michigan. FBB, florbetaben; FBP, florbetapir; FTP, flortaucipir; EroSWM, eroded supratentorial white matter; CereWhole, whole cerebellum; CereInf, inferior cerebellum.

15525279, 2024, 11, Downloaded from https://alz-journals.onlinelibrary.wiley.com/doi/10.1002/alz.14303, Wiley Online Library on [17/06/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 1 Image types in ADNI4 and their approach to de-facing.

Modality	Image type	Approach in ADNI4
MRI	T1-weighted	De-faced[3]
MRI	T2-weighted (FLAIR and SPACE/CUBE/VISTA)	De-faced[3]
MRI	T2*-weighted (single and multi-echo)	De-faced[3]
MRI	Arterial spin labelling	De-faced (was not indicated in[3], but newly released sequences have changed)
MRI	High-resolution hippocampal T2-weighted	De-facing not needed (face is outside the field of view)
MRI	Diffusion MRI	De-facing not needed[3]
MRI	Functional MRI (task-free)	De-facing not needed[3]
PET	Amyloid PET (Florbetapir and Florbetaben)	De-faced[2]
PET	Tau PET (Flortaucipir and MK-6240)	De-faced[2]

relatively smaller deviations. After de-facing, preprocessing, and analyzing each image, the measurements from PET de-faced with mri_reface 0.3.4 were compared to those from mri_reface 0.3.3, and no significant differences were observed. The PET core decided to proceed with 0.3.4 for images going forward.

Summary of secondary validation: From the sum of these results, the groups at UCL, UC Davis, Mayo Clinic, UC Berkeley, and U. Michigan each agreed that de-facing with *mri_reface* did not cause any concerning effects on their ADNI standard pipelines, and they agreed with ADNI's plan to conduct de-facing of applicable images in ADNI4.

3.3 | Implementation of *mri_reface* for ADNI at Mayo Clinic

The *mri_reface* and database teams from the MRI core at Mayo Clinic created workflows for de-facing of ADNI images. These include the following steps: (1) automated nightly downloads of all ADNI4 MRI and PET images from the central repository (LONI); (2) automated classification of all MRI and PET images into broad image type categories using DICOM header information, needed to determine whether the image type needs de-facing and what template should be used by mri_reface during de-facing; (3) automated running of mri_reface on all applicable image types (see Table 1); (4) visual quality control (QC) by a team of trained image analysts who inspect all de-faced images to ensure complete replacement of the face and complete retention of the brain; (5) automated uploads of all QC-passed, de-faced images back to the central repository at LONI; (6) automated reporting for all received DICOM series to track their statuses in these steps. All ADNI images received by this pipeline are in DICOM format, from which they are automatically converted to NiFTI format using dcm2niix,[23] automatically run through the latest version of mri_reface, and these de-faced NiFTI files are automatically converted back to DICOM format using nii2dicom, a utility written by the mri_reface team and included with mri_reface, prior to their automated re-upload to LONI. We show this de-facing process in the context of the ADNI imaging data flow in Figure 7, and we describe this in the next two paragraphs.

MRI data flow: For MRI series, these de-facing steps are conducted before any other post-upload procedures, including their subsequent quality control checking by a different, downstream team at Mayo Clinic. Images that pass that subsequent QC for quality and for compliance with ADNI standard protocols are then released for quantitative analyses and download by ADNI users. De-faced image types include all T1-weighted, T2-weighted, T2-FLAIR-weighted, and T2*-weighted (including ME-GRE) MRI. Based on previous research, diffusion MRI and functional MRI are considered minimal risk for re-identification[4] and are released without de-facing. High-resolution hippocampus scans are also released without de-facing because these image types do not contain faces within the image FOV. Although previous work[4] had found that perfusion (arterial spin labelling) MRI had minimal risk of re-identification, we subsequently discovered that recently released PCASL sequences from Siemens and Philips have much more identifiable face imagery when compared to their earlier PASL sequences and when compared to the existing PCASL sequence from GE. Consequently, ADNI will change their original plans and de-face all perfusion (arterial spin labeling [ASL]) images in ADNI4 as well, but this will not begin until sufficient data are collected from these new sequence types that effective de-facing can be developed and validated. All ADNI4 perfusion MRI are being held in quarantine, until then. An overview of MRI in ADNI, and its associated workflows, is available in this issue (Jack et al., *Overview of ADNI MRI*).

