Received: 7 February 2025 | Revised: 13 August 2025 | Accepted: 21 August 2025

DOI: 10.1002/alz.70712

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

RESEARCH ARTICLE

Development and evaluation of image preprocessing pipelines for the Centiloid method on Down Syndrome data

Weiquan Luo[1] | **Davneet S Minhas**[2] | **Ethan D Rubenstein**[3] | **David N Situ**[1] | **Sarah K Royse**[2] | **Beau M Ances**[4] | **Bradley T Christian**[5,6] | **Ann D Cohen**[7] | **Benjamin L Handen**[8] | **William E Klunk**[9] | **Dana L Tudorascu**[7] | **Shahid Zaman**[10] | **Charles M Laymon**[1,2] | **Alzheimer's Biomarkers Consortium–Down Syndrome (ABC-DS) Investigators**

[1]Department of Bioengineering, University of Pittsburgh, Swanson School of Engineering, Pittsburgh, Pennsylvania, USA

[2]Department of Radiology, University of Pittsburgh, PET Center, Presbyterian Hospital, Pittsburgh, Pennsylvania, USA

[3]School of Computing and Information, University of Pittsburgh, Information Sciences Building, Pittsburgh, Pennsylvania, USA

[4]Department of Neurology, Washington University, St. Louis, Missouri, USA

[5]Department of Psychiatry, Waisman Center, University of Wisconsin-Madison, Madison, Wisconsin, USA

[6]Department of Medical Physics, University of Wisconsin-Madison, Madison, Wisconsin, USA

[7]Department of Psychiatry, University of Pittsburgh, Pittsburgh, Pennsylvania, USA

[8]Department of Psychiatry, University of Pittsburgh, Quantum Building, Pittsburgh, Pennsylvania, USA

[9]Department of Psychiatry, University of Pittsburgh, 1422 UPMC Western Psychiatric Hospital, Pittsburgh, Pennsylvania, USA

[10]Department of Psychiatry, University of Cambridge, Douglas House, Cambridge, UK

Correspondence
Charles M Laymon, Department of Radiology, UPMC Presbyterian Hospital, 9th Floor B-wing, 200 Lothrop Street, Pittsburgh, PA 15213, USA.
Email: cml14@pitt.edu

Funding information
National Institute on Aging and the National Institute for Child Health and Human Development, Grant/Award Numbers: U01 AG051406, U01 AG051412, U19 AG068054; Investigation of Co-occurring conditions across the Lifespan to Understand Down Syndrome; Alzheimer's Disease Research Centers Program, Grant/Award Numbers: P50 AG008702, P30 AG062421, P50 AG16537, P50 AG005133, P50 AG005681, P30 AG062715, P30 AG066519, P30 AG066468, P30 AG072973; Eunice Kennedy Shriver Intellectual and Developmental Disabilities Research Centers Program, Grant/Award

Abstract

BACKGROUND: Centiloid provides a standardized process to quantify brain amyloid in which a subject's T1 magnetic resonance imaging (MRI) and amyloid positron emission tomography (PET) scans are registered and warped to Montreal Neurological Institute 152 space using prescribed procedures. The method has a high failure rate in Down syndrome (DS) subjects from the Neurodegeneration in Aging Down Syndrome (NiAD) project. We evaluate imaging preprocessing methods (PMs) to improve the DS success rate.

METHODS: PMs were constructed from combinations of image origin reset, filtering, MRI bias correction, and MRI skull stripping. Centiloid results were evaluated for adherence to standards using The Global Alzheimer's Association Interactive Network dataset. PMs were also evaluated using the NiAD dataset to judge their suitability for the DS population. DS PM evaluation procedures were developed corresponding to those specified for non-DS populations.

Number: P50 HD105353; National Center for Advancing Translational Sciences, Grant/Award Numbers: UL1 TR001873, UL1 TR002373, UL1 TR001414, UL1 TR001857, UL1 TR002345, UL1 TR002366; National Centralized Repository for Alzheimer Disease and Related Dementias, Grant/Award Number: U24 AG21886; DS-Connect; National Institute on Aging; National Institute for Health and Care Research (NIHR) Cambridge Biomedical Research Centre, Grant/Award Number: NIHR203312; NIHR Applied Research Collaboration East of England

RESULTS: Five accepted PMs improved the Centiloid-processing success rate in the DS cohort from 61.3% to 95.6%.

DISCUSSION: The identified combinations of preprocessing steps substantially improved the success rate of Centiloid processing in DS.

KEYWORDS
amyloid imaging, Centiloid, Down Syndrome, image processing, MRI, PET

Highlights

- Image preprocessing pipeline is proposed for Centiloid analysis of DS.
- Preprocessing pipelines are evaluated for adherence to Centiloid standards.
- Pipelines are evaluated for improvement in yield of usable imaging data.
- Preprocessing of amyloid imaging data resulted in a large yield improvement.

1 | BACKGROUND

Alzheimer's disease (AD) pathogenesis is widely believed to be driven by processes that lead to the aggregation and deposition of the amyloid beta peptide (Aβ) in the brain, and the presence of Aβ plaques is a defining pathological feature of AD. In humans, the production of Aβ is brought about by the proteolysis of amyloid precursor protein (APP), which is coded by the APP gene located on chromosome 21.[1] People with Down syndrome (DS) are subject to overproduction of APP and are predisposed to increased Aβ accumulation and early onset AD, which has a histopathology similar to sporadic, late-onset AD.[2,3] Here, we report on a study using data from the Neurodegeneration in Aging Down Syndrome (NiAD) project, which was a multisite study of DS participants that included an investigation of the natural history of amyloid burden and has now transitioned into the Alzheimer's Biomarker Consortium-Down syndrome (ABC-DS).[4–10] The ABC-DS includes a core that actively engages with and recruits from diverse populations.

To attain maximum utility of datasets formed by combining data from different sites, it is necessary to minimize (standardize) and then account for (harmonize) site-specific differences in characteristics that affect quantitation. Various standardization and harmonization strategies have been developed for PET.[11,12] The Centiloid method[13] was established specifically to provide standardization in quantifying brain Aβ burden. In addition to specifying a scale for reporting results, it provides a method of image processing, and this paper was motivated specifically by our experience with this aspect of the procedure.

Within NiAD, we observed a higher Centiloid processing failure rate in participants with DS compared to individuals without DS (Figure 1). Presumably, this is partially due to anatomic differences between DS participants and the non-DS brain images used in the formation of the Montreal Neurological Institute (MNI) 152 template, which is critical to the Centiloid process. Relative to participants without DS, the DS brain has an overall smaller volume and a disproportionately smaller cerebellar volume.[14] In addition, the frequency of serious motion artifacts observed in T1 magnetic resonance imaging (MRI) scans in this cohort is relatively high. Further, NiAD image data were acquired at five collection sites using five different combinations of MRI and PET scanners with different characteristics, introducing site- and scanner-specific variability (Table S1).

