

RESEARCH

Open Access



Centiloid values from deep learning-based CT parcellation: a valid alternative to freesurfer

Yeo Jun Yoon^{1†}, Seungbeom Seo^{1†}, Sangwon Lee², Hyunkeong Lim², Kyobin Choo³, Daesung Kim⁴, Hyunkyung Han⁴, Minjae So¹, Hosung Kang⁵, Seongjin Kang², Dongwoo Kim^{2,6}, Young-gun Lee⁷, Dongho Shin⁸, Tae Joo Jeon^{2,9*} and Mijin Yun^{2*}

Abstract

Background Amyloid PET/CT is essential for quantifying amyloid-beta (A β) deposition in Alzheimer's disease (AD), with the Centiloid (CL) scale standardizing measurements across imaging centers. However, MRI-based CL pipelines face challenges: high cost, contraindications, and patient burden. To address these challenges, we developed a deep learning-based CT parcellation pipeline calibrated to the standard CL scale using CT images from PET/CT scans and evaluated its performance relative to standard pipelines.

Methods A total of 306 participants (23 young controls [YCs] and 283 patients) underwent 18 F-florbetaben (FBB) PET/CT and MRI. Based on visual assessment, 207 patients were classified as A β -positive and 76 as A β -negative. PET images were processed using the CT parcellation pipeline and compared to FreeSurfer (FS) and standard pipelines. Agreement was assessed via regression analyses. Effect size, variance, and ROC analyses were used to compare pipelines and determine the optimal CL threshold relative to visual A β assessment.

Results The CT parcellation showed high concordance with the FS and provided reliable CL quantification ($R^2 = 0.99$). Both pipelines demonstrated similar variance in YCs and effect sizes between YCs and ADCL. ROC analyses confirmed comparable accuracy and similar CL thresholds, supporting CT parcellation as a viable MRI-free alternative.

Conclusions Our findings indicate that the CT parcellation pipeline achieves a level of accuracy similar to FS in CL quantification, demonstrating its reliability as an MRI-free alternative. In PET/CT, CT and PET are acquired sequentially within the same session on a shared bed and headrest, which helps maintain consistent positioning and adequate spatial alignment, reducing registration errors and supporting more reliable and precise quantification.

Keywords Alzheimer's disease, Centiloid, Florbetaben, Amyloid imaging

[†]Yeo Jun Yoon and Seungbeom Seo contributed equally to this work as co-first authors.

*Correspondence:
Tae Joo Jeon
TJEONNM@yuhs.ac
Mijin Yun
YUNMIJIN@yuhs.ac

Full list of author information is available at the end of the article



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Background

Amyloid positron emission tomography / computed tomography (PET/CT) imaging has emerged as a pivotal tool for quantifying amyloid-beta ($A\beta$) deposition in vivo, with significant implications for understanding Alzheimer's disease (AD) pathology [1]. This quantification enables the early detection of $A\beta$ plaques and facilitates tracking their progression over time, which is critical for diagnosis, prognosis, and monitoring treatment effects [2, 3]. The importance of accurately quantifying $A\beta$ burden has become even more pronounced with the emergence of monoclonal antibodies such as lecanemab and donanemab, which target aggregated $A\beta$ to reduce its accumulation in the brain [4, 5]. However, variability in tracers, imaging protocols, and analysis methods hampers cross-center comparisons and meta-analyses. This variability also complicates the establishment of universal diagnostic thresholds, highlighting the critical need for standardized approaches to ensure consistency and broader applicability in clinical and research settings.

The Centiloid (CL) project was initiated to standardize the quantification of $A\beta$ burden across PET/CT imaging centers, addressing variability in tracers and analysis methods [6]. By establishing a standardized pipeline for cortical standardized uptake value ratios (SUVr) and enabling cross-calibration of different tracers, this approach integrates data onto a unified scale. This framework facilitates comparisons across studies and enhances the utility of amyloid PET imaging in both clinical and research settings.

Traditional PET quantification has relied on magnetic resonance imaging (MRI) as an anatomical reference, but this requirement presents several challenges. MRI may be unavailable due to cost, patient burden, or contraindications such as claustrophobia or medical implants, potentially excluding certain participants and introducing selection bias [7, 8]. To address these limitations, several studies have explored MRI-free quantification methods, utilizing PET-based or CT-based approaches by developing standardized templates and aligning PET images with these templates for quantification [9, 10]. Previous studies have noted that template-based methods may face limitations in patients with significant brain atrophy, including reduced spatial accuracy and increased susceptibility to partial volume effects, particularly when spatial normalization fails to account for severe anatomical distortions [8, 11].

To address these issues, several studies have attempted CT-based segmentation using deep learning models trained on SPM-derived tissue labels, direct application of SynthSeg to CT, or CT-to-MR translation followed by MR segmentation. However, these approaches either lack fine anatomical detail, are prone to segmentation failures, or require multi-step pipelines [12–14]. In contrast,

a recently developed deep learning-based (DL-based) CT parcellation method provides a robust alternative to traditional MRI-based approaches by directly segmenting CT images according to Desikan-Killiany-Tourville (DKT) atlas, thereby overcoming the limitations of previous methods [15]. This DL-based approach, which employs three independent 2D UNet-based segmentation models to capture multi-view anatomical features, provides accurate definitions of cortical gyri and maintains reliable quantification even in severely deformed brains, such as those with atrophy or hydrocephalus. The acquisition of PET and CT in the same session using a shared scanner bed and headrest helps maintain spatial alignment, thereby reducing errors and improving reliability. This approach overcomes the limitations of MRI-free methods while remaining accessible and cost-effective, making amyloid PET quantification feasible even for individuals unable to undergo MRI. It may serve as a useful alternative when MR-based processing is not feasible and anatomical variation is a concern.

In this study, we implemented a DL-based CT parcellation pipeline calibrated to the standard CL scale using CT images from PET/CT scans and determined a cut-off value based on agreement with visual reads. We then evaluated its performance by comparing it with CL scales derived from previously established pipelines, assessing agreement with visual reads, and comparing effect sizes between AD and non-AD groups.

Methods

Participants

We recruited 23 young controls (YCs) and 283 patients who presented to our memory clinic with memory complaints and underwent ^{18}F -florbetaben (FBB) PET/CT and MRI between January 2017 and December 2022. YCs, defined as those younger than 45 years, had no history of neurologic or psychiatric disorders, confirmed normal cognitive function after a clinical evaluation, and were $A\beta$ -negative on PET imaging. Of the 283 patients, 207 were $A\beta$ -positive and 76 were $A\beta$ -negative based on visual assessment. In the $A\beta$ -negative patients ($n=76$), 70 patients had amnesic mild cognitive impairment (aMCI), and 6 patients had dementia. Among $A\beta$ -positive patients ($n=207$), 55 patients had AD dementia, and 152 patients had aMCI. Patients who were $A\beta$ -positive and diagnosed with either dementia or aMCI were collectively referred to as the AD-related cognitive impairment (ADCI) group. AD dementia was diagnosed based on the National Institute on Aging-Alzheimer's Association criteria for probable AD [16]. Participants with aMCI met the criteria proposed by Petersen et al. [17]: (1) subjective memory complaints, (2) relatively normal performance in other cognitive domains, (3) normal activities of daily living (ADL), (4) objective memory impairment below -1.5

SD on either verbal or visual memory tests, and (5) not demented. The following exclusion criteria were applied to the 283 patients: (1) presence of structural brain lesions – such as brain tumors, multiple lacunar infarcts, or cerebral infarctions on MRI – and (2) a time gap of over one year between the FBB PET/CT scans and MRI.