PET data flow: All PET image types in ADNI (amyloid PET and tau PET) carry sufficient risk of re-identification to warrant de-facing.[3] University of Michigan first inspects each series to determine whether it passes quality control and complies with ADNI standard protocols, and only the approved series are released to the Mayo group's queue for de-facing. The Mayo group performs de-facing and de-facing QC on these unprocessed, raw PET images (individual, dynamic frames), and de-faced PET are uploaded back to LONI, where they are received again by University of Michigan to perform subsequent PET preprocessing. The data flow for PET differs from MRI because PET QC is performed at University of Michigan, a different site from where de-facing is performed (Mayo Clinic), while MRI QC is performed at Mayo Clinic (the same site where de-facing is performed) and, therefore, can

15525279, 2024, 11, Downloaded from https://alz-journals.onlinelibrary.wiley.com/doi/10.1002/alz.14303, Wiley Online Library on [17/06/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

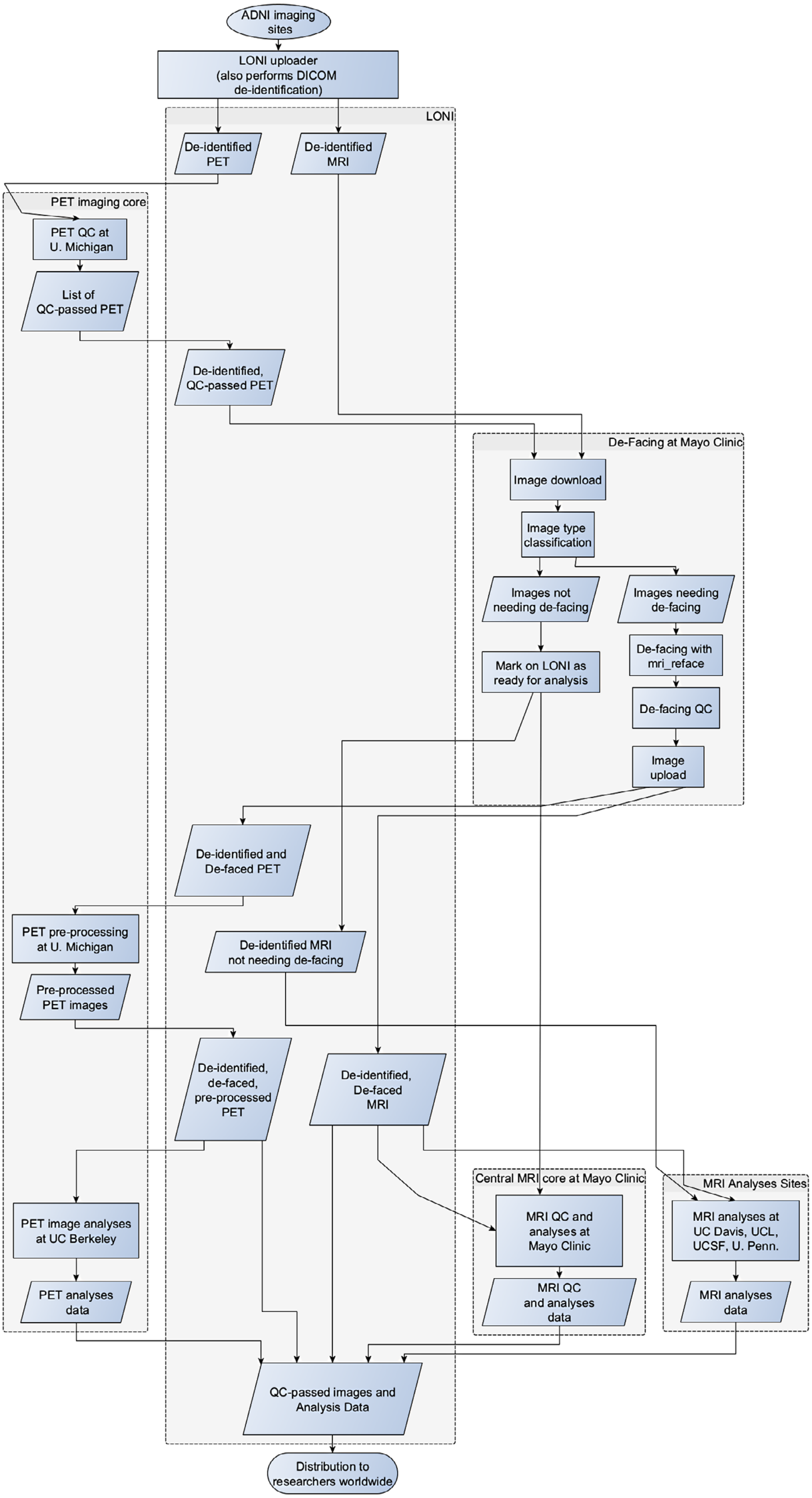

FIGURE 7 Flowchart of ADNI image, image analysis, and QC data, as it relates to the de-facing process.

15525279, 2024, 11, Downloaded from https://alz-journals.onlinelibrary.wiley.com/doi/10.1002/alz.14303, Wiley Online Library on [17/06/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

avoid an additional data transfer across sites. The de-faced "raw" PET series, and PET images from all four stages of ADNI-standard PET preprocessing of these de-faced series, are all made available for download by ADNI users, and the stage-4 preprocessed de-faced PET images are analyzed quantitively by the UC Berkeley team. An overview of PET in ADNI, and its associated workflows, is available in this issue (Landau et al., this issue).

4 | DISCUSSION