The Centiloid methodology comprises several steps: image processing, quantification of tracer uptake (i.e., application of regions of interest [ROIs]), and conversion of results to the Centiloid scale. Each of these should be examined and optimized for a distinctive cohort. Here, we address the first of these steps, image processing, upon which the others depend. Although extended procedures for validating fundamentally new pipelines (e.g., those that use different templates or registration, warping algorithms, quantitation ROIs, or positron emission tomography [PET]-quantitation harmonization strategies) for producing Centiloid results have been specified,[13] our focus was to develop alternative *preprocessing* pipelines that improve the Centiloid success rate but that remain within the bounds of the standard Centiloid method. This research consisted of two major components: (1) validation of alternative PET and magnetic resonance image processing pipelines on the publicly available standard Centiloid [^{11}C]PiB 50- to 70-min dataset and (2) evaluation of the same pipelines on a set of participants acquired under the NiAD study with newly developed rigorous quality assurance (QA) criteria. As a consequence, the procedures that were developed, while motivated by needs related to our DS cohort, are not specific to DS.

In what follows, we develop and evaluate pipelines with the goal of having a set of preprocessing tools that can be used interchangeably. For this work, it is necessary to process all scans in our datasets through all pipelines to judge the quality of each result and to compare variances in Centiloid values between the pipelines. Having established interchangeability, in typical practice, an analyst would not run all pipelines for an analysis but would stop once a scan was successfully processed. In the discussion, we outline a potential order for running the pipelines based on their expected success rate.

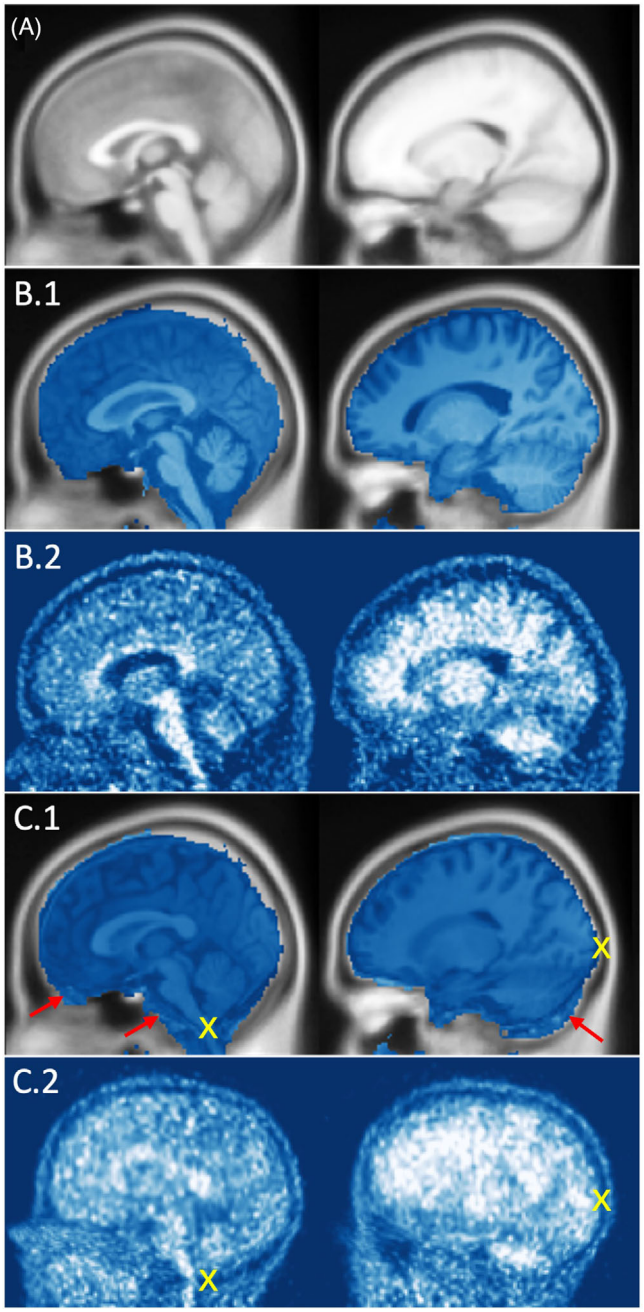


FIGURE 1 Two sagittal slices of examples of warped magnetic resonance images cropped with template brain mask and positron emission tomography (PET) images in Montreal Neurological Institute (MNI) template space: (A) MNI template image as reference (gray); (B) successfully processed scan set; (B.1) T1 magnetic resonance image (blue) with good spatial normalization overlaid on template image (gray); (B.2) PET image with good coregistration and good normalization to corresponding T1 magnetic resonance image; (C) unsuccessfully processed scan; (C.1) magnetic resonance image (blue) with failed spatial normalization overlaid on template image in which brain was shrunk on normalization and non-brain tissue was aligned to template brain tissue marked with red arrows; (C.2) PET image with failed coregistration and normalization to corresponding T1 magnetic resonance image due to the PET image being pitched relative to the magnetic resonance image. The green crosses shown in C.1 and C.2 mark the same final coordinates in each image and illustrate a mismatch between the two images.

2 | METHODS

2.1 | Demographics and imaging data

The standard Centiloid dataset used in component 1 (validation of alternative PET and magnetic resonance image processing pipelines on the publicly available standard Centiloid [^{11}C]PiB 50- to 70-min dataset) was downloaded from The Global Alzheimer's Association Interactive Network (GAAIN) website.[15] The PET scans in the standard Centiloid dataset include a group of 34 young control participants (age range 22 to 43 years) and 45 participants diagnosed with AD (age range 50 to 89 years).[13] The version downloaded from GAAIN is the recommended dataset for evaluating Centiloid procedures, in which each scan consists of a single 50- to 70-min post-injection frame. A T1-weighted MRI scan for each participant is included in the dataset. The dataset is accompanied by gold standard Centiloid values that were determined in the original analysis of the data.[13]

For component 2 (evaluation of pipelines on a set of participants acquired under the NiAD study), NiAD scans (N = 318), available at the LONI website (https://ida.loni.usc.edu), were used. Of these, 271 were from participants with DS, and 47 were scans of sibling controls without DS. All downloaded datasets included a T1 structural MRI scan for anatomic reference and PET images acquired 50 to 70 min after injection in four 5-min frames. The LONI website contains several versions of each scan with various levels of processing. Although it would have been possible to use preprocessed scans available at LONI for this study, we also performed separate analyses, unrelated to the current work, in which we used the original (not preprocessed) scans. Thus, in the current evaluation, we also used the original multiframe scans as uploaded by the various sites. PET scans were inspected for frame-to-frame motion, and, when necessary, motion was corrected using PMOD software (PMOD Technologies, Zurich, Switzerland). The motion-corrected frames were then averaged to produce a single-frame image representing [^{11}C]PiB uptake over the 50- to 70-min post-injection period.

2.2 | Centiloid processing pipeline

Scans from both components were processed through a standard pipeline starting with the magnetic resonance and single-frame [^{11}C]PiB images. Scans were also reprocessed through the alternative pipelines implemented after various processing steps, described below. In all, for each scan, seven versions of processing pipelines were implemented.