Image acquisition

FBB PET/CT scans were acquired using the Discovery 600 system (GE Healthcare, Milwaukee, WI). A total of 300 MBq of FBB was administered intravenously, and PET imaging was performed 90 min post-injection for 20 min, followed by a CT scan for attenuation correction. The CT scan was performed in spiral mode with acquisition parameters of 0.8 s per rotation, 120 kVp, 200 mA, a slice thickness of 3.27 mm, a collimation of 10 mm, and a table feed of 9.375 mm per rotation. The PET images (matrix size: 256×256 , voxel size: $0.98 \text{ mm} \times 0.98 \text{ mm} \times 3.27 \text{ mm}$) were reconstructed using the ordered subset expectation maximization (OSEM) algorithm with 4 iterations and 32 subsets, incorporating attenuation, scatter, and random corrections. A Gaussian smoothing filter with a full-width at half-maximum (FWHM) of 4 mm was applied to the reconstructed images. Non-contrast T1-weighted MRI scans were obtained on a 3-T MRI system (Ingenia CX or Achieva; Philips Healthcare, Best, Netherlands) with a matrix of 256×256 , a field of view ranging from 230 to 240 mm, a slice thickness of 1.2 mm, and a repetition time/echo time (TR/TE) of 6.9/3.2 milliseconds.

Visual assessment

Visual assessment of FBB PET/CT scans was performed by two expert nuclear medicine physicians (M. Yun and T.J. Jeon), both blinded to the amyloid PET quantification results. The evaluation employed a standardized approach based on a regional cortical tracer uptake (RCTU) scoring system applied to four brain regions including lateral temporal cortex, frontal cortex, posterior cingulate cortex/precuneus, and parietal cortex. The RCTU assessments were then integrated into the overall brain amyloid plaque load (BAPL) score, leading to a binary classification of the scans [18, 19]. A BAPL score of 1 (no A β load) is classified as an A β -negative PET scan, while BAPL scores of 2 (minor A β load) and 3 (significant A β load) are classified as A β -positive PET scans.

Data processing - Centiloid pipeline

All PET and MRI images were processed using an implementation adapted from the standard CL pipeline with SPM 12, based on procedures described in [5]. In brief, each subject's MRI image was segmented and normalized to MNI space using SPM12, and the PET image was co-registered to the corresponding MRI and subsequently

normalized using the same transformation parameters. The standard cortical and whole cerebellum reference regions of interest (ROIs) were downloaded from the GAAIN website (www.gaain.org). Finally, we calculated the CL pipeline-derived SUVR ($\text{SUVR}_{\text{standard}}$) in MNI space using the standard cortical target region and the whole cerebellum as reference region.

Local pipeline validation procedures

As described in section 2.2.2 of Klunk [6], the first step of a Level 2 analysis begins with a replication of the Level-1 analysis. To validate our local pipeline, we downloaded the reference ^{11}C -Pittsburgh Compound-B (PiB) PET (50–70 min) dataset from GAAIN. Briefly, this dataset includes 34 young controls (YC) and 45 older adults with clinically diagnosed AD, which serve as the CL scale's anchor points of 0 and 100 units, respectively. We processed the downloaded dataset using our local pipeline, yielding local Level-1 outcomes. We then compared these local CL values against published CL values. All processing was performed using SPM12 (version 12; <https://www.fil.ion.ucl.ac.uk/spm/software/spm12/>).

Data processing - FreeSurfer pipeline

We analyzed MRIs using FreeSurfer (FS) v7.4 to generate a native-space FS atlas for each MRI (<http://surfer.nmr.mgh.harvard.edu>). We subsequently co-registered FBB PET images to their corresponding MRIs with SPM12's "Coregister: Estimate and Reslice" tool using default parameters. We then sampled the PET images to assess the mean tracer uptake in target cortical areas using a cortical mask made up of FS-defined frontal, cingulate, lateral parietal, and lateral temporal regions as previously described [20]. We then assessed the mean tracer uptake in four reference regions: the whole cerebellum (WC), cerebellar gray matter (CG), pons, and a composite reference region. The composite reference region was created by taking the unweighted average of FS-defined whole cerebellum, brainstem, and eroded subcortical white matter, following previously established methods [21, 22]. The eroded subcortical white matter mask was generated by first smoothing a binarized FS-defined subcortical white matter image to the 8mm^3 resolution of the FBB PET image, followed by thresholding at 0.70 to selectively erode white matter-defining voxels near gray matter. Using mean tracer uptake in the cortical target areas and these reference regions, we calculated FS pipeline-derived SUVRs (SUVR_{FS}) for each FBB PET scan.

Data processing - CT parcellation pipeline

We performed CT-based FBB quantification using our previously developed DL-based CT parcellation method, utilizing commercially available software (NCM-brain v2.5, Newcure M, Seoul, Korea). Briefly, the pipeline

employs three independent 2D U-Net segmentation networks—one each for axial, sagittal, and coronal CT slices—each configured with 256 feature channels per layer to accommodate CT's lower soft-tissue contrast [15]. The outputs of these three models are ensembled to generate a complete 1 mm isotropic 3D brain parcellation, which is then co-registered to PET space. PET images were sampled using the same cortical mask as in the FS pipeline to calculate mean tracer uptake in the target cortical areas. We used the same four reference regions employed in the FS pipeline – WC, CG, pons, and the composite reference region. Using the mean tracer uptake in the cortical target areas and these reference regions, we calculated CT parcellation pipeline-derived SUVRs (SUVR_{CT}) for each FBB PET scan. The co-registration process was omitted as FBB and CT images were acquired with consistent positioning in a single PET/CT session, allowing adequate alignment.

Centiloid conversion

We performed a regression on the corresponding SUVR_{FS} and SUVR_{CT} values. From these equations, we determined the intercept (b) and slope (m), which were then used to convert SUVR_{CT} into SUVR_{FS}. The conversion equations for SUVR_{FS} to CL, specifically for WC and composite reference region, have already been published [20]. Using these equations, the “calculated” SUVR_{FS} were subsequently converted into CL units.

For the standard CL pipeline, the FBB tracer is a well-established surrogate tracer for PiB, and the conversion equation from the CL pipeline-derived FBB-SUVR to PiB-SUVR has been validated in previous studies [23]. Thus, the FBB-SUVR_{standard} values were subsequently converted into “calculated” PiB-SUVR_{standard} values. Using Eq. 2.2.3 from Klunk et al. [6] and substituting our Level-1 YC-0 and AD-100 CL values, these “calculated” PiB-SUVR_{standard} values were converted into CL units.

$$CL = 100 \times (PiB - SUVR_{standard} - 1.012) / (2.077 - 1.012)$$

The standard CL pipeline uses WC as the reference region and generates a single CL value per patient, abbreviated as CL_{standard}. In contrast, the FS and CT parcellation pipelines, each of which uses four reference regions, produce CL values abbreviated as CL_{FS} and CL_{CT}, respectively. Specifically, the FS pipeline produces the following values: CL_{FS-WC} (WC), CL_{FS-CG} (CG), CL_{FS-pons} (pons), and CL_{FS-comp} (composite reference region); and for the CT parcellation pipeline: CL_{CT-WC} (WC), CL_{CT-CG} (CG), CL_{CT-pons} (pons), and CL_{CT-comp} (composite reference region).