This document describes the rationale and exhaustive analyses that led ADNI to implement de-facing with mri_reface on all applicable image types in ADNI4, along with details of this implementation. Although re-identification of brain images from face imagery was not considered a sufficiently compelling risk to implement "de-facing" in earlier phases of ADNI, recent leaps in imaging and in automated face recognition technologies led to increased concerns,[1] and ADNI chose to reconsider this policy before the start of ADNI4. After several phases of validation, *mri_reface* was chosen as the de-facing approach for ADNI, and de-facing and QC workflows were implemented by the ADNI MRI Core at Mayo Clinic. The findings of ADNI's validations, presented in this work, were consistent with now-previously published works showing that de-facing programs generally have negligible effects on various research pipelines that analyze T1-weighted MRI.[8,3,24–26] They are also consistent with several previous comparisons in which mri_reface performed well relative to other programs.[8,26,27] ADNI's decision to de-face applicable images using mri_reface is consistent with many other large imaging studies that made this decision in the past several years, including SCAN (Standardized Centralized Alzheimer's and Related Dementias Neuroimaging), A4 (Anti-Amyloid Treatment in Asymptomatic Alzheimer's Disease), ALLFTD (ARTFL-LEFFTDS Longitudinal Frontotemporal Lobar Degeneration), NAPS2 (North American Prodromal Synucleinopathy Consortium for RBD, Stage 2), MCSA (Mayo Clinic Study of Aging), and others. However, it should be noted that these studies' decisions were not all made independently, as some chose to trust ADNI and mirrored their approach, and some have shared key personnel with ADNI or with authors of mri_reface. Other large brain imaging studies of aging and dementia that also perform de-facing, but using other software, include the UK-Biobank, HABS (Harvard Aging Brain Study), Prevent-AD (Pre-symptomatic Evaluation of Experimental or Novel Treatments for Alzheimer Disease), and HCP (Human Connectome Project) (in some phases).