2.2.1 | Standard pipeline

The standard Centiloid processing pipeline[13] (also see the Centiloid processing document available at GAAIN[15]) includes steps to align each PET image to its corresponding T1-weighted MRI image and then to warp the MRI image (and co-warp the PET image) to the MNI152

template. Tracer concentrations are then extracted from the warped PET images using ROIs defined in MNI template space.

The original Centiloid procedure relies on the MNI152 template and is carried out using SPM8.[16] Briefly, to prepare for the warping process, the subject MRI scan is first rigidly aligned to the MNI 152 template using the SPM8 "Coregister" function with default parameters. For the reference image (alignment target), we use an in-house white matter (WM) and gray matter (GM) only template (TPM_MRI.nii), constructed from the SPM8/MNI152 standard tissue priors (spm8/tpm/TPM.nii). The process for generating this synthetic, MNI-aligned template was carried out using the image math function of SPM. It was calculated as (GM tissue probability map) + 1.7× (WM probability map).

The precise value of the factor (1.7) differentiating WM from GM is not critical in this application, as the (default) objective function used in the coregistration operation is normalized mutual information.[17] The purpose of using the synthetic MNI T1 reference image is to provide a brain-tissue-only registration target that is less likely to be confounded by extra-brain tissue (e.g., neck); additionally, the synthetic T1 target does well with skull-stripped magnetic resonance images.

Following rigid registration of the subject MRI image to the MNI 152 template via the synthetic target, the subject PET image is rigidly registered to the realigned subject MRI image and, consequently, to the MNI152 template, again using the SPM8 "Coregister" function. Next, a non-linear warping of the MRI image to the MNI 152 template via the SPM8 Unified Segmentation method is performed, and the resultant deformation parameters are applied to the PET image to warp it into alignment with the template.[18]

Tracer concentrations are extracted from the warped PET image using the standard Centiloid cortical ROI and whole cerebellum ROI. The cortical-to-cerebellum tissue ratio is then converted to Centiloid units via a specified linear transformation.[13] Our in-house implementation of the original Centiloid procedure is considered the standard pipeline for both GAAIN and NiAD datasets.

2.2.2 | Alternative pipelines

The first alternative (A_0) differs from the other alternatives in that it is a slight modification of the first step of the standard Centiloid procedure rather than a true preprocessing pipeline. However, for the convenience of presentation, we categorize this procedure with the other alternative pipelines. The standard procedure calls for manually setting the image origins to the anterior commissure before performing registrations.[15] With an eye toward pipeline automation, in the A_0 pipeline, we simply replace the manual method with an in-house method, in which the origin is set at the center of the field of view (FOV). The GAAIN dataset was processed using the A_0 pipeline, and the resulting values were compared to the published Centiloid values. As we show in the results, we found only slight differences between the standard pipeline with a manual origin setting and the A_0 pipeline with the automatic origin setting.

The other five alternative pipelines all include A_0 with one or more additional processing steps. Three processing steps in various combinations were investigated for their ability to improve the success rate of the Centiloid procedure for this group of participants (*1: smoothing [SMO]*): A smoothed PET image was produced by filtering with SPM8 using a Gaussian kernel of 2 × 2 × 2 mm full width at half maximum to moderate noise effects during PET-MRI registration. This kernel size was chosen to be small compared to PET resolution, but it has not been optimized. The smoothed PET image was used only for the coregistration. The unsmoothed PET was then realigned using the parameters generated and was used for any following steps, including the extraction of concentrations (*2: N4 bias field correction [N4]*): ANTsPy's N4 bias field correction with default settings was used to reduce MRI image intensity inhomogeneities.[19,20] (*3: MRI skull-stripping [SS]*): Before registration, the MRI scan is cropped by a subject-specific brain mask that combines the binarized probability tissue maps of gray matter, white matter, and cerebrospinal fluid (CSF) generated by SPM8 unified segmentation.[18] In pipelines involving modifications to the MRI scan (N4 and SS), the modified MRI scan was used for both the PET-MRI registration and for the MRI image warping.

Combinations of SMO, N4, or both with A_0 are designated as A_SMO, A_N4, and A_N4_SMO, respectively. Our preliminary examination of the pipelines indicated that skull stripping worked poorly if N4 bias correction was not performed first. Therefore, in the formal analysis, skull stripping of the MRI image was not performed without N4 bias correction. Adding SS to A_N4 and A_N4_SMO yielded A_N4_SS and A_N4_SMO_SS pipelines, respectively. With a defined ordering of processing steps, this yielded a total of six candidate alternative pipelines based on the standard pipeline. We note that while smoothed versions of the PET scan are used in some pipelines, the registration/warping parameters generated are applied to the original, unsmoothed PET, which is then used for the determination of the Centiloid value.

2.3 | Evaluation

2.3.1 | Overview

One purpose of our evaluation was to establish the interchangeability of the alternative pipelines with the standard Centiloid pipeline. The GAAIN dataset is the standard for evaluating new pipelines, and the reproducibility of a newly implemented unmodified pipeline is established through a comparison of outcomes determined by the new method to the published values. Thus, as a step in evaluating reproducibility, we compared results from all our pipelines with the published results using the established criteria for acceptance (Section 2.3.2).

Additional evaluation is required to validate the process for the multisite DS dataset. The philosophy for the evaluation of pipeline reproducibility using the NiAD data was similar to that of the GAAIN analysis. However, the evaluation schema for GAAIN data cannot be directly applied to the NiAD, where the ground-truth Centiloid

outcomes are not available. Thus, to determine the degree to which the various pipelines are interchangeable, that is, to determine the degree of reproducibility of results between different pipelines, it was first necessary to judge whether a particular scan had been successfully processed by a particular pipeline. Consequently, as described in the following subsections, there are two parts to the evaluation procedure: (1) quality assurance (QA), where the quality of the transformation of the scans to the MNI-152 template is determined using visual quality assurance criteria, and (2) reproducibility, where the Centiloid outcome measures from each alternative pipeline are compared to the results of the standard pipeline. Reproducibility comparisons are made using only those scans that passed the QA procedure. Thus, a fundamental difference between the GAAIN and NiAD evaluations is that, whereas a visual QA assessment is, of necessity, applied to the NiAD data, no such evaluation is applied to the GAAIN analysis.

2.3.2 | Quality assurance (NiAD)

QA methodology is not specified in the original Centiloid publication. In the Centiloid processing guideline (https://www.gaaindata.org/data/centiloid/Centiloid_Processing.docx), a manual quality check is recommended after each registration and spatial normalization step. However, applying the prescribed manual quality check at multiple points throughout the Centiloid processing pipeline is burdensome, especially for very large datasets, and was not feasible in this study given the multitude of alternative Centiloid pipelines tested. Therefore, we developed and implemented a new QA procedure performed solely on the final spatially normalized PET and MRI scans.