Statistical analysis

Demographic characteristics, including both continuous variables and ordinal variables, were compared using the Kruskal-Wallis test, with post-hoc comparisons conducted using Dunn's test with Bonferroni correction. For two-group comparisons, the Wilcoxon rank-sum test was used. To compare values against zero, the Wilcoxon signed-rank test was performed. Correlations between SUVR values were assessed using linear regression analysis. The effect size between the ADCI and YC groups was calculated with the following equation:

$$\text{Effect Size} = (\mu_p - \mu_n) / \sqrt{(N_p \sigma_p^2 + N_n \sigma_n^2) / (N_p + N_n - 2)}$$

where μ_p , μ_n are the average SUVR in the ADCI and YC groups. σ_p^2 , σ_n^2 are the variance of the ADCI and YC groups. N_p and N_n are the number of participants in the ADCI and YC groups. Effect sizes are computed separately for each reference region within the same pipeline, and their 95% confidence intervals (CIs) were estimated using non-parametric bootstrapping with 10,000 resamples. Differences in effect sizes were considered statistically significant when the corresponding CIs did not overlap.

Within the YC group, we evaluated the variance of each imaging pipeline and reference. First, Levene's test was performed to assess homogeneity of variance among CL values obtained using WC – CL_{standard}, CL_{CT-WC}, and CL_{FS-WC}. When significant differences were detected, post-hoc pairwise F-tests were conducted. Next, pairwise F-tests were performed to compare the variances between the two reference regions within each pipeline (CL_{CT-WC} versus CL_{CT-comp} and CL_{FS-WC} versus CL_{FS-comp}). *P*-values for all pairwise comparisons were adjusted using the Bonferroni correction for multiple comparisons.

Receiver operating characteristic (ROC) curve analysis was performed to determine the CL cutoff that showed the highest agreement with the visual read. Youden's *J* index was calculated as the sum of sensitivity and specificity minus one, and the threshold maximizing this index was selected as the optimal cutoff. Classification performance was further evaluated using accuracy, sensitivity, specificity, and the area under the ROC curve (AUC). All statistical analyses were performed using the R software (version 4.0, <http://www.r-project.org>).

Results

Demographics

The baseline demographics and clinical characteristics of the participants are summarized in Table 1. There was a significant difference in age between the

Table 1 Demographics and clinical characteristics of the participants

	YC	Aβ-negative	ADCI	p-value
N	23	76	207	N/A
Age, years ^{a, b}	30.3 ± 7.98	71.8 ± 7.70	72.0 ± 8.18	< 0.001
Females, n (%)	14 (61%)	61 (80.3%)	138 (66.7%)	0.054
K-MMSE ^{a, b, c}	29.8 ± 0.491	23.8 ± 4.19	21.9 ± 4.39	< 0.001
CDR global score ^{a, b, c}				< 0.001
0	23	0	0	
0.5	0	70	151	
1	0	5	45	
≥ 2	0	1	11	

Note: Values are presented as number (%) for categorical variables, mean ± SD for continuous variables, and median (IQR) for ordinal variables (CDR score)

Abbreviations: YC, young control; Aβ, amyloid beta; ADCI, Alzheimer's disease-related cognitive impairment; K-MMSE, Korean version of the Mini-Mental State Examination; CDR, Clinical Dementia Rating

^a Significant differences between YC and Aβ-negative

^b Significant differences between YC and ADCI

^c Significant differences between Aβ-negative and ADCI

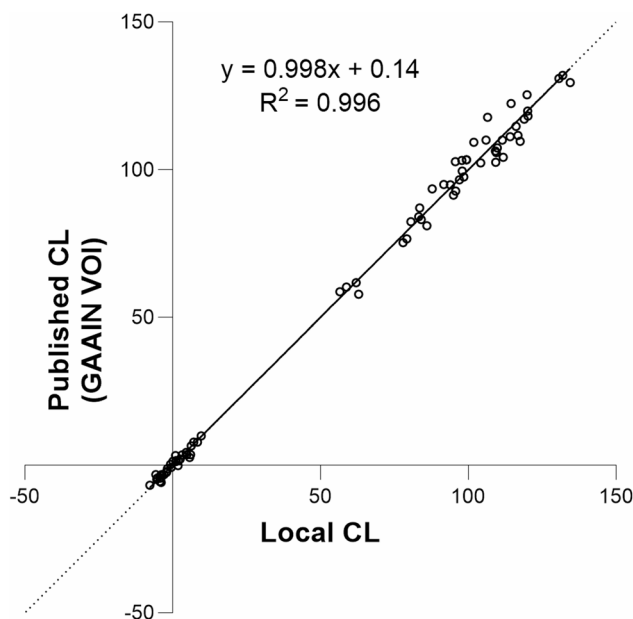


Fig. 1 Plot of CL outcomes derived from Level-1 analysis of the standard 34 YC-0 and 45 AD-100 scans vs. published CL values. Dashed unity lines have been added to facilitate visual comparison between the axes. The equation and R^2 indicate that the local Centiloid pipeline was applied correctly. Abbreviation: CL, Centiloid; VOI: volume of interest

YC group and both the Aβ-negative and ADCI groups. The Korean version of the Mini-Mental State Examination (K-MMSE) scores were highest in the order YC, Aβ-negative, and ADCI, with significant differences among all three groups. In contrast, Clinical Dementia Rating (CDR) global scores were highest in ADCI, followed by Aβ-negative and then YC, with significant differences observed across the three groups.

Local pipeline validation

Linear regression of the local standard CL level 1 outcomes against published CL outcomes yielded a fit equation with slope = 0.998, intercept = 0.140, and correlation coefficient (R^2) = 0.996 (Fig. 1). The fit exceeded the minimum specified acceptance criteria (i.e., $R^2 > 0.98$, slope between 0.98 and 1.02, and intercept between -2 and +2), confirming that the local results derived from SPM12 were comparable with published CL from GAAIN.

FreeSurfer versus CT parcellation pipeline

MRI and CT images were segmented using the FS and CT parcellation pipelines. Both pipelines demonstrated accurate anatomical segmentation, allowing reliable ROI-based quantification (Fig. 2). However, in one case, a susceptibility artifact on MRI resulted in inaccurate segmentation by FS; in another case, inaccurate co-registration between MRI and PET led to unreliable PET quantification; both cases were removed from further analysis (Fig. 3). Following segmentation, the cortical mask of the target regions was applied, and $SUVR_{CT}$, $SUVR_{FS}$ were calculated for each reference region. After plotting $SUVR_{FS}$ and $SUVR_{CT}$, we performed a regression analysis. Regression analysis between the FS pipeline and the CT parcellation pipeline demonstrated high concordance across all reference regions. Results for WC and composite reference regions are shown in Fig. 4, and results for CG and pons are provided in Supplementary Fig. 1.

Conversion equations

$SUVR_{FS}$ and $SUVR_{CT}$ were plotted and regressed against each other for WC and the composite reference region. The slope and intercept values derived from the regression equations were used to convert $SUVR_{CT}$ into “calculated” $SUVR_{FS}$. Using the previously known equation described in previous work, $SUVR_{FS}$ values were then converted into CL [20].