While this paper is largely a historical retrospective, the analyses presented here for the first time have a unique strength in being, to our knowledge, the first published evaluation of de-facing software in a sample that was specifically enriched for inclusion of under-represented groups, and it is notable that the findings were largely consistent with previous analyses of less diverse samples. However, these analyses also had some limitations. The independent committee was selected based on their involvement in ADNI and perceived lack of bias with respect to any one de-facing approach, but its leaders were also not previously involved in face de-identification research, which may have led to suboptimal evaluation design. For example, the committee did not consider the need to remove ears when selecting the four candidate methods for preliminary evaluation, but ears were a clear driver of the recognizability evaluation by the three raters. To find and operationalize a de-facing solution for ADNI as soon as possible, the committee and ADNI leadership did not hold an open "grand challenge" competition, instead limiting the comparison to the four pre-selected methods. Had an open competition been held, it is possible that other methods not considered by the committee would have won. In another example, the committee designed face recognition experiments based on non-experts' visual ratings of perceived "similarity" between original and de-faced images, assuming that these visual ratings are directly correlated with rates of successful recognition by automated face recognition algorithms (i.e., re-identification risk). Although this relationship is intuitive and logical (missing faces cannot be recognized), we acknowledge that its strength and linearity have not been quantified. The committee did not have access to a dataset with paired MRIs and photos to directly measure face recognition performance, nor to leading face recognition software. These limitations were shared with many other validations of de-facing software at the time.[11,24,28,29] For comparisons using leading automated face recognition systems, we refer readers to other works.[3,4,8] A reader should also note that since the need for de-facing was considered established at that point, the visual rater "recognizability" experiments were designed primarily for comparing relative re-identification risk across de-facing algorithms, rather than absolute quantification.

Reviewers of this manuscript also pointed out that some analyses could have been improved (e.g., by adding Bland-Altman analyses, analyses of variability in addition to central tendency, or including additional statistical tests), but because this paper primarily describes a history, the authors felt it would be inappropriate to make these changes in the present. Another potential limitation is that while ADNI has implemented de-facing for a wide range of MRI and PET images, the initial comparison of multiple algorithms was limited only to T1-weighted MRI.

From the extensive ADNI-specific validations presented in this work for the first time, and from other previously published validations of de-facing,[8,27,30] ADNI's generous research volunteers should be assured that ADNI maintains a strong and leading commitment to protecting the privacy of their identities and collected health information, and researchers using ADNI data should be assured that their receiving only "de-faced" imaging data does not cause any significant effects on brain image analyses measured with a multitude of software.

ACKNOWLEDGMENTS

Data collection and sharing for the Alzheimer's Disease Neuroimaging Initiative (ADNI) is funded by the National Institute on Aging (National Institutes of Health Grant U19 AG024904). The grantee organization is the Northern California Institute for Research and Education. In the past, ADNI has also received funding from the National Institute of Biomedical Imaging and Bioengineering, the Canadian Institutes of Health Research, and private sector contributions through the Foundation for the National Institutes of Health (FNIH)

15525279, 2024, 11, Downloaded from https://alz-journals.onlinelibrary.wiley.com/doi/10.1002/alz.14303, Wiley Online Library on [17/06/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

including 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 authors gratefully acknowledge the members of the teams who developed the four de-facing methods selected for preliminary evaluation, who applied the methods to the common dataset and/or helped the ADNI de-facing committee do so: Vladimir Fonov and Louis Collins (MNI); Fidel Alfaro Amargo and Stephen Smith (Oxford); Daniel Marcus (WUSTL). The authors additionally acknowledge funding sources for the specific analyses in this work, including R01 AG068206, RF1 AG056014, R01 AG474069, P30 AG072979, R01 AG070592.