Two QA checks were performed, facilitated by custom image display software. (1) The MRI-TEMPLATE registration QA check evaluates the spatial normalization of MRI image to the template space. The criteria for success were good conformity of gray matter/CSF boundaries, minimal misclassification of brain tissue as meninges or vice versa, and good matching of the cerebellum. (2) The PET-MRI registration QA check evaluates PET-to-MRI registration in template space. The criteria for success were that there must be minimal left-right (LR), anterior-posterior (AP), or inferior-superior (IS) shifts visible at the edges of the brain. Specifically, a good match of the cerebellum, corpus callosum, brainstem, and ventricles is needed.

A team of five raters was employed in this study. A series of training sessions was held to develop competency and interrater reliability in the application of the pass/fail standards before the actual evaluation. The team manager compiled a set of demonstration scans and instructed the team on how to review and rate the quality for both QA tasks. The manager also assembled a set of 200 scans, divided into four batches, for training and reader reliability optimization. The five raters independently scored a batch of scans. Following this, a review session was held to discuss any scans that did not have consensus. This pattern was followed for all four batches of training data.

A four-point rating system corresponding to quality classifications of definite failure, likely failure, likely pass, and definite pass was used to evaluate both QA components. For the ease of evaluation workflow, the classifications of each QA category were contracted to a binary outcome (i.e., pass or fail).

Actual QA evaluation involved rating the NiAD registration for all seven pipelines of the 318 scans for a total of 2226 image sets. For efficiency, each QA component was evaluated in a single phase. In phase 1, each image set was randomly assigned to two readers for evaluation of the MRI-template registration quality. If both readers rated the normalization as a failure (consensus fail), then the PET-MRI registration quality was not assessed in phase 2. Phase 2 focused on the within-subject PET-MRI linear registration. For a scan to pass, both readers had to rate it as a pass (consensus pass). Compared to phase 2, the weaker pass requirement for phase 1 allows more leeway for subjectivity in the judgment of the warping evaluation. Assignments were balanced so that each reader pair rated approximately the same number of image sets. Scan assignments were not maintained between the two phases but were chosen randomly for each. Each reader rated approximately 890 scans in phase 1 and 843 in phase 2.

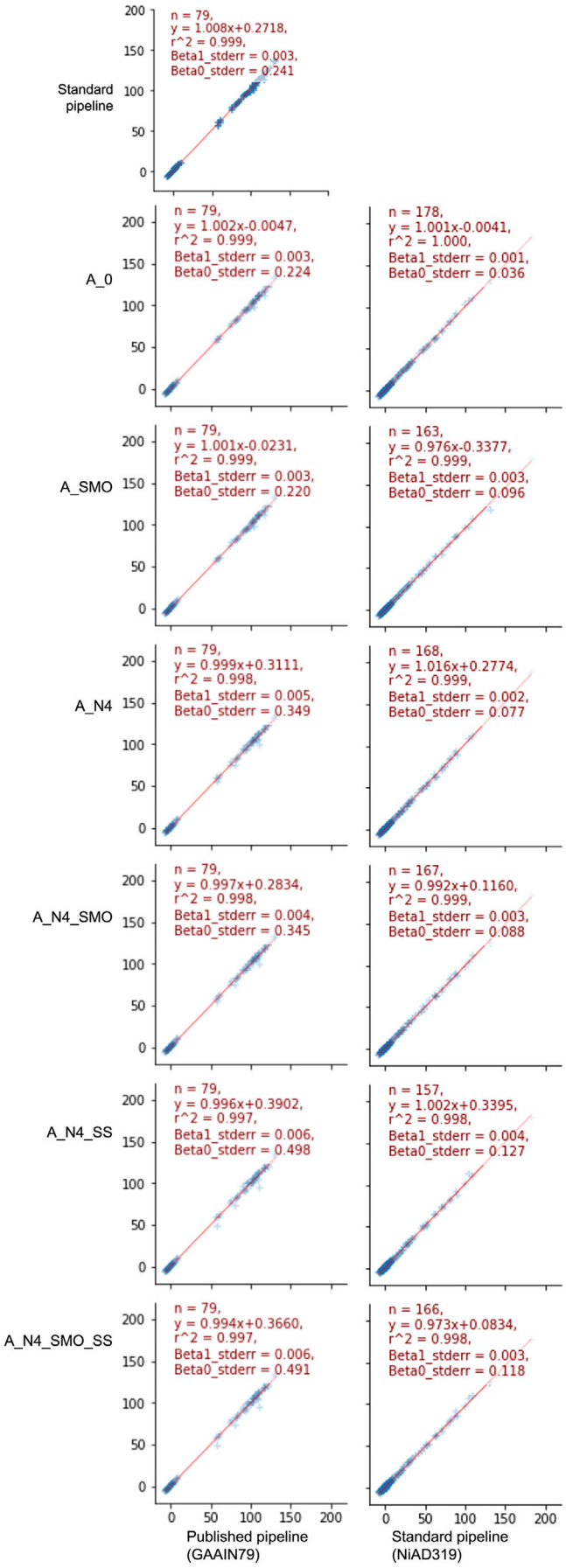
As described, PET-MRI registration and MRI-template warping were evaluated for each scan. We did not formally evaluate pipeline performance (e.g., skull strip quality) at a scale finer than this as part of the current study.

2.3.3 | Reproducibility (GAAIN and NiAD)

The reproducibility of each of the various pipelines was evaluated by comparing Centiloid results from each alternative pipeline to reference Centiloid values. For the GAAIN data, linear regressions were performed between results from each of the seven pipelines and the published Centiloid reference values, in each case using all 79 GAAIN participants. For the NiAD data, values from the six alternative pipelines were compared to values from our standard pipeline, which are taken to be the reference values. In the NiAD analysis, only those scans that passed the QA procedure for both the pipeline being assessed and the standard pipeline were included in the reproducibility evaluation. No such cuts were made in the GAAIN analysis given the ground-truth published values.

The criteria for passing the pipeline reproducibility evaluation were the reproducibility criteria as described in Section 2.3 of Klunk et al. 2015: (1) slope between 0.98 and 1.02; (2) intercept between −2 and 2 Centiloid; and (3) $R^2 > 0.98$ based on a linear regression between Centiloid outcomes from each of the pipelines and reference values. Additionally, for the NiAD cohort, we performed correlation analyses stratified by Centiloid value, where subjects were classified as low amyloid (Centiloid < 12.2), intermediate (Centiloid between 12.2 and 24.4), and high (Centiloid > 24.4).

Method agreement was further quantified for each pipeline pair by calculating bootstrapped intraclass correlation coefficients (ICCs) (two-way mixed effects, absolute agreement, single rater model, calculated using the ICC(2, 1) formula[21]) and Bland–Altman plots.