For the standard CL pipeline, $FBB-SUVR_{standard}$ values were converted into “calculated” $PiB-SUVR_{standard}$ values [23]:

$$PiB - SUVR_{standard} = (FBB - SUVR_{standard} - 0.39)/0.61$$

$PiB-SUVR_{standard}$ were finally converted into CL units. The final conversion equations are as summarized in Table 2.

Effect size, variability, and distribution of CL values

The variance in YC reflects the noise inherent in the quantification method and has been used to evaluate both pipelines and reference regions [6, 24]. In our

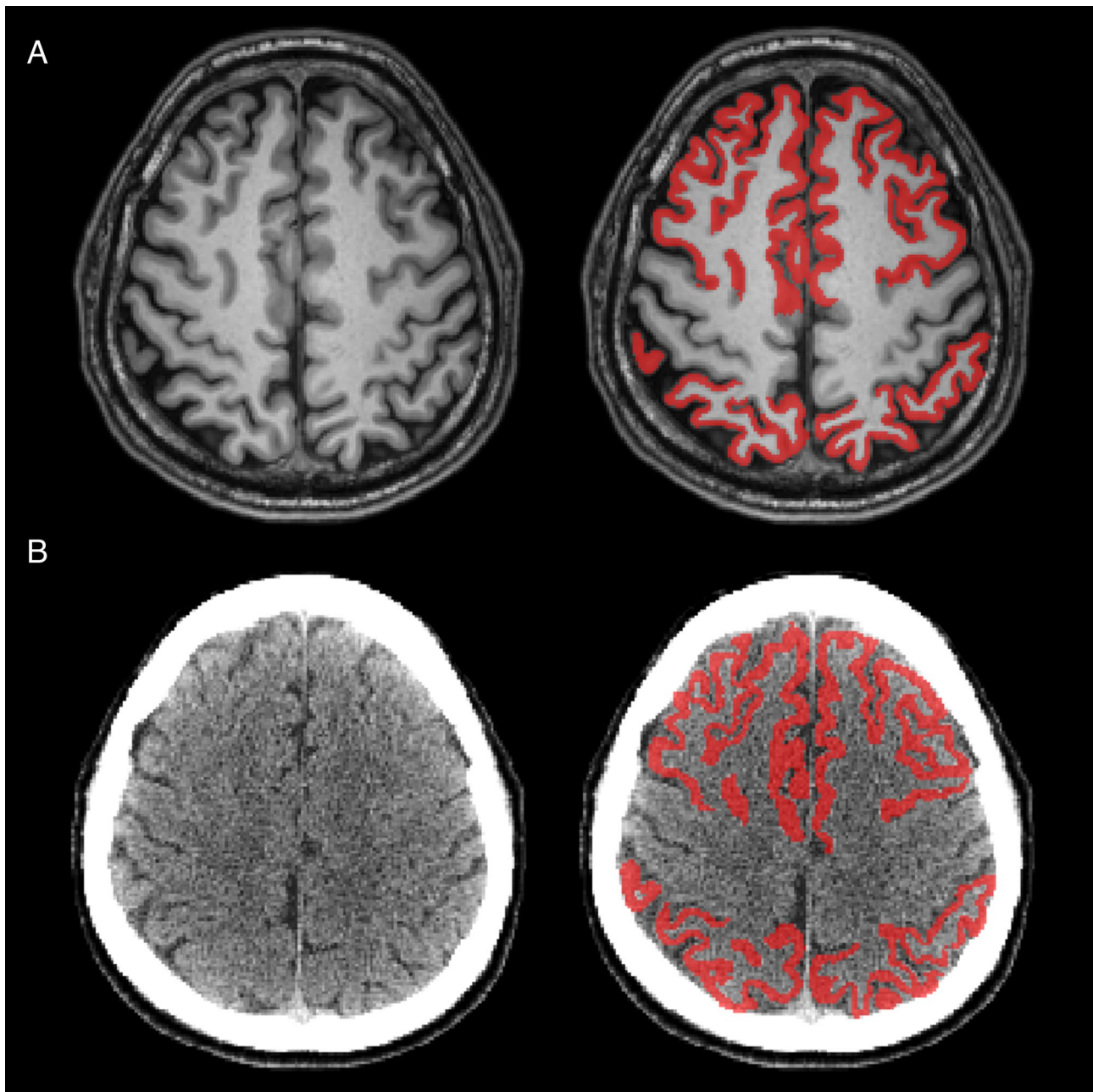


Fig. 2 Cortical mask overlaid on original MRI and CT images. **(A)** Representative images of MRI after segmentation using the FS pipeline. Left: original MRI scan; Right: same scan with the cortical mask overlaid, delineating the frontal, cingulate, lateral parietal, and lateral temporal regions. **(B)** Representative CT images following segmentation using the CT parcellation pipeline. Left: original CT scan; Right: same scan with the cortical mask overlaid, delineating the corresponding target regions. Abbreviation: FS, FreeSurfer

study, when using WC as the reference region, both CL_{FS} and CL_{CT} consistently showed the lowest variance among YC groups (Table 3). When comparing different pipelines using WC as the reference region, no significant difference was observed between CL_{FS-WC} and CL_{CT-WC} ($p=0.73$); however, both CL_{FS-WC} ($p=0.038$) and CL_{CT-WC} ($p=0.017$) exhibited significantly lower variance than $CL_{standard}$.

In addition, the effect size between the ADCI and YC groups was highest with WC in the FS (3.15, 95% CI: 2.62–3.66) and CT parcellation (3.17, 95% CI: 2.65–3.69) pipelines. When comparing the pipelines using WC as the reference region, the effect sizes of $CL_{standard}$ (3.10, 95% CI: 2.58–3.61), CL_{FS-WC} (3.15, 95% CI: 2.62–3.66) and CL_{CT-WC} (3.17, 95% CI: 2.65–3.69) did not differ significantly. Overall, WC was determined to be the optimal reference region in the CT parcellation pipeline, which