CONFLICT OF INTEREST STATEMENT

Christopher G. Schwarz: receives research funding from the NIH, related and unrelated to this work. Mark Choe: reports no disclosures. Stephanie Rossi: reports no disclosures. Sandhitsu R. Das: reports no disclosures. Ranjit Ittyerah: reports no disclosures. Evan Fletcher: reports no disclosures. Pauline Maillard: reports no disclosures. Baljeet Singh: reports no disclosures. Danielle J. Harvey: receives research funding from the NIH, related and unrelated to this work, serves as statistical advisor to PLOS ONE. Ian B. Malone: is supported by grants to his institution from NIH and is an employee of the Dementia Research Centre which is supported by Alzheimer's Research UK, Brain Research Trust, and The Wolfson Foundation. Lloyd Prosser: works for the Dementia Research Centre (DRC) which is supported by: Alzheimer's Research UK; Brain Research Trust; the Wolfson Foundation. L. Prosser reports no additional disclosures to those that support the DRC. Matthew L. Senjem: reports no disclosures. Leonard C. Matoush: reports no disclosures. Chadwick P. Ward: reports no disclosures. Carl M. Prakaashana: reports no disclosures. Susan M. Landau: receives research funding from the NIH, related and unrelated to this work, is on the DSMB/SAB for KeifeRx and the NIH IPAT study, has received speaking honoraria from Eisai and IMPACT-AD, has consulted for Banner Health, Vaccinex, and has received travel funding and other research support from IMPACT AD and the Alzheimer's Association. Robert A. Koeppe: reports no disclosures. JiaQie Lee: reports no disclosures. Charles DeCarli: reports no disclosures. Michael W. Weiner: serves on Editorial Boards for Alzheimer's & Dementia, and the Journal for Prevention of Alzheimer's Disease (JPAD). He has served on Advisory Boards for Acumen Pharmaceutical, Alzheon, Inc., Amsterdam UMC; MIRIADE, Cerecin, Merck Sharp & Dohme Corp., NC Registry for Brain Health, and REGEnLIFE. He also serves on the USC ACTC grant which receives funding from Eisai. He has provided consulting to Boxer Capital, LLC, Cerecin, Inc., Clario, Dementia Society of Japan, Dolby Family Ventures, Eisai, Guidepoint, Health and Wellness Partners, Indiana University, LCN Consulting, MEDA Corp., Merck Sharp & Dohme Corp., NC Registry for Brain Health, Prova Education, T3D Therapeutics, University of Southern California (USC), and WebMD. He has acted as a speaker/lecturer for China Association for Alzheimer's Disease (CAAD) and Taipei Medical University, as well as a speaker/lecturer with academic travel funding provided by: AD/PD Congress, Amsterdam UMC, Cleveland Clinic, CTAD Congress, Foundation of Learning; Health Society (Japan), Kenes, U. Penn, U. Toulouse, Japan Society for Dementia Research, Korean Dementia Society, Merck Sharp & Dohme Corp., National Center for Geriatrics and Gerontology (NCGG; Japan), University of Southern California (USC). He holds stock options with Alzeca, Alzheon, Inc., ALZPath, Inc., and Anven. Dr Weiner received support for his research from the following funding sources: National Institutes of Health (NIH)/NINDS/National Institute on Aging (NIA), Department of Defense (DOD), California Department of Public Health (CDPH), University of Michigan, Siemens, Biogen, Hillblom Foundation, Alzheimer's Association, Johnson & Johnson, Kevin and Connie Shanahan, GE, VUmc, Australian Catholic University (HBI-BHR), The Stroke Foundation, and the Veterans Administration. Clifford R. Jack, Jr.: reports no disclosures. William J. Jagust: Jagust receives research funding from the NIH and Genentech, has consulted for Lilly, Biogen, Clario and Eisai and holds equity in Molecular Medicine and Optoceutics. Paul A. Yushkevich: receives research funding from the NIH, related and unrelated to this work. Duygu Tosun: receives research funding from the National Institute on Aging (NIH). Author disclosures are available in the supporting information.

CONSENT STATEMENT

This study used only previously de-identified data from ADNI, and therefore it is not human subjects research. All human subjects provided informed consent for their original ADNI data acquisitions.