3 | RESULTS

3.1 | Pipeline reproducibility regression statistics

Linear regression results are depicted in Figure 2, which shows subject Centiloid values from each pipeline plotted against the corresponding Centiloid reference value. Data points for all 79 GAAIN subjects are shown in the GAAIN plots, whereas for NiAD results are plotted for NiAD subjects that passed the QA procedure for both pipelines (i.e., the standard pipeline and the particular preprocessing pipeline). The corresponding regression statistics for each pipeline compared to the reference are shown in Table 1. The first plot in Figure 2 and the first row of Table 1 provide the results of our standard pipeline using the GAAIN data set. Tables S2–S4 contain regression results for the NiAD cohort when stratified by amyloid burden.

All pipelines passed the reproducibility criteria using the GAAIN data. All pipelines passed reproducibility criteria for the NiAD data except for A_N4_SMO_SS (slope = 0.973). Although the slope for this pipeline for NiAD data is outside the lower limit by less than 0.01, we reject it for the NiAD data because it technically does not meet the reproducibility criteria. The slope fit limits are specified to two decimal places,[13] so we consider pipeline A_SMO (slope = 0.976) to have passed. Results from the ICC analysis indicate near-perfect agreement for every pipeline pair, with all ICCs being above 0.998. More details can be found in Table S5.

Additional results describing pipeline correspondence in terms of mean Centiloid difference between pipeline pairs and standard deviation of Centiloid difference are presented in Tables S6 and S7, respectively. These results are illustrated by Bland–Altman plots shown in Figure S1.

3.2 | NiAD QA and pipeline benefit

Table 2 summarizes quality assurance results and rater comparisons on a per-scan basis. The initial set of 2226 MRI/PET scan pairs (318 unique scans processed by seven pipelines) was reduced to 2140 after evaluation of the MRI-template normalization, where 86 scans were given a fail rating by both readers.

Table 3 summarizes quality assurance results and scan acceptance on a per-pipeline basis, with percentages of the full set of 318 scans. The first column lists the number of scans for which at least one rating was "pass" in the MRI-template quality check, which was the necessary

FIGURE 2 Scatter plots of Centiloid values from each of the pipelines (*y*-axis) versus the reference Centiloid values (*x*-axis). Results are shown for the GAAIN dataset (first column) and NiAD dataset (second column). Centiloid reference values for the GAAIN data are those published on the GAAIN website; those for the NiAD data are from our standard pipeline. Regression models are presented in each subplot with additional statistics: sample size (*n*), R-squared (*r*^2), standard error of estimate slope (Beta1_sterr), and intercept (Beta0_sterr). Regression fits are shown as the line in each of the scatter plots. NiAD, Neurodegeneration in Aging Down Syndrome.

TABLE 1 Linear regression parameters and R^2 between each of the pipelines and the reference values for the GAAIN dataset and the NiAD dataset.

	GAAIN79 (ref: published values)			NiAD319 QA (ref: standard pipeline)		
Pipeline	Slope	Intercept	R^2	Slope	Intercept	R2
standard pipeline	1.008	0.272	1	-	-	-
A_0	1.002	−0.005	0.999	1.001	−0.004	1.000
A_SMO	1.001	−0.023	0.999	0.976	−0.338	0.999
A_N4	0.999	0.311	0.998	1.016	0.277	0.999
A_N4_SMO	0.997	0.283	0.998	0.992	0.116	0.999
A_N4_SS	0.996	0.390	0.997	1.002	0.339	0.998
A_N4_SMO_SS	0.994	0.366	0.997	0.973	0.083	0.998

Note: Reference values for the GAAIN data set are those published on the GAAIN website. Reference values for the NiAD data are from the standard pipeline.
Abbreviation: NiAD, Neurodegeneration in Aging Down Syndrome.

TABLE 2 Comparison of NiAD scan pass rates for five raters using the two categories of QA criteria.

	MRI -Template	PET-MRI
Total scans evaluated	2226	2104
Number of scans evaluated by a single rater	889–892	839–845
Pass rate of each of five raters	91.0%, 94.9%, 93.9%, 91.6%, 94.1%	80.2%, 84.2%, 78.4%, 85.5%, 73.3%

Note: The table shows the percentage of passing scans on a per-rater basis for both of the evaluations conducted. Percentages listed in the PET-MRI column are based on the number of scans that survived the MRI-Template assessment and made it to the PET-MRI evaluation component. As explained in the text, each scan was evaluated using multiple pipelines, and the pass rates shown here are across all pipelines and scans, that is, the pass rates include multiple evaluations of the same scan in different pipelines. Thus, the values presented are lower than the percentage of scans that passed at least one pipeline. The rating comparison of individual rater-pairs is shown in Appendix D.
Abbreviation: MRI, magnetic resonance imaging; NiAD, Neurodegeneration in Aging Down Syndrome; PET, positron emission tomography; QA, quality assurance.

TABLE 3 The summary of quality assurance results and scan acceptance on a per-pipeline basis.

	Passed in MRI-template	Final pass
Standard pipeline	286 (89.9%)	195 (61.3%)
A_0	287 (90.3%)	212(66.7%)
A_SMO	284 (89.3%)	212(66.7%)
A_N4	309 (97.2%)	227(71.4%)
A_N4_SMO	309 (97.2%)	245(77.0%)
A_N4_SS	315 (99.1%)	213(67.0%)
A_N4_SMO_SS (rejected)	314 (98.7%)	249(78.3%)
6 accepted pipelines	317 (99.7%)**	303 (95.2%)**
All 7 pipelines	317 (99.7%)**	304 (95.6%)**

**Passed at least one processing pipeline

condition for the scan to be evaluated on the PET-MRI quality check. In the second column, the final pass category lists the number of scans that received at least one pass in the MRI-template category, followed by a consensus pass in the PET-MRI check. Figure S2 illustrates the level of agreement between raters in these analyses.

To be counted as a pass in the grouped result of "six accepted pipelines" and "all seven pipelines," a scan must have had a pass in at least one of the included pipelines. The availability of five alternative pipelines in addition to the standard pipeline allowed us to increase our final pass rate from 61.3% to 95.2% in the NiAD group of subjects. As apparent in Table 3, the standard pipeline is effective for MRI-template registrations, and most of the benefit of additional pipelines was due to the improved success rate in the PET-MRI registration. Had we included pipeline A_N4_SMO_SS, which technically falls outside of the reproducibility criteria based on NiAD data, it would have generated only a 0.4 percentage point improvement in the final pass rate.

3.3 | Pipeline and site comparisons

Figure 3A shows the pipeline-specific success rate in two QA phases on NiAD data. Our lowest-yielding pipeline was the standard pipeline with an overall success rate of 61.3%. Our highest-yielding pipeline A_N4_SMO_SS resulted in an overall success rate of 78.3%. However, A_N4_SMO_SS was just below the Centiloid slope reproducibility criteria of 0.98 (slope = 0.973). The best pipeline that satisfied all Centiloid reproducibility criteria, A_N4_SMO, resulted in an overall success rate of 77.0%. The A_N4_SS pipeline had a lower passing rate than A_N4_SMO and A_N4_SMO _SS for all sites.