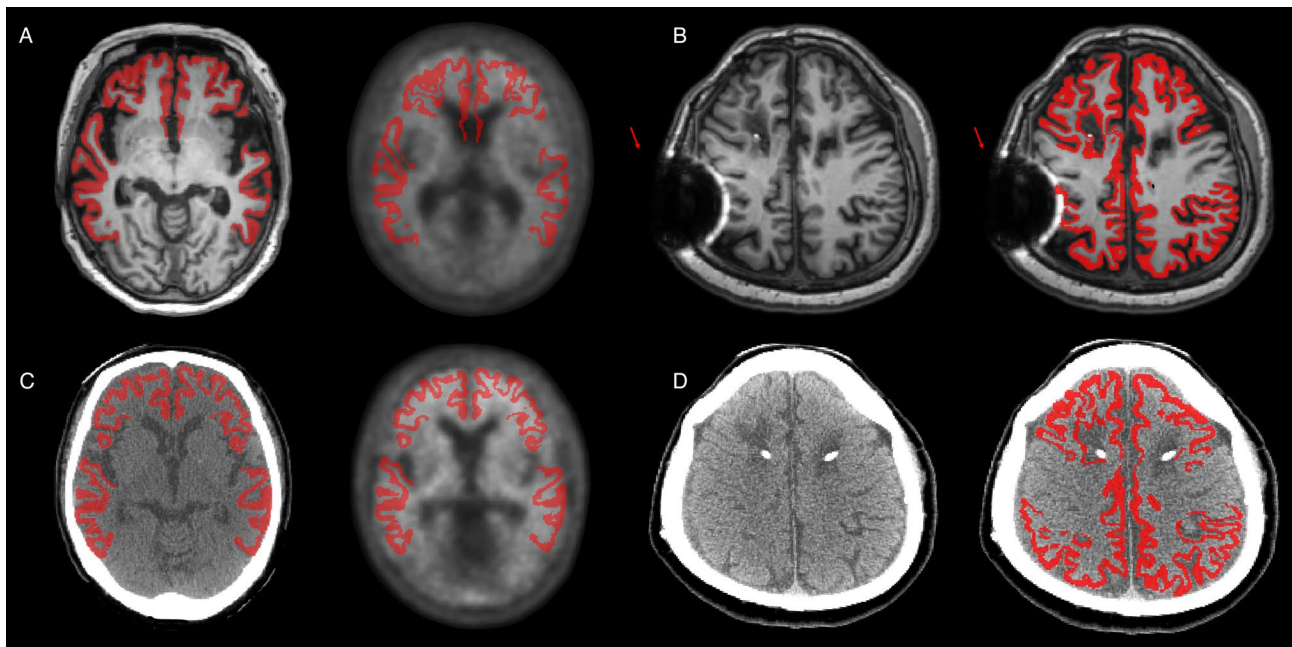


Fig. 3 Representative cases that were excluded from regression analysis. **(A)** The outlier case from the FS pipeline, showing suboptimal co-registration between MRI and PET/CT. Left: original MRI with the FS-derived cortical mask; Right: co-registered PET/CT image in which the mask does not align properly. **(B)** Case with a pronounced susceptibility artifact on MRI. Left: original MRI displaying the artifact; Right: same MRI with the FS-based cortical mask applied. **(C)** Corresponding CT images for the outlier in **(A)**. Left: original CT scan; Right: CT parcellation-derived cortical mask overlaid on the co-registered PET/CT image. **(D)** Corresponding CT image for the artifact case in **(B)**. Left: original CT scan; Right: CT parcellation-derived cortical mask overlaid on the PET/CT image. Both cases illustrated in **(A)** and **(B)** were excluded from the regression analysis. Abbreviation: FS, FreeSurfer

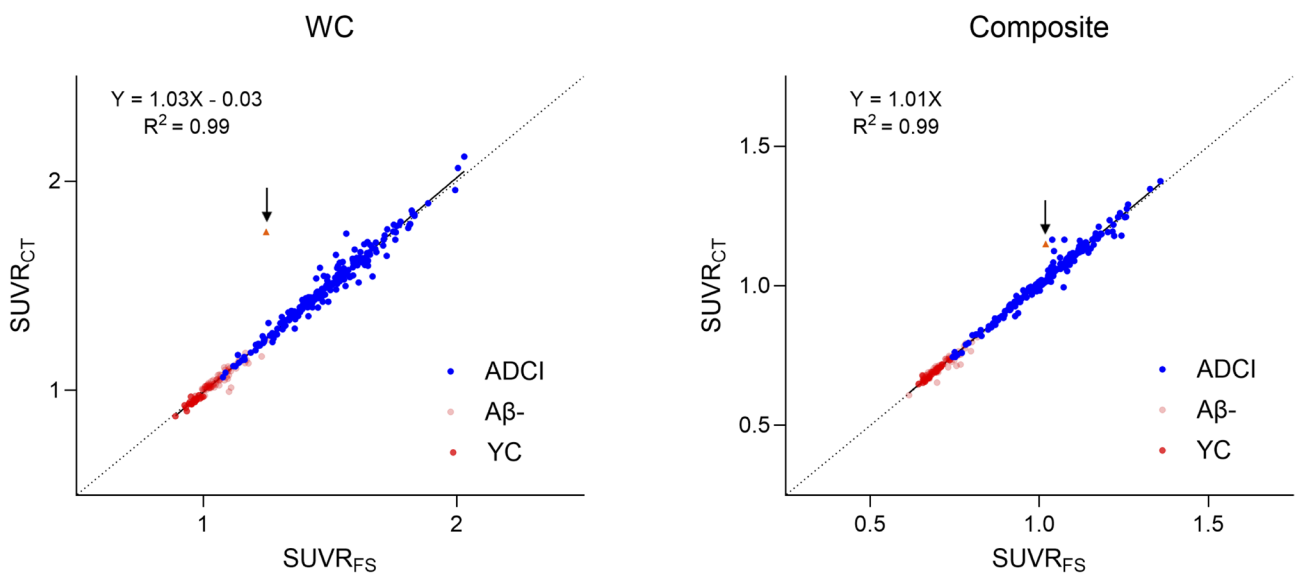


Fig. 4 Linear regression of $SUVR_{FS}$ against $SUVR_{CT}$ for the whole cerebellum (WC) and the composite reference region. The scatter plots display the relationship between $SUVR_{FS}$ and $SUVR_{CT}$ for each reference region. Dashed unity lines have been added to facilitate visual comparison between the axes. An outlier with inaccurate co-registration (as shown in Fig. 3A and C) is indicated by an orange triangle and black arrow in the scatter plot. Abbreviation: FS, FreeSurfer; $SUVR_{FS}$, FS pipeline-derived standardized uptake value ratio; $SUVR_{CT}$, CT parcellation pipeline-derived standardized uptake value ratio; WC, whole cerebellum; Composite, composite reference region

Table 2 CL conversion equations with FS and CT-based pipelines for each reference region

Reference region	FS pipeline	CT parcellation pipeline	Standard CL pipeline
WC	CL = 157.15 × SUVR – 151.87	CL = 153.05 × SUVR – 143.63	CL = 154.0 × SUVR – 155.1
Composite	CL = 244.20 × SUVR – 170.80	CL = 234.01 × SUVR – 168.12	N/A

Abbreviations: CL, Centiloid; FS, FreeSurfer; SUVR, standardized uptake value ratio; SUVR, standardized uptake value ratio; WC, whole cerebellum; Composite, composite reference region

Table 3 CL values across reference regions and groups

Reference region	CL _{FS}		CL _{CT}		CL _{standard}
	WC	Comp	WC	Comp	WC
ADCI					
Mean	85.27	78.53	85.15	72.28	83.81
SD	28.12	28.67	27.91	27.93	30.70
YC					
Mean	0.91	-3.78	0.99	-7.73	-6.66
SD	3.85^a	5.61	3.57^a	5.25	6.83
Effect size	3.15	3.01	3.17	3.01	3.10

Abbreviations: CL, Centiloid; FS, FreeSurfer; CL_{FS}, FS pipeline-derived CL value; CL_{CT}, CT parcellation pipeline-derived CL value; CL_{standard}, Centiloid pipeline-derived CL value; ADCI, patients with AD dementia and mild cognitive impairment; YC, young control; WC, whole cerebellum; Comp, composite reference region

^a Significant difference from CL_{standard}

demonstrated performance that was similar to either FS or the standard CL pipeline.