REFERENCES

1. Schwarz CG, Kremers WK, Therneau TM, et al. Identification of anonyvmous MRI research participants with face recognition software. *N Engl J Med*. 2019.
2. Provenzano D, Loew MH, Goyal S, Rao YJ. Unmasking a privacy concern: potential identification of patients in an immobilization mask from 3-dimensional reconstructions of simulation computed tomography. *Practical Radiation Oncology*. 2021:S1879850021002356. doi:10.1016/j.prro.2021.09.006
3. Schwarz CG, Kremers WK, Lowe VJ, et al. Face recognition from research brain PET: an unexpected PET problem. *Neuroimage*. 2022;258:119357. doi:10.1016/j.neuroimage.2022.119357
4. Schwarz CG, Kremers WK, Arani A, et al. A face-off of MRI research sequences by their need for de-facing. *Neuroimage*. 2023;276:120199. doi:10.1016/j.neuroimage.2023.120199
5. Prior FW, Brunsden B, Hildebolt C, et al. Facial recognition from volume-rendered magnetic resonance imaging data. *IEEE Trans Inf Technol Biomed*. 2009;13:5-9. doi:10.1109/TITB.2008.2003335
6. Mazura JC, Juluru K, Chen JJ, Morgan TA, John M, Siegel EL. Facial recognition software success rates for the identification of 3D surface reconstructed facial images: implications for patient privacy and security. *J Digit Imaging*. 2012;25:347-351. doi:10.1007/s10278-011-9429-3

15525279, 2024, 11, Downloaded from https://alz-journals.onlinelibrary.wiley.com/doi/10.1002/alz.14303, Wiley Online Library on [17/06/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