Figure 3B further shows the pipeline-specific success rate stratified by acquisition site. The results of the MRI-template phase were good for all pipelines, except for site 1, which is due to lower-quality MRI scans compared to other sites. N4 bias correction was beneficial to the data collected from all sites, but especially for site 1.

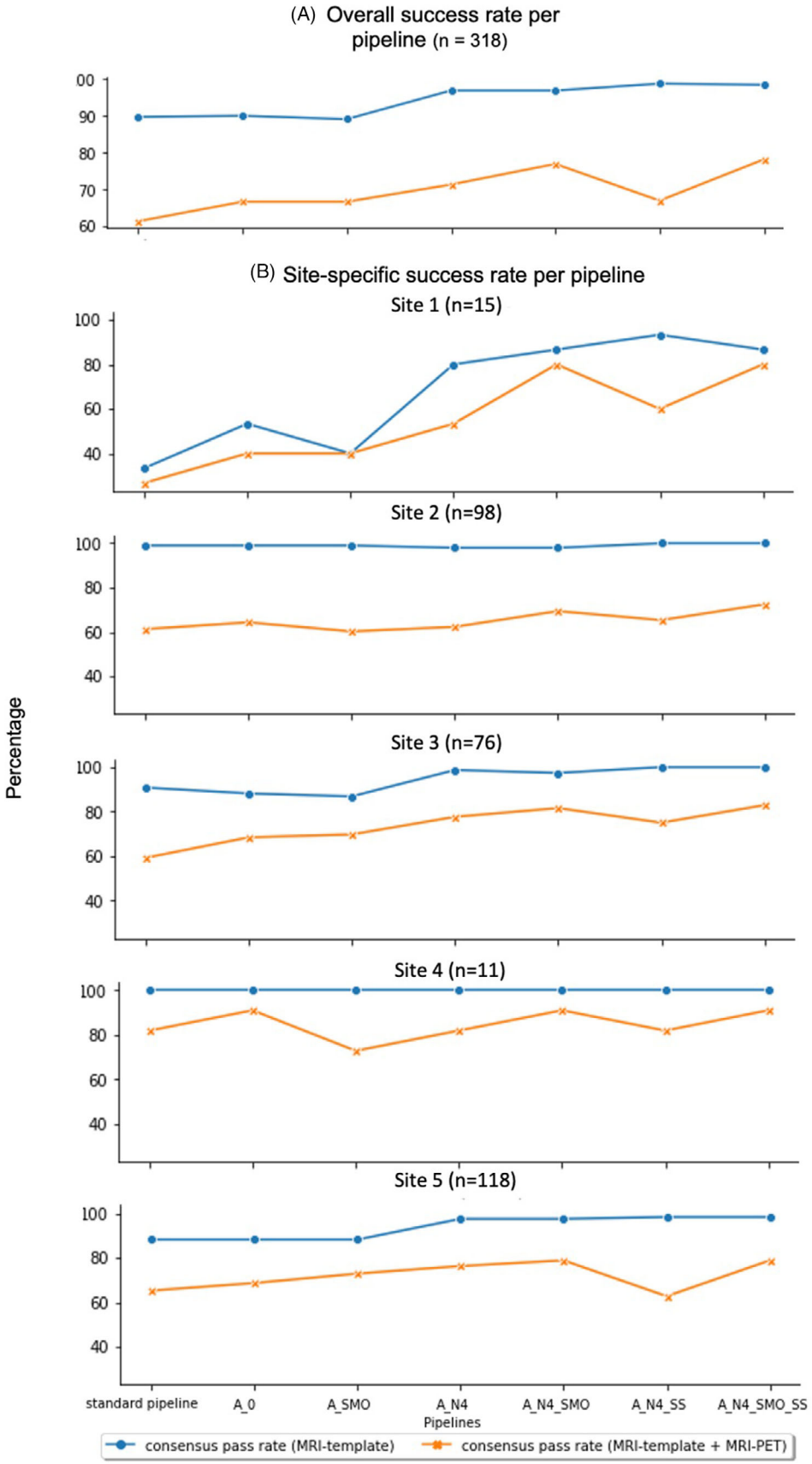


FIGURE 3 Overall (A) and site-specific (B) consensus pass rate per pipeline. Each subplot shows pass rate per pipeline using MRI-template criteria (blue) and the final pass rate using both MRI-template and PET-MRI criteria (yellow) for five different sites. MRI, magnetic resonance imaging; PET, positron emission tomography.

4 | DISCUSSION

The direct motivation for this work was to identify methods for increasing the success rate of Centiloid processing of multisite DS scans. Within the context of our validation procedure, we processed every scan set in our data collection through every pipeline. However, as discussed in what follows, this is not a necessary or recommended procedure for practical pipeline use.

4.1 | Use of pipelines

The regression analysis using GAAIN data shows that our in-house standard pipeline and six alternative pipelines achieve the established reproducibility criteria within acceptable limits. Importantly, by meeting reproducibility standards, each of the pipelines can, by definition, be used interchangeably in non-DS populations. In an independent dataset comprising mostly participants with DS, we employed an extensive visual QA evaluation before comparing our standard pipeline using the NiAD data. As a result of the evaluation, the A_N4_SMO_SS pipeline narrowly missed the slope reproducibility criteria of 0.98 with a slope of 0.973 in the DS evaluation. Failure to meet reproducibility criteria suggests that this pipeline is not suitable for this DS cohort. It also suggests that the reproducibility of a pipeline is subject to the cohort group. We do note, however, that A_N4_SMO_SS passed reproducibility standards in using the GAAIN dataset and can be considered a validated pipeline for non-DS subjects.

There is a clear advantage in having several pipelines available. Processing a dataset with all pipelines and combining results from accepted interchangeable pipelines could yield the highest possible success rate. Based on the NiAD component in this study, the use of a single accepted pipeline results in a success rate ranging from 61.3% to 77.0%, or a failure rate of about one-fourth of these data. However, combining results from multiple pipelines, validated to produce interchangeable results based on the Centiloid reproducibility criteria, leads to a 95.6% success rate.

In practice, running all seven pipelines on all imaging data within a cohort is burdensome. A potential strategy for the application of alternative pipelines would be to first use the standard pipeline, which is the simplest procedure. If this is unsuccessful, we propose to apply pipelines for a DS cohort dataset in the following order based on the reproducibility criterion and the gain in success rate with respect to the standard pipeline:

1. Standard pipeline
2. A_N4_SMO: results in the best gain (+15.7%) in success rate among all accepted pipelines with a good R-squared (0.999).
3. A_0: is closest to the standard pipeline with a perfect R^2 (1.000) and a moderate gain (+5.4%).
4. A_N4: has the second-highest gain (+10.1%),
5. A_N4_SS and A_SMO: has moderate gain (+5.7%, +5.4%),

Including the standard pipeline, a total of six pipelines are listed in the foregoing scheme. The A_N4_SMO_SS pipeline is rejected and not included for DS cohort data processing because its slope (0.973) technically fails the reproducibility criteria.