ROC analysis results

Both the FS and CT parcellation pipelines showed comparable performance in the ROC analysis. ROC curves were generated based on visual read outcomes. The AUC was 0.994 versus 0.995 for the FS and CT parcellation pipelines, respectively, with the optimal thresholds identified via Youden’s J index were 29.9 and 29.7. At these thresholds, FS and CT parcellation exhibited comparable performance: accuracy – 0.964 versus 0.967; sensitivity – 0.957 versus 0.957; and specificity – 0.980 versus 0.990, respectively. These results suggest that both pipelines provide equally robust classification performance for identifying Aβ-positive scans, with minimal differences in AUC, accuracy, sensitivity, and specificity.

Discussion

In this study, we calculated the CL conversion equation for our recently developed DL-based CT parcellation method. Using the derived equation, CT parcellation pipeline-derived CL scales showed high agreement with the CL scales generated by both FS and the CL pipeline. Notably, when using WC as the reference region, CL_{FS} and CL_{CT} demonstrated high concordance, with an R² of 0.99. Furthermore, the optimal thresholds for the two pipelines were similar (29.9 for CL_{FS} and 29.7 for CL_{CT}), falling within the previously reported range of 17 to 40 CL, with most studies citing values between 25 and 35 CL [9, 25–29]. This consistency reinforces the validity of our approach, and its alignment with established visual

cutoffs supports the clinical potential of the CT parcellation pipeline as a viable alternative to the FS pipeline.

Early brain PET quantification relied on template-based methods, aligning PET images to a standardized framework for ROI-based analysis. MRI-free approaches using CT or PET templates have been explored but struggle with accurately defining intricate cortical structures and are prone to distortions in cases with severe atrophy or hydrocephalus [8, 11]. Parcellation-based methods, like FS, improve accuracy by extracting precise ROIs from MRI. Our CT parcellation method retains these advantages while eliminating the need for MRI, enhancing accessibility and robustness [15]. First, the CT parcellation pipeline enables rapid and cost-effective CL quantification for a broader patient population. Unlike MRI, which requires long scan times and may be inaccessible due to cost or contraindications [7, 8], CT scans are acquired simultaneously during PET/CT scans, significantly reducing acquisition time and improving accessibility. This is particularly beneficial following the introduction of monoclonal antibody therapies, as regular Aβ imaging is essential for tracking Aβ accumulation. By enabling Aβ quantification without high-resolution MRI, this approach may facilitate more frequent treatment monitoring, ultimately improving AD management. Second, in PET/CT scans, CT and PET images are acquired sequentially within the same session using a shared bed and headrest, which helps maintain consistent positioning and facilitates spatial alignment. This typically results in improved co-registration, supporting more precise quantification. In contrast, MRI and PET are susceptible to co-registration failures due to factors such as noise, limited spatial resolution, and time gaps between acquisitions. For instance, we observed a case where accurate co-registration between MRI and PET failed and resulted in inaccurate quantification (Fig. 3). Although SPM12, used in the FS pipeline, has been reported to outperform other tools such as NiftyReg and Vinci in MRI-PET co-registration accuracy, the FS pipeline is still not entirely free from co-registration issues [30].

As previously mentioned, YC variance and the effect size between ADCI and YC are key variables used in cross-sectional studies for comparing reference regions and pipelines. Klunk et al. previously investigated these two variables to identify the optimal reference region for the standard CL pipeline [6]. In this study, the YC

variance of the standard CL pipeline was significantly higher than that of both the FreeSurfer and CT parcellation pipelines. Since YC variance reflects the noise inherent in a quantification method, the template-based method may be more susceptible to noise—possibly due to its inability to perform quantification along accurate cortical gyri—compared to the parcellation-based method. This increased variance might also result from the fact that the cortical ROI in the standard CL pipeline was derived from PiB PET data, which may not fully apply to FBB PET. Further studies are needed to determine the superiority between template-based and parcellation-based methods.

Biologically, the cerebellum is favorable because it is free of A β deposition and exhibits nondisplaceable activity similar to the target cortical area [31, 32]. Previous cross-sectional and longitudinal studies have further supported WC as the most reliable reference region for FBB [24, 33]. In parallel, recent studies on longitudinal A β quantification using PiB and ¹⁸F-Florbetapir have indicated that a composite reference region, which combines WC, brainstem, and eroded subcortical white matter, may offer improved longitudinal stability [20, 21, 34]. In this study, we compared WC and the composite reference region; although the difference was not statistically significant, WC consistently yielded better performance in terms of YC variance. This may reflect the fact that subcortical white matter is more susceptible to atrophy, vascular lesions, and signal spillover [21]. However, our data were acquired in a single-center dataset with uniform acquisition settings. In multi-center studies where scanner types and participant positioning vary, the composite reference region may offer better longitudinal consistency.

Regarding the use of CT in neurodegenerative disease screening, CT-based assessments have shown significant potential, with recent advancements in DL techniques further enhancing their capabilities. CT quantification has demonstrated performance comparable to MRI in measuring brain atrophy and white matter lesions [35]. Additionally, CT-based volumetric measures can differentiate patients with neurodegenerative diseases from healthy controls and are strongly associated with cognitive, biochemical, and neuroimaging markers [36]. Our CT parcellation-based CL approach further extends this potential by enabling quantitative amyloid PET analysis without the need for MRI, making amyloid burden assessment more accessible and cost-effective. Moreover, emerging technologies such as photon-counting CT offer high-resolution imaging with lower radiation exposure [37]. As CT resolution and soft tissue contrast continue to improve, its utility in screening not only for AD but also for other neurodegenerative diseases may increase further. Additionally, extracting more information from

the CT performed simultaneously with PET could enable a more comprehensive assessment in a single scan, providing significant benefits for patients.

This study has several limitations. First, we increased the tube current to 200 mAs for high-quality parcellation, resulting in an effective dose of 0.39 mSv from CT. However, this additional dose is minimal compared to the total PET/CT effective dose, which exceeds 4 mSv [19]. Several recent studies have investigated CT image denoising techniques aimed at enabling reliable anatomical analysis from low-dose CT scans [38, 39]. Applying such denoising methods prior to parcellation may allow our current pipeline to be extended to dose-reduced CT images. In addition, recent studies have demonstrated that PET tracer doses can be substantially reduced to as low as 12.5% of the original dose without compromising quantitative accuracy in CL scaling [40]. Together, these complementary strategies may help minimize overall radiation exposure in PET/CT protocols and improve the feasibility of the proposed method in diverse research and clinical settings. Second, as this is a single-center study, validation in different centers, with various PET tracers and scanner modes, is required to establish broader applicability. Incorporating a more diverse training dataset – encompassing scans with structural abnormalities and those from multiple scanners and tracers – could improve robustness and broaden clinical applicability. Leveraging publicly available imaging datasets that include PET/CT and MR scans could enable external validation and help assess generalizability across diverse populations and acquisition protocols. Furthermore, our pipeline is designed specifically for PET/CT data and is not currently applicable to PET/MRI. Given the increasing adoption of PET/MRI scanners, extending compatibility to this modality is an important direction for future development. Third, although CT and PET are acquired sequentially using the same headrest and positioning setup, our pipeline is not entirely free from co-registration error due to potential patient movement. In addition, while the CL standard pipeline is based on SPM8, our analysis employed SPM12. Despite strong agreement in Level 1 validation, segmentation differences between versions may introduce methodological variability. Fourth, we did not directly compare our CT-based pipeline to PET-only template methods (e.g., rPOP). Although our high concordance with MRI-guided FreeSurfer supports the accuracy of CT parcellation, a head-to-head evaluation against PET-only approaches would help determine which method enables more accurate quantification. Finally, longitudinal validation of the CL scale is necessary to establish its robustness in tracking A β changes over time. While the CT parcellation pipeline has demonstrated strong performance in cross-sectional studies, its stability and its comparability to other pipelines for

monitoring disease progression remain unclear. Further studies are required to assess whether CL_{CT} remains stable in $A\beta$ -negative individuals and whether it can detect subtle but clinically meaningful changes in $A\beta$ -positive individuals over time.