7. Alfaro-Almagro F, Jenkinson M, Bangerter NK, et al. Image processing and Quality Control for the first 10,000 brain imaging datasets from UK Biobank. *Neuroimage*. 2018;166:400-424. doi:10.1016/j.neuroimage.2017.10.034
8. Schwarz CG, Kremers WK, Wiste HJ, et al. Changing the face of neuroimaging research: comparing a new MRI de-facing technique with popular alternatives. *Neuroimage*. 2021;231. doi:10.1016/j.neuroimage.2021.117845
9. Fonov VS, Collins LD. BIC defacing algorithm. *bioArxiv*. 2018. doi:10.1101/275453
10. Tremblay-Mercier J, Madjar C, Das S, et al. Open science datasets from PREVENT-AD, a longitudinal cohort of pre-symptomatic Alzheimer's disease. *Neuroimage Clin*. 2021;31:102733. doi:10.1016/j.nicl.2021.102733
11. Milchenko M, Marcus D. Obscuring surface anatomy in volumetric imaging data. *Neuroinformatics*. 2013;11:65-75. doi:10.1007/s12021-012-9160-3
12. Marcus DS, Harms MP, Snyder AZ, et al. Human Connectome Project informatics: quality control, database services, and data visualization. *Neuroimage*. 2013;80:202-219. doi:10.1016/j.neuroimage.2013.05.077
13. Marinescu RV, Oxtoby NP, Young AL, et al. TADPOLE Challenge: accurate Alzheimer's disease prediction through crowdsourced forecasting of future data. *Predict Intell Med*. 2019;11843:1-10. doi:10.1007/978-3-030-32281-6_1
14. Fischl B. FreeSurfer. *Neuroimage*. 2012;62:774-781. doi:10.1016/j.neuroimage.2012.01.021
15. Freeborough PA, Fox NC. The boundary shift integral: an accurate and robust measure of cerebral volume changes from registered repeat MRI. *IEEE Trans Med Imaging*. 1997;16:623-629. doi:10.1109/42.640753
16. Vemuri P, Senjem ML, Gunter JL, et al. Accelerated vs. unaccelerated serial MRI based TBM-SyN measurements for clinical trials in Alzheimer's disease. *Neuroimage*. 2015;113:61-69. doi:10.1016/j.neuroimage.2015.03.026
17. Yushkevich PA, Pluta JB, Wang H, et al. Automated volumetry and regional thickness analysis of hippocampal subfields and medial temporal cortical structures in mild cognitive impairment: automatic Morphometry of MTL Subfields in MCI. *Hum Brain Mapp*. 2015;36:258-287. doi:10.1002/hbm.22627
18. Fletcher E, DeCarli C, Fan AP, Knaack A. Convolutional neural net learning can achieve production-level brain segmentation in structural magnetic resonance imaging. *Front Neurosci*. 2021;15:683426. doi:10.3389/fnins.2021.683426
19. Fletcher E, Carmichael O, DeCarli C, MRI non-uniformity correction through interleaved bias estimation and B-spline deformation with a template. 2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, San Diego, CA: IEEE; 2012, p. 106-9. doi:10.1109/EMBC.2012.6345882
20. Fletcher E, Singh B, Harvey D, Carmichael O, DeCarli C. Adaptive image segmentation for robust measurement of longitudinal brain tissue change. *EMBC*. 2012;2012:5319-5322. doi:10.1109/EMBC.2012.6347195
21. Maillard P, Lu H, Arfanakis K, et al. Instrumental validation of free water, peak-width of skeletonized mean diffusivity, and white matter hyperintensities: markVCID neuroimaging kits. *Alz & Dem Diag Ass & Dis Mo*. 2022;14:e12261. doi:10.1002/dad2.12261
22. Landau S, Jagust W. ADNI Florbetapir processing methods. *ADNI PET Image Analysis*. 2015.
23. Li X, Morgan PS, Ashburner J, Smith J, Rorden C. The first step for neuroimaging data analysis: dICOM to NIfTI conversion. *J Neurosci Methods*. 2016;264:47-56. doi:10.1016/j.jneumeth.2016.03.001
24. Buimer EEL, Schnack HG, Caspi Y, et al. De-identification procedures for magnetic resonance images and the impact on structural brain measures at different ages. *Hum Brain Mapp*. 2021:1-13. doi:10.1002/hbm.25459
25. de Sitter A, Visser M, Brouwer I, et al. Facing privacy in neuroimaging: removing facial features degrades performance of image analysis methods. *Eur Radiol*. 2020;30:1062-1074. doi:10.1007/s00330-019-06459-3
26. Rubbert C, Wolf L, Turowski B, et al. Impact of defacing on automated brain atrophy estimation. *Insights Imaging*. 2022;13:54. doi:10.1186/s13244-022-01195-7
27. Gao C, Landman BA, Prince JL, Carass A. Reproducibility evaluation of the effects of MRI defacing on brain segmentation. *J Med Imag*. 2023;10. doi:10.1117/1.JMI.10.6.064001
28. Silva JM, Guerra A, Silva JF, Pinho E, Costa C, Face De-Identification Service for Neuroimaging Volumes. Proceedings—IEEE Symposium on Computer-Based Medical Systems 2018;2018-June:141-5. doi:10.1109/CBMS.2018.00032
29. Bischoff-Grethe A, Ozyurt IB, Busa E, et al. A technique for the deidentification of structural brain MR images. *Hum Brain Mapp*. 2007;28:892-903. doi:10.1002/hbm.20312
30. Schwarz CG, Kremers WK, Weigand SD, et al. Effects of de-facing software mri_reface on utility of imaging biomarkers used in Alzheimer's disease research. *NeuroImage: Clinical*. 2023;40:103507. doi:10.1016/j.nicl.2023.103507

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Schwarz CG, Choe M, Rossi S, et al. Implementation and validation of face de-identification (de-facing) in ADNI4. *Alzheimer's Dement*. 2024;20:8048–8061. https://doi.org/10.1002/alz.14303

APPENDIX: COLLABORATORS

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 link to 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.

15525279, 2024, 11, Downloaded from https://alz-journals.onlinelibrary.wiley.com/doi/10.1002/alz.14303, Wiley Online Library on [17/06/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License