As the various pipelines can be considered interchangeable, the preceding priority list ordering is only for efficiency and is based on our cohort of subjects. A different ordering may be more appropriate in other cases and could be site-specific in multisite studies. For example, the characteristics of the MRI scans from site 1 (which tended to have motion artifacts) resulted in a high MRI/template failure rate in the standard pipeline and in fact drove the priority list toward a preference for running the A_N4_SMO pipeline first. Without site 1, the impetus for running this pipeline first, while still high, is reduced.

4.2 | Processing step analysis

A common step in all analyses is the initial alignment of the subject PET and MRI scans. We introduced a simple procedure for this with the goal of producing automated pipelines. Although other methods are possible, such as using the center of mass of each image as the origin, we chose to align the centers of the images. This approach works well for appropriately positioned patients and may be less subject to confounds from neck tissue found in this cohort, particularly if the MRI and PET scanners have different FOV lengths. In practice, manual alignment could be substituted in cases where the automated alignment is problematic.

The skull-stripping step completely removes information outside of the brain in MRI. Early in this work, we found that the skull-stripping step had to be coupled to the N4 step to produce success rates at the same level as other pipelines. Application of skull stripping without bias field correction was found to improperly remove regions of true brain, causing cascading errors in the later steps. Therefore, the performance of skull stripping without N4 was not formally investigated.

Markiewicz et al.[22] suggested that post-processing or PET reconstruction steps, apart from those used for quantitation, could be employed to facilitate PET/MRI scan registration. We adopted this strategy for the SMO pipeline in which the PET image was smoothed specifically for use in the PET/MRI registration step. We observed that smoothing the PET was a double-edged sword. It reduced the high-frequency noise in PET images but also degraded high-frequency features such as tissue boundaries that are valuable to the PET-MRI registration process. This is evident in the mixed results on the success rate of using the A_SMO pipeline. The A_SMO pipeline improved the success rates for sites 1, 3, and 5 but worsened them for sites 2 and 4. The negative effect was mitigated by coupling smoothing and N4. The SMO element of the preprocessing pipelines has adjustable parameters (the smoothing kernel). We chose the $2 \times 2 \times 2$ mm smoothing kernel used in the SMO operation to be relatively small compared to the scanner resolution. Given the foregoing discussion and our mixed success rate with A_SMO, a conceivably fruitful extension of this work

would be to optimize the kernel, possibly in a site-specific way. Further, it must be recognized that the use of different PET tracers, different PET reconstruction methods, or different reconstruction parameters (e.g., number of subsets), as well as image smoothing carried out, for example, for multisite harmonization[23] (not done in this analysis), will modify the effect of the SMO pipeline element, which, ideally, should be optimized for each situation.

The effect of N4 was found to vary by site. Applying N4 processing on MRI scans improved the consensus pass rate for the template/MRI normalization for three sites (1, 3, and 5) but not for two sites (2 and 4). This is likely related to differences in the character of bias field inhomogeneity observed in MRI images from various sites (Figure S3).

4.3 | Observations

PET-MRI registration is a substantial cause of pipeline failure. Various factors can affect the uncertainty of the registration.[22] There are many existing approaches to the problem, such as alternative registration methods and removing outlier voxel intensity. Currently, we have not evaluated the performance of the various candidate methods of registration with the DS cohort. However, our work shows that relatively simple steps can improve PET-MRI registrations and improve the success rate of PET-MRI registration by pipeline aggregation.

We reexamined scans from the ABC-DS cohort that resulted in points lying outside of the 95% prediction band (not shown) of our linear regression analyses. As described, all scans passed quality checks for both the standard pipeline and the alternative pipeline under consideration. Nevertheless, our inspection suggests that differences in the results between the two regressed pipelines for these points may frequently be due to subtle differences in the registration, particularly observable in the cerebellum.

4.4 | Additional considerations

Centiloid processing of DS subjects presents challenges. The current work provides preprocessing strategies that, in aggregate, substantially improve the success rate in this group. Nevertheless, there may be additional steps that could reduce the already small failure rate. As the focus of this first study was to develop automated pipelines, performance quality was assessed at the two major end-of-pipeline checkpoints, PET/MRI registration and MRI/MNI warping, which are common to all. In cases where an image set fails in all pipelines, it would be useful to identify points in the various pipelines where an intervention might improve results.

While we have employed approaches that fall within the standard Centiloid method, it is likely that alternative strategies that venture outside of the standard procedure would also yield improved results and perhaps reduce the need for multiple parallel pipelines. As noted, a formal process for validating fundamentally new pipelines has been specified.[13] Given that the PET-to-MRI registration was the most frequent cause of failure, an updated registration method is an obvious direction for investigation.[22] Additional improvements could result from a revised MRI normalization procedure. In particular, a DS-specific template method is being developed.

5 | CONCLUSION

Compatibility between the original Centiloid procedure, our standard pipeline, and five modified pipelines has been established for DS data using the newly defined QA criteria. Additional processing steps greatly improved the success rate of Centiloid processing for DS imaging. The success rate for processing NiAD scans increased from 61.3% (n = 195) with the standard Centiloid pipeline to 95.6% (n = 304) by combining the standard pipeline and five verified interchangeable modifications.

ACKNOWLEDGMENTS

The ABC-DS is funded by the National Institute on Aging and the National Institute for Child Health and Human Development (U01 AG051406, U01 AG051412, U19 AG068054) and the Investigation of Co-occurring conditions across the Lifespan to Understand Down Syndrome (NIH INCLUDE Project). The work contained in this publication was also supported through the following National Institutes of Health (NIH) programs: The Alzheimer's Disease Research Centers Program (P50 AG008702, P30 AG062421, P50 AG16537, P50 AG005133, P50 AG005681, P30 AG062715, P30 AG066519, P30 AG066468 and P30 AG072973), the Eunice Kennedy Shriver Intellectual and Developmental Disabilities Research Centers Program (P50 HD105353), the National Center for Advancing Translational Sciences (UL1 TR001873, UL1 TR002373, UL1 TR001414, UL1 TR001857, UL1 TR002345, UL1 TR002366), the National Centralized Repository for Alzheimer Disease and Related Dementias (U24 AG21886), and DS-Connect® (The Down Syndrome Registry) supported by the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD). In Cambridge, UK, this research was supported by the National Institute for Health and Care Research (NIHR) Cambridge Biomedical Research Centre (NIHR203312) and the NIHR Applied Research Collaboration East of England.

The authors are grateful to the ABC-DS study participants, their families and care providers, and the ABC-DS research and support staff for their contributions to this study. This manuscript has been reviewed by ABC-DS investigators for scientific content and consistency of data interpretation with previous ABC-DS study publications. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH, the NIHR, or the UK Department of Health and Social Care.

CONFLICTS OF INTEREST STATEMENT

GE Healthcare holds a license from the University of Pittsburgh to the PiB PET technology used in this study and pays license fees directly to the University of Pittsburgh. Author WEK, as a co-inventor of PiB, receives payments from the university based on these license fees as

governed by university policy. Funding for the study was provided by the sources listed.