Abbreviations

$A\beta$	Amyloid-beta
AD	Alzheimer's disease
ADCI	Alzheimer's disease-related cognitive impairment
ADL	Activities of daily living
aMCI	Amnesic mild cognitive impairment
AUC	Area under the ROC curve
BAPL	Brain amyloid plaque load
CDR	Clinical dementia rating
CI	Confidence interval
CL	Centiloid
MRI	Magnetic resonance imaging
FBB	^{18}F -florbetaben
FS	FreeSurfer
K-MMSE	Korean version of the Mini-Mental State Examination
PiB	^{11}C -Pittsburgh Compound-B
RCTU	Regional cortical tracer uptake
ROC	Receiver operating characteristic
ROI	Region of interest
SUVr	Standardized uptake value ratios
WC	Whole cerebellum

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13195-025-01860-1>.

Supplementary Material 1

Acknowledgements

The authors are grateful to all participants who took part in this study.

Author contributions

Y. Yoon, S. Seo, and M. Yun contributed to the conceptualization and design of the study. Material preparation and data collection were performed by Y. Yoon, S. Seo, D. Kim, H. Han, and S. Kang. Data analysis was conducted by Y. Yoon, K. Choo, S. Lee, H. Lim, and H. Kang. The first draft of the manuscript was written by Y. Yoon and S. Seo. Revisions for content, including medical writing, were performed by Y. Yoon, S. Seo, M. So, D. Kim, D. Shin, and Y. Lee. M. Yun and T.J. Jeon equally supervised the study and contributed to the final revision of the manuscript. All authors read and approved the final manuscript.

Funding

This research was supported by a grant from the Korea Dementia Research Project through the Korea Dementia Research Center (KDRC), funded by the Ministry of Health & Welfare and the Ministry of Science and ICT, Republic of Korea (grant number RS-2022-KH127174); by the Technology Development Program (RS-2024-00467478 and RS-2024-00508913) funded by the Ministry of SMEs and Startups (MSS, Korea); and by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (RS-2025-00520546); and the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea (RS-2022-KH130344).

Data availability

No datasets were generated or analysed during the current study.

Declarations

Ethics approval

This study was performed in accordance with the principle of the institutional and/or national research committee and with the Declaration of Helsinki.

Approval was granted by the institutional review board of Severance Hospital (IRB no. 2024-3483-001).

Competing interests

The authors declare the following financial interests and personal relationships, which may be considered as potential competing interests: Author Mijin Yun has equity in Newcure M, Inc. No other author has reported a potential conflict of interest relevant to this article.

Author details

- ¹Yonsei University College of Medicine, 50-1 Yonsei-ro, Seodaemun-gu, Seoul 03722, Republic of Korea
- ²Department of Nuclear Medicine, Yonsei University College of Medicine, 50-1 Yonsei-ro, Seodaemun-gu, Seoul 03722, Republic of Korea
- ³Department of Computer Science, Yonsei University, 50 Yonsei-ro, Seodaemun-gu, Seoul 03722, Republic of Korea
- ⁴Department of Artificial Intelligence, Yonsei University, 50 Yonsei-ro, Seodaemun-gu, Seoul 03722, Republic of Korea
- ⁵ELTEC College of Engineering, Ewha Woman's University, 52 Ewhayeodae-gil, Seodaemun-gu, Seoul 03760, Republic of Korea
- ⁶Department of Nuclear Medicine, Hallym University Sacred Heart Hospital, 22 Gwanpyeong-ro, Dongan-gu, Anyang 14068, Republic of Korea
- ⁷Department of Neurology, Ilsan Paik Hospital, Inje University College of Medicine, 170 Juhwa-ro, Ilsanseo-gu, Goyang 10380, Republic of Korea
- ⁸Massachusetts College of Pharmacy & Health Sciences, Boston, USA
- ⁹Department of Nuclear Medicine, Gangnam Severance Hospital, Yonsei University College of Medicine, 211 Eonjuro, Gangnam-gu, Seoul 06273, Republic of Korea

Received: 14 March 2025 / Accepted: 1 September 2025

Published online: 30 September 2025

References

- Lammertsma AA. Forward to the past: the case for quantitative PET imaging. *J Nucl Med*. 2017;58(7):1019–24.
- Rosenberg PB, Wong DF, Edell SL, Ross JS, Joshi AD, Brasic JR, et al. Cognition and amyloid load in alzheimer disease imaged with Florbetapir F 18(AV-45) positron emission tomography. *Am J Geriatr Psychiatry*. 2013;21(3):272–8.
- Kim SE, Lee B, Park S, Cho SH, Kim SJ, Kim Y, et al. Clinical significance of focal α -amyloid deposition measured by (18F)-flutemetamol PET. *Alzheimers Res Ther*. 2020;12(1):6.
- Zhang Y, Chen H, Li R, Sterling K, Song W. Amyloid beta-based therapy for alzheimer's disease: challenges, successes and future. *Signal Transduct Target Ther*. 2023;8(1):248.
- Park KW. Anti-amyloid antibody therapies for alzheimer's disease. *Nucl Med Mol Imaging*. 2024;58(4):227–36.
- Klunk WE, Koeppe RA, Price JC, Benzinger TL, Devous MD, Sr., Jagust WJ, et al. The centiloid project: standardizing quantitative amyloid plaque estimation by PET. *Alzheimers Dement*. 2015;11(1):1–e151.
- Delhomme C, Njeim M, Varlet E, Pechmajou L, Benamer N, Cassan P, et al. Automated external defibrillator use in out-of-hospital cardiac arrest: current limitations and solutions. *Arch Cardiovasc Dis*. 2019;112(3):217–22.
- Landau SM, Ward TJ, Murphy A, Iaccarino L, Harrison TM, La Joie R, et al. Quantification of amyloid beta and tau PET without a structural MRI. *Alzheimers Dement*. 2023;19(2):444–55.
- Iaccarino L, La Joie R, Koeppe R, Siegel BA, Hillner BE, Gatsonis C, et al. rPOP: robust PET-only processing of community acquired heterogeneous amyloid-PET data. *NeuroImage*. 2022;246:118775.
- Bourgeat P, Villemagne VL, Dore V, Brown B, Macaulay SL, Martins R, et al. Comparison of MR-less PiB SUVr quantification methods. *Neurobiol Aging*. 2015;36(Suppl 1):S159–66.
- Kim SJ, Ham H, Park YH, Choe YS, Kim YJ, Jang H, et al. Development and clinical validation of CT-based regional modified centiloid method for amyloid PET. *Alzheimers Res Ther*. 2022;14(1):157.
- Srikrishna M, Pereira JB, Heckemann RA, Volpe G, van Westen D, Zettergren A, et al. Deep learning from MRI-derived labels enables automatic brain tissue classification on human brain CT. *NeuroImage*. 2021;244:118606.

