Improved accuracy of amyloid PET quantification with adaptive template-based anatomical standardization

Yuma Tsubaki[1], Takayoshi Kitamura[2], Natsumi Shimokawa[1], Go Akamatsu[3], Masayuki Sasaki[1], Japanese Alzheimer's Disease Neuroimaging Initiative*

[1]Department of Health Sciences, Graduate School of Medical Sciences, Kyushu University, Fukuoka, Japan

[2]Department of Health Sciences, School of Medicine, Kyushu University, Fukuoka, Japan

[3]National Institute of Radiological Sciences, National Institutes for Quantum and Radiological Science and Technology (NIRS-QST), Chiba, Japan

Corresponding author

Masayuki Sasaki, MD, PhD

Department of Medical Quantum Science, Faculty of Medical Sciences, Kyushu University

3-1-1 Maidashi, Higashi-ku, Fukuoka 812-8582, Japan

TEL: +81-92-642-6746, FAX: +81-92-642-6723

E-mail: msasaki@hs.med.kyushu-u.ac.jp

First author

Yuma Tsubaki

Department of Medical Quantum Science, Graduate School of Medical Sciences, Kyushu University

3-1-1 Maidashi, Higashi-ku, Fukuoka 812-8582, Japan

TEL: +81-92-642-6746, FAX: +81-92-642-6723

E-mail: tsubaki.yuma.650@s.kyushu-u.ac.jp

Abstract

Amyloid positron emission tomography (PET) noninvasively visualizes amyloid-β (Aβ) accumulation in the brain. Visual binary reading is the standard method for interpreting amyloid PET, while objective quantitative evaluation is required in research and clinical trials. Anatomical standardization is important for quantitative analysis, and various standard templates are used for this purpose. To address the large differences in the radioactivity distribution between amyloid-positive and amyloid-negative participants, an adaptive template method has been proposed for the anatomical standardization of amyloid PET. In this study, we investigated the difference between the adaptive template method and the single template methods (use of a positive or a negative template) in amyloid PET quantitative evaluation, focusing on the accuracy in diagnosing Alzheimer's disease (AD). **Methods:** A total of 166 participants (58 normal controls (NCs), 62 participants with mild cognitive impairment (MCI), and 46 patients with AD) who underwent [^{11}C] Pittsburgh Compound B PET (^{11}C-PiB) through the Japanese Alzheimer's Disease Neuroimaging Initiative study were examined. For the anatomical standardization of ^{11}C-PiB PET images, we applied three methods: a positive template-based method, a negative template-based method, and adaptive template-based method. The positive template was created by averaged four patients with AD and seven patients with MCI PET images. Conversely, the negative template was created by averaged eight participants of NC PET images. In the adaptive template-based method, either of the

3

templates was used on the basis of the similarity (normalized cross-correlation (NCC)) between the individual standardized image and the corresponding template. Empirical PiB-prone region-of-interest was used to evaluate specific regions where Aβ accumulates. The reference region was the cerebellar cortex, whereas the evaluated regions were the posterior cingulate gyrus and precuneus as well as the frontal, lateral temporal, lateral parietal, and occipital lobes. The mean cortical standardized uptake value ratio (mcSUVR) was calculated for quantitative evaluation. **Results**: The NCCs of single template-based methods (the positive template or negative template) showed a significant difference between NC, MCI and AD ($P < 0.05$), whereas the NCC of the adaptive template-based method did not ($P > 0.05$). The mcSUVR exhibited significant differences between NC, MCI and AD in all methods ($P < 0.05$). The area under curve by receiver operating characteristic analysis between the positive group (MCI and AD) and NC was not significantly different in mcSUVR among all templates. With regard to diagnostic accuracy based on mcSUVR, the sensitivity of the negative and adaptive template-based methods was superior to that of the positive template-based method ($P < 0.05$); however, there was no significant difference in specificity between them. **Conclusions**: In the diagnostic accuracy for AD by amyloid PET quantitative evaluation, the adaptive template-based anatomical standardization method outperformed the single template-based methods.

Key words: Alzheimer's disease; amyloid PET; anatomical standardization; adaptive template method

INTRODUCTION

Dementia is a syndrome due to diseases of the brain in which shows disturbance of multiple

higher cortical functions (*1*). The most common type of dementia is Alzheimer's disease (AD). It

accounts for more than 50% of cases of primary disease causing dementia in Japan (*2*). The number of

affected patients is expected to reach approximately 5.0 million by 2025 (*2,3*).

The cause of AD is thought to be the neuronal degeneration induced by the accumulation of

amyloid-β (Aβ) plaques and phosphorylated tau protein. Such an accumulation is considered to begin

before the onset of cognitive impairment (*4*). Amyloid positron emission tomography (PET)

noninvasively visualizes Aβ plaques in the brain. The standard method of interpreting amyloid PET

scans is visual binary reading, and objective quantitative evaluation is required in research and clinical

trials. Anatomical standardization (i.e., spatial normalization) is essential for the quantitative evaluation

of amyloid PET; for this process, a standard brain template is required. The standard templates are

defined in the standard space (e.g., Montreal Neurological Institute standard space), and various types of

templates, such as a magnetic resonance (MR) T1 template and a fluorodeoxyglucose PET template

were used for anatomical standardization (*5*). While MR-based methods are standard for anatomical

standardization of brain PET, PET-based methods have also been employed in numerous studies for

practical reasons (*6,7*). Amyloid PET demonstrates the activity distribution patterns that differ between

amyloid-positive and amyloid-negative images. To account for the difference in the activity distributions, Akamatsu et al. developed an adaptive template-based method, which involves the use of multiple templates, both positive and negative templates (*8*). In the adaptive template-based method, the template that is most similar to the subject image is selected and used for anatomical standardization. In some studies, the adaptive template-based method has been employed for amyloid PET anatomical standardization (*9,10*). However, the difference between the adaptive template-based method and the single template-based method in amyloid PET quantitative evaluation has not been well elucidated.

In this study, the influence of the different anatomical standardization methods (the adaptive template-based method and the single template-based methods) on amyloid PET quantitative evaluation is investigated, focusing on the diagnostic accuracy for AD.

MATERIALS AND METHODS

Participants

We retrospectively analyzed the data from 166 participants who underwent [^{11}C] Pittsburgh Compound-B (^{11}C-PiB) PET examination through the Japanese Alzheimer's Disease Neuroimaging Initiative (J-ADNI) study (*11*). The J-ADNI study is a multi-institutional research project on Alzheimer's disease led by the Ministry of Health, Labour and Welfare and the New Energy and Industrial

Technology Development Organization in Japan. (*12*). The J-ADNI study was approved by the ethics

committee of each institution for data acquisition and an additional written informed consent for this

retrospective study was obtained. The data utilized in this study were provided by the National

Bioscience Database Center and were retrospectively analyzed. In addition, the requirement for written

informed consent was waived. This study was approved by the ethics committee of Kyushu University,

Fukuoka, Japan (approval no. 30-174).

All participants were native Japanese speakers, and their mean age was 70.5 ± 6.3 years

(range: 60–84 years). The participants consisted of 58 normal controls (NCs), 62 participants with mild

cognitive impairment (MCI), and 46 patients with AD as shown in Table 1. The diagnoses of MCI and

probable AD were based on the clinical criteria presented by National Institute of Neurological and

Communicative Disorders and the Alzheimer's Disease and Related Disorders Association

(NINCDS-ADRDA). The Mini-Mental State Examination–Japanese (MMSE-J), the Clinical Dementia

Rating Scale–Japanese (CDR-J), and the Wechsler Memory Scale-R, Logical Memory II, corrected for

education (WMS-R), were used to classify the early stages of dementia. The NC participants scored

24–30 on the MMSE-J, 0 on the CDR-J, and above the cutoff values on the WMS-R. The MCI

participants scored 24–30 on the MMSE-J, 0.5 on the CDR-J, and below the cutoff values on the

WMS-R. The AD patients scored 20–26 on the MMSE-J, 0.5 or 1 on the CDR-J, and below the cutoff

7

values on the WMS-R.

Imaging protocol of ^{11}C-PiB PET

^{11}C-PiB PET was performed using a total of 10 different PET camera models by 3 vendors as presented in Table 2 (*13*). PET images were reconstructed with data from 50 to 70 min after ^{11}C-PiB injection (555 ± 185 MBq). For attenuation correction, the segmented attenuation correction method by 6 min transmission scan or CT-based method was used depending on scanner types including stand-alone PET scanners and hybrid PET/CT scanners. Of the ^{11}C-PiB PET images, 88 were classified as visually positive, 68 as visually negative, and 10 as visually equivocal by three expert nuclear medicine physicians (*14*). All physicians have specialized in neuroimaging more than 15 years. The results of the visual interpretation were based on the official judgment by J-ADNI PET Core (*14*). In this study, the equivocal images were analyzed with the positive images (number of images: 98).

Workflow of the quantitative evaluation method

The positive and negative templates were created in a previous study (*8*). Eleven typical positive images (4 AD and 7 MCI participants with high ^{11}C-PiB accumulation) were averaged to generate positive template, and eight typical negative images (8 NC participants) were averaged to generate the negative template (*8*).

Figure 1 presents the workflow of the quantitative evaluation method (*8*). These processes

were performed using the PMOD software ver. 3.7 (PMOD Technologies LLC., Zürich, Switzerland).

First, the PET images were anatomically standardized to either the positive or negative template. In the

adaptive template-based method, the PET images were standardized to both templates, and the template

most similar to the subject image, according to the normalized cross-correlation (NCC), was selected.

Second, the transformation vector of the anatomical standardization was recorded. Third, regions of

interest (ROIs) were inversely transformed to an individual PET images using a transformation vector.

We calculated standardized uptake value ratio (SUVR) using the Empirical PiB-prone ROI (EPP-ROI)

that was generated to evaluate regions where Aβ specifically accumulates (*8*).

The NCC was calculated to evaluate similarities between anatomically standardized images

and the respective templates (*8*). In addition, the SUVR in five brain regions (posterior cingulate gyrus

and precuneus, frontal lobe, lateral temporal lobe, lateral parietal lobe, and occipital lobe) were

calculated; the reference region was the cerebellar cortex. The average SUVRs of the five regions were

referred to as mean cortical SUVR (mcSUVR).

Statistical analysis

JMP Pro 13 (SAS Institute Inc., Cary, NC, USA) was used for statistical analysis. The

Steel–Dwass test was employed to analyze the significance of the differences between NCCs and

mcSUVR of the three groups (NC, MCI and AD). The McNemar test was used for analyzing the

significance for diagnostic ability. The significance level was set to $P < 0.05$. The cutoff value for differential diagnosis was obtained according to maximum the Youden index (sensitivity + specificity – 1).

RESULTS

Concordance rate between visual evaluation and used template

Table 3 presents the concordance rate between the visual evaluation and the template that was used. When the adaptive template was used, the concordance of the adopted template with visual evaluation was 89.2%, and the association coefficient was 0.803.

NCC in relation to different templates

The results of the NCCs are presented in Figure 2. When the negative template was used, the mean NCCs of NC, MCI and AD were 0.754 ± 0.122, 0.654 ± 0.143, and 0.580 ± 0.106, respectively. The NCCs of the three groups significantly differed ($P < 0.05$). Thus, among the three groups, the NC who were visually negative had the highest NCC. When the positive template was used, the mean NCCs were 0.548 ± 0.130 for NC, 0.701 ± 0.142 for MCI, and 0.777 ± 0.098 for AD. The results significantly differed among the three groups ($P < 0.05$): the NCCs were higher for the positive participants (MCI and AD) compared with NC. When the adaptive template-based method was employed, the mean NCCs

were 0.778 ± 0.102 for NC, 0.791 ± 0.072 for MCI, and 0.803 ± 0.050 for AD. All three groups exhibited high NCCs, which did not differ significantly.

The mcSUVR in the different anatomical standardization methods

The mcSUVR obtained using the different anatomical standardization methods are presented in Figure 3. When the positive template was used, the mcSUVR was 1.48 ± 0.33 for NC, 1.86 ± 0.46 for MCI and 2.12 ± 0.45 for AD. When the negative template was used, it was 1.35 ± 0.26 for NC, 1.68 ± 0.42 for MCI and 1.93 ± 0.44 for AD. On the other hand, the mcSUVR of NC, MCI and AD was 1.37 ± 0.33, 1.80 ± 0.50 and 2.10 ± 0.47, respectively, when the adaptive template was used. The mean mcSUVR difference between the groups was greatest in adaptive template. The mcSUVR differed significantly among the three groups (NC, MCI, and AD) in all methods ($P < 0.05$), although the difference among the groups was the largest when the adaptive template-based method was used.

Diagnostic ability in different anatomical standardization methods

Figure 4 and Table 4 show the results of receiver operating characteristic (ROC) analysis for differentiating positive (MCI and AD) and negative (NC). The area under the curves (AUCs) were not significantly different among the anatomical standardization methods; however, that of the adaptive template-based method was slightly larger than those of the single template-based methods. The diagnostic ability of each method is presented in Table 4. The adaptive and negative template-based

11

methods exhibited significantly higher sensitivity than did the positive template-based method ($P < 0.05$). Neither the specificity nor accuracy of all the methods differed significantly; however, the accuracy of the adaptive template-based method was the highest.

DISCUSSION

In this study, we examined the influence of the different anatomical standardization methods on the ^{11}C-PiB PET quantitative evaluation, focusing on the diagnostic accuracy for AD. The adaptive template method was shown to be useful in improving the diagnostic performance of AD, as the mcSUVR difference between the groups was greater than that of the single template method.

The adaptive template method exhibited high agreement (89.2%) between visual classification and the used template. The mcSUVR of NC, MCI, and AD differed significantly in all the standardization methods. For the differential diagnosis, the adaptive template-based method had the highest accuracy among the methods.

The adaptive template-based method exhibited higher NCCs in all the groups than did the single template-based methods. Lundqvist et al. and Bourgeat et al. examined ^{18}F-flutemetamol PET and ^{11}C-PiB PET, respectively (*9,10*). They reported that the template most similar to visual evaluation adopted by using cross-correlation coefficient and normalized mutual information enhanced the

quantitative accuracy (*9,10*). Therefore, the adaptive template method is considered to successfully adopt

a similar template to individual PET image.

The mcSUVR of the three groups differed significantly ($P < 0.05$); the mcSUVR of the

adaptive template-based method was the largest compared with those of the other methods (Figure 3).

For sensitivity and accuracy based on the mcSUVR, the adaptive template method outperformed the

single template methods. This is probably because the cutoff value was calculated to maximize accuracy.

The negative and adaptive templates had the same cutoff values, whereas the positive template had a

higher cutoff value (Table 4). This value (cutoff value = 1.8) is high compared to previous study (cutoff

value = 1.5) (*15*) and may include a large amount of false negative cases (low sensitivity). The

specificities of the templates did not differ significantly, although the specificity of the positive template

was higher. The ROC curves were not significantly different among the three methods, although the

AUC of the adaptive template-based method was superior to those of the single template-based methods.

Edison et al. reported that the use of a conventional MR-based template and a single PET

template yielded comparable results in the ^{11}C-PiB PET quantitative analysis (*16*). Moreover, several

studies have demonstrated that the use of multiple PET templates for anatomical standardization

improved quantitative accuracy. Bourgeat et al. reported no significant difference between the SUVRs

using the adaptive PET atlas approach and the MR-based SUVR, although the SUVR obtained using the

13

single mean PET atlas approach was significantly different from the MR-based SUVR (*10*). Kang et al.

developed a deep learning-based approach to generate multiple adaptive ^{11}C-PiB PET templates (*17*).

This approach also significantly enhanced the quantitative accuracy of PET-based anatomical

standardization (*17*). Our results suggested that the adaptive template-based method can provide

sufficient accuracy for amyloid PET anatomical standardization, although only two templates (positive

and negative) were used.

This study had some limitations. First, the number of participants was small. Thus, further

examination of a larger number of participants is needed to yield more robust results. Second, two

templates were examined in this study. In the adaptive template-based method, increasing the number of

templates with various types of accumulation has the potential to improve the accuracy of anatomical

standardization. Third, the PET data were acquired more than 10 years ago. Because recent PET

scanners can provide higher resolution images, the difference between positive and negative images

might be clear. Therefore, we expect the adaptive template-based method is effective for current PET

images as well.

CONCLUSION

We have examined the influence of the different anatomical standardization methods on the

amyloid PET quantitative evaluation, focusing on the diagnostic accuracy for AD. The adaptive

template-based method slightly improved the diagnostic accuracy in comparison with the single

template-based method.

DISCLOSURE

No potential conflict of interest relevant to this article was reported.

REFERENCES

1. World Health Organization. International Statistical Classification of Diseases and Related Health Problems, 10th Revision. 1993.

2. Akatsu H, Takahashi M, Matsukawa N, et al. Subtype analysis of neuropathologically diagnosed patients in a Japanese geriatric hospital. J Neurol Sci. 2002;196(1-2):63-9.

3. Ninomiya T. Research on future projection of the population of the elderly with dementia in Japan. Ministry of Health, labour, and welfare. 2015. https://mhlw-grants.niph.go.jp/niph/search/NIDD00.do?resrchNum=201405037A [accessed 18 Nov 2020] (in Japanese)

4. Jack CR Jr, Knopman DS, Jagust WJ, et al. Hypothetical model of dynamic biomarkers of the Alzheimer's pathological cascade. Lancet Neurol. 2010;9(1):119-28.

5. Evans AC, Janke AL, Collins DL, et al. Brain templates and atlases. Neuroimage. 2012;62(2):911-22.

6. Gispert JD, Pascau J, Reig S, et al. Influence of the normalization template on the outcome of statistical parametric mapping of PET scans. Neuroimage. 2003;19(3):601-12.

7. Della Rosa PA, Cerami C, Gallivanone F, et al. A standardized [18F]-FDG-PET template for spatial normalization in statistical parametric mapping of dementia. Neuroinformatics. 2014;12(4):575-93.

8. Akamatsu G, Ikari Y, Ohnishi A, et al. Automated PET-only quantification of amyloid deposition with adaptive template and empirically pre-defined ROI. Phys Med Biol. 2016;61(15):5768-80.

9. Lundqvist R, Lilja J, Thomas BA, et al. Implementation and validation of an adaptive template registration method for 18F-flutemetamol imaging data. J Nucl Med. 2013;54(8):1472-8.

10. Bourgeat P, Villemagne VL, Dore V, et al. Comparison of MR-less PiB SUVR quantification methods. Neurobiol Aging. 2015;36 Suppl 1:S159-66.

11. Shimokawa N, Akamatsu G, Kadosaki M, et al. Feasibility study of a PET-only amyloid quantification method: a comparison with visual interpretation. Ann Nucl Med. 2020;34(9):629-635.

12. Iwatsubo T. Japanese Alzheimer's Disease Neuroimaging Initiative: present status and future. Alzheimers Dement. 2010;6(3):297-9.

13. Ikari Y, Nishio T, Makishi Y, et al. Head motion evaluation and correction for PET scans with 18F-FDG in the Japanese Alzheimer's disease neuroimaging initiative (J-ADNI) multi-center study. Ann Nucl Med. 2012;26(7):535-44.

14. Yamane T, Ishii K, Sakata M, et al. Inter-rater variability of visual interpretation and comparison with quantitative evaluation of 11C-PiB PET amyloid images of the Japanese Alzheimer's Disease Neuroimaging Initiative (J-ADNI) multicenter study. Eur J Nucl Med Mol Imaging. 2017;44(5):850-857.

15. Jagust WJ, Bandy D, Chen K, et al. The Alzheimer's Disease Neuroimaging Initiative positron emission tomography core. Alzheimers Dement. 2010;6(3):221-229.

16. Edison P, Carter SF, Rinne JO, et al. Comparison of MRI based and PET template based approaches in the quantitative analysis of amyloid imaging with PIB-PET. Neuroimage. 2013;70:423-33.

17. Kang SK, Seo S, Shin SA, et al. Adaptive template generation for amyloid PET using a deep learning approach. Hum Brain Mapp. 2018;39(9):3769-3778.

Table

Table 1. Characteristics of the participants who underwent ^{11}C-PiB PET

Characteristic	Clinical diagnosis		
	NC	MCI	AD
n (male/female)	58 (30/28)	62 (30/32)	46 (21/25)
Age, mean ± SD (range)	66.4 ± 4.5 (60–80)	71.4 ± 5.5 (60–82)	74.4 ± 6.3 (62–84)
NINCDS-ADRDA	—	—	Probable AD
MMSE-J, mean ± SD (range)	29.3 ± 1.1 (24–30)	26.7 ± 1.8 (24–30)	22.2 ± 1.8 (20–26)
CDR-J	0	0.5	0.5 or 1.0
WMS-R	Above the cutoff	Below the cutoff	Below the cutoff
Visually positive, n	14	41	43
Visually negative, n	44	21	3

n = number of participants; NC = normal control; MCI = mild cognitive impairment; AD = Alzheimer's

disease; NINCDS-ADRDA = National Institute of Neurological and Communicative Disorders and the

Alzheimer's Disease and Related Disorders Association; MMSE-J = Mini-Mental State

Examination–Japanese; CDR-J = Clinical Dementia Rating Scale–Japanese; WMS-R, Wechsler Memory

Scale–R, Logical Memory II, corrected for education.

Table 2. PET scanners and reconstruction parameters used for ^{11}C-PiB PET in the J-ADNI study

PET scanner		Reconstruction parameters		
Vendor	**Model**	**Algorithm**	**Iteration**	**Subset**
GE	Advance	Iterative (FORE+OSEM)	6	16
GE	Discovery ST Elite	Iterative (VUE Point plus)	2	40
Shimadzu	Eminence SOPHIA G/X	FORE+DRAMA	4	N/A
Shimadzu	Eminence SOPHIA B/L	FORE+DRAMA	4	N/A
Shimadzu	Eminence G/X	FORE+DRAMA	4	N/A
Shimadzu	HEADTOME V	Iterative (FORE+OSEM)	4	16
Siemens	ECAT ACCEL	Iterative (FORE+OSEM)	6	16
Siemens	ECAT EXACT HR+	Iterative (FORE+OSEM)	4	16
Siemens	Biograph 6	Iterative (FORE+OSEM)	4	16
Siemens	Biograph 16	Iterative (FORE+OSEM)	4	14

J-ADNI = Japanese Alzheimer's Disease Neuroimaging Initiative; N/A = not available.

Table 3. Breakdown of visual evaluation of participants

Participant			Template used: number of images		
Visual evaluation	**Clinical diagnosis**	***n***	**Positive**	**Negative**	**Adaptive**
Positive	NC	14	58	0	7
	MCI	41	62	0	35
	AD	43	46	0	38
	Total	98	166	0	80
Negative	NC	44	0	58	51
	MCI	21	0	62	27
	AD	3	0	46	8
	Total	68	0	166	86
Concordance rate			59.0%	41.0%	89.2%
Coefficient of association			not	not	0.80

n = number of participants; NC = normal control; MCI = mild cognitive impairment; AD = Alzheimer's

disease.

Table 4. Comparison of the diagnostic ability

Template	AUC	Cutoff	Sensitivity	Specificity	Accuracy
Positive	0.806	1.80	0.657	0.862	0.729
Negative	0.801	1.40	0.750*	0.793	0.765
Adaptive	0.815	1.40	0.759*	0.793	0.771

AUC = area under the curve. *$P < 0.05$ (difference from the positive template).

FIGURE 1. Workflow of PET-only quantitative evaluation method. First, the PET images are anatomically standardized to either templates. There are three methods: the positive template method, the negative template method, and the adaptive template method. Second, the transformations vector used for standardization is calculated. Third, the EPP-ROI is inverse transformed to individual PET image by using transformations vector.