CONSENT STATEMENT

Data for this project were obtained from the NiAD partition of the LONI data archive. Each NiAD site acquires data with Institutional Review Board approval and with written informed consent.

REFERENCES

1. Kang J, Lemaire H-G, Unterbeck A, et al. The precursor of Alzheimer's disease amyloid A4 protein resembles a cell-surface receptor. *Nature*. 1987;325(6106):733-736.
2. Wisniewski KE, Wisniewski HM, Wen GY. Occurrence of neuropathological changes and dementia of Alzheimer's disease in Down's syndrome. *Ann. Neurol*. 1985;17(3):278-282.
3. Mann DMA, Yates PO, Marcyniuk B. Alzheimer's presenile dementia, senile dementia of Alzheimer type and Down's Syndrome in middle age form an age related continuum of pathological changes. *Neuropathol Appl Neurobiol*. 1984;10(3):185-207.
4. Lao PJ, Handen BL, Betthauser TJ, et al. Longitudinal changes in amyloid positron emission tomography and volumetric magnetic resonance imaging in the nondemented Down syndrome population. *Alzheimers Dement*. 2017;9(1):1-9.
5. Lao PJ, Betthauser TJ, Hillmer AT, et al. The effects of normal aging on amyloid-β deposition in nondemented adults with Down syndrome as imaged by carbon 11–labeled Pittsburgh compound B. *Alzheimers Dement*. 2016;12(4):380-390.
6. Cohen AD, McDade E, Christian B, et al. Early striatal amyloid deposition distinguishes Down syndrome and autosomal dominant Alzheimer's disease from late-onset amyloid deposition. *Alzheimers Dement*. 2018;14(6):743-750.
7. Hartley SL, Handen BL, Devenny DA, et al. Cognitive functioning in relation to brain amyloid-β in healthy adults with Down syndrome. *Brain*. 2014;137(9):2556-2563.
8. Zammit MD, Laymon CM, Betthauser TJ, et al. Amyloid accumulation in Down syndrome measured with amyloid load. *Alzheimers Dement*. 2020;12(1):e12020.
9. Handen BL, Lott IT, Christian BT, et al. The Alzheimer's Biomarker Consortium-Down Syndrome: Rationale and methodology. *Alzheimers Demen*. 2020;12(1):e12065.
10. Handen BL, Mapstone M, Hartley S, et al. The Alzheimer's Biomarker Consortium-Down Syndrome (ABC-DS): a 10-year report. *Alzheimers Dement*. 2025;21(5):e70294.
11. Akamatsu G, Tsutsui Y, Daisaki H, Mitsumoto K, Baba S, Sasaki M. A review of harmonization strategies for quantitative PET. *Ann. Nuclear Med*. 2023;37(2):71-88.
12. Minhas DS, Yang Z, Muschelli J, et al. Statistical methods for processing neuroimaging data from two different sites with a Down Syndrome population application. *Inform Process Manage Uncertain Knowledge-Based Syst*. 2020;1239:367-379.
13. Klunk WE, Koeppe RA, Price JC, et al. The Centiloid Project: Standardizing quantitative amyloid plaque estimation by PET. *Alzheimers Dementia*. 2015;11(1):1.
14. Pinter JD, Eliez S, Schmitt JE, Capone GT, Reiss AL. Neuroanatomy of Down's Syndrome: A High-Resolution MRI Study. *Amer J Psych*. 2001;158(10):1659-1665.
15. Network TGAsAI. Centiloid project [cited 2023 Dec 10]. Available from: https://gaain.org/centiloid-project
16. Functional Imaging Laboratory UQSIoN. SPM8 [cited 2023 Dec 10]. Available from: https://www.fil.ion.ucl.ac.uk/spm/software/spm8/
17. Studholme C, Hill DLG, Hawkes DJ. An overlap invariant entropy measure of 3D medical image alignment. *Pattern Recogn*. 1999;32(1):71-86.
18. Ashburner J, Friston KJ. Unified segmentation. *NeuroImage*. 2005;26(3):839-851.
19. Tustison NJ, Avants BB, Cook PA, et al. N4ITK: Improved N3 Bias Correction. *IEEE Trans Med Imag*. 2010;29(6):1310-1320.
20. Avants BB, Tustison NJ, Stauffer M, Song G, Wu B, Gee JC. The Insight ToolKit image registration framework. *Front Neuroinform*. 2014;8:44.
21. Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *J Chiroprac Med*. 2016;15(2): 155-163.
22. Markiewicz PJ, Matthews JC, Ashburner J, et al. Uncertainty analysis of MR-PET image registration for precision neuro-PET imaging. *NeuroImage*. 2021;232: 117821.
23. Joshi A, Koeppe RA, Fessler JA. Reducing between scanner differences in multi-center PET studies. *NeuroImage*. 2009;46(1):154-159.

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: Luo W, Minhas DS, Rubenstein ED, et al. Development and evaluation of image preprocessing pipelines for the Centiloid method on Down Syndrome data. *Alzheimer's Dement*. 2025;21:e70712.
https://doi.org/10.1002/alz.70712

APPENDIX

Collaborators

Alzheimer's Biomarker Consortium–Down Syndrome (ABC-DS) Investigators

Beau M. Ances, MD PhD; Howard F. Andrews, PhD; Karen Bell, MD; Rasmus M. Birn, PhD; Adam M. Brickman, PhD; Peter Bulova, MD; Jeff Burns, MD; Amrita Cheema, PhD; Kewei Chen, PhD; Bradley T. Christian, PhD; Isabel Clare, PhD; Ann D. Cohen, PhD; Eric W. Doran, MS; Tatiana M. Foroud, PhD; Benjamin L. Handen, PhD; Jordan Harp, PhD; Sigan L. Hartley, PhD; Elizabeth Head, PhD; Denise Head, PhD; Christy Hom, PhD; Lawrence Honig, MD; Milos D. Ikonomovic, MD; Sterling C Johnson, PhD; M. Ilyas Kamboh, PhD; David Keator, PhD; Julia K. Kofler, MD; William Charles Kreisl, MD; Sharon J. Krinsky-McHale, PhD; Florence Lai, MD; Patrick Lao, PhD; Charles Laymon, PhD; Joseph Hyungwoo Lee, PhD; Ira T. Lott, MD; Victoria Lupson, PhD; Mark Mapstone, PhD; Davneet Singh Minhas, PhD; Neelesh Nadkarni, MD; Sid O'Bryant, PhD; Deborah Pang, MPH; Melissa Petersen, PhD; Julie C. Price, PhD; Lauren Ptomey, PhD; Margaret Pulsifer, PhD; Michael S. Rafii, MD PhD; Herminia Diana Rosas, MD; Frederick Schmitt, PhD; Nicole Schupf, PhD; Wayne P. Silverman, PhD; Dana L. Tudorascu, PhD; Rameshwari Tumuluru, MD; Badri Varadarajan, PhD; Michael A. Yassa, PhD; Shahid Zaman, MD PhD; Fan Zhang, PhD