13. Huisman S, Maspero M, Philippens M, Verhoeff J, David S. Validation of synthseg segmentation performance on CT using paired MRI from radiotherapy patients. *NeuroImage*. 2024;303:120922.
14. Kim D, Choo K, Lee S, Kang S, Yun M, Yang J. Automated quantification of brain PET in PET/CT using deep learning-based CT-to-MR translation: a feasibility study. *Eur J Nucl Med Mol Imaging*. 2025;52(8):2959–67.
15. Choo K, Joo J, Lee S, Kim D, Lim H, Kim D et al. Deep learning-based precontrast CT parcellation for MRI-free brain amyloid PET quantification. *Clin Nucl Med*. 2025;50(5):e262–e270.
16. McKhann GM, Knopman DS, Chertkow H, Hyman BT, Jack CR Jr., Kawas CH, et al. The diagnosis of dementia due to alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's association workgroups on diagnostic guidelines for alzheimer's disease. *Alzheimers Dement*. 2011;7(3):263–9.
17. Petersen RC, Smith GE, Waring SC, Ivnik RJ, Tangalos EG, Kokmen E. Mild cognitive impairment: clinical characterization and outcome. *Arch Neurol*. 1999;56(3):303–8.
18. Seibyl J, Catafau AM, Barthel H, Ishii K, Rowe CC, Leverenz JB, et al. Impact of training method on the robustness of the visual assessment of 18F-Florbetaben PET scans: results from a phase-3 study. *J Nucl Med*. 2016;57(6):900–6.
19. Sabri O, Seibyl J, Rowe C, Barthel H. Beta-amyloid imaging with florbetaben. *Clin Transl Imaging*. 2015;3(1):13–26.
20. Royse SK, Minhas DS, Lopresti BJ, Murphy A, Ward T, Koeppe RA, et al. Validation of amyloid PET positivity thresholds in centiloids: a multisite PET study approach. *Alzheimers Res Ther*. 2021;13(1):99.
21. Landau SM, Fero A, Baker SL, Koeppe R, Mintun M, Chen K, et al. Measurement of longitudinal beta-amyloid change with 18F-florbetapir PET and standardized uptake value ratios. *J Nucl Med*. 2015;56(4):567–74.
22. Sabri O, Sabbagh MN, Seibyl J, Barthel H, Akatsu H, Ouchi Y, et al. Florbetaben PET imaging to detect amyloid beta plaques in alzheimer's disease: phase 3 study. *Alzheimers Dement*. 2015;11(8):964–74.
23. Rowe CC, Dore V, Jones G, Baxendale D, Mulligan RS, Bullich S, et al. (18) F-Florbetaben PET beta-amyloid binding expressed in centiloids. *Eur J Nucl Med Mol Imaging*. 2017;44(12):2053–9.
24. Cho SH, Choe YS, Park S, Kim YJ, Kim HJ, Jang H, et al. Appropriate reference region selection of (18)F-florbetaben and (18)F-flutemetamol beta-amyloid PET expressed in centiloid. *Sci Rep*. 2020;10(1):14950.
25. Collij LE, Salvado G, Shekari M, Lopes Alves I, Reimand J, Wink AM, et al. Visual assessment of [(18)F]flutemetamol PET images can detect early amyloid pathology and grade its extent. *Eur J Nucl Med Mol Imaging*. 2021;48(7):2169–82.
26. Collij LE, Salvado G, de Wilde A, Altomare D, Shekari M, Gispert JD, et al. Quantification of [(18)F]florbetaben amyloid-PET imaging in a mixed memory clinic population: the ABIDE project. *Alzheimers Dement*. 2023;19(6):2397–407.
27. La Joie R, Ayakta N, Seeley WW, Borys E, Boxer AL, DeCarli C, et al. Multisite study of the relationships between antemortem [(11)C]PIB-PET centiloid values and postmortem measures of alzheimer's disease neuropathology. *Alzheimers Dement*. 2019;15(2):205–16.
28. Collij LE, Bischof GN, Altomare D, Bader I, Battle M, Vallez Garcia D, et al. Quantification supports amyloid PET visual assessment of challenging cases: results from the AMYPAD diagnostic and patient management study. *J Nucl Med*. 2025;66(1):110–6.
29. Amadoru S, Dore V, McLean CA, Hinton F, Shepherd CE, Halliday GM, et al. Comparison of amyloid PET measured in centiloid units with neuropathological findings in alzheimer's disease. *Alzheimers Res Ther*. 2020;12(1):22.
30. Markiewicz PJ, Matthews JC, Ashburner J, Cash DM, Thomas DL, De Vita E, et al. Uncertainty analysis of MR-PET image registration for precision neuro-PET imaging. *NeuroImage*. 2021;232:117821.
31. Schmidt ME, Chiao P, Klein G, Matthews D, Thurfjell L, Cole PE, et al. The influence of biological and technical factors on quantitative analysis of amyloid PET: points to consider and recommendations for controlling variability in longitudinal data. *Alzheimers Dement*. 2015;11(9):1050–68.
32. Catafau AM, Bullich S, Seibyl JP, Barthel H, Ghetti B, Leverenz J, et al. Cerebellar amyloid-beta plaques: how frequent are they, and do they influence 18F-Florbetaben SUV ratios?? *J Nucl Med*. 2016;57(11):1740–5.
33. Bullich S, Villemagne VL, Catafau AM, Jovalekic A, Koglin N, Rowe CC, et al. Optimal reference region to measure longitudinal amyloid-beta change with (18)F-Florbetaben PET. *J Nucl Med*. 2017;58(8):1300–6.
34. Schwarz CG, Senjem ML, Gunter JL, Tosakulwong N, Weigand SD, Kemp BJ, et al. Optimizing PiB-PET SUVR change-over-time measurement by a large-scale analysis of longitudinal reliability, plausibility, separability, and correlation with MMSE. *NeuroImage*. 2017;144:113–27.
35. Kaipainen AL, Pitkanen J, Haapalinna F, Jaaskelainen O, Jokinen H, Melkas S, et al. A novel CT-based automated analysis method provides comparable results with MRI in measuring brain atrophy and white matter lesions. *Neuroradiology*. 2021;63(12):2035–46.
36. Srikrishna M, Ashton NJ, Moscoso A, Pereira JB, Heckemann RA, van Westen D, et al. CT-based volumetric measures obtained through deep learning: association with biomarkers of neurodegeneration. *Alzheimers Dement*. 2024;20(1):629–40.
37. Esquivel A, Ferrero A, Mileto A, Baffour F, Horst K, Rajiah PS, et al. Photon-counting detector CT: key points radiologists should know. *Korean J Radiol*. 2022;23(9):854–65.
38. Sadia RT, Chen J, Zhang J. CT image denoising methods for image quality improvement and radiation dose reduction. *J Appl Clin Med Phys*. 2024;25(2):e14270.
39. Yang Q, Yan P, Zhang Y, Yu H, Shi Y, Mou X, et al. Low-dose CT image denoising using a generative adversarial network with Wasserstein distance and perceptual loss. *IEEE Trans Med Imaging*. 2018;37(6):1348–57.
40. Young P, Heeman F, Axelsson J, Collij LE, Hitzel A, Sanaat A, et al. Impact of simulated reduced injected dose on the assessment of amyloid PET scans. *Eur J Nucl Med Mol Imaging*. 2024;51(3):734–48.

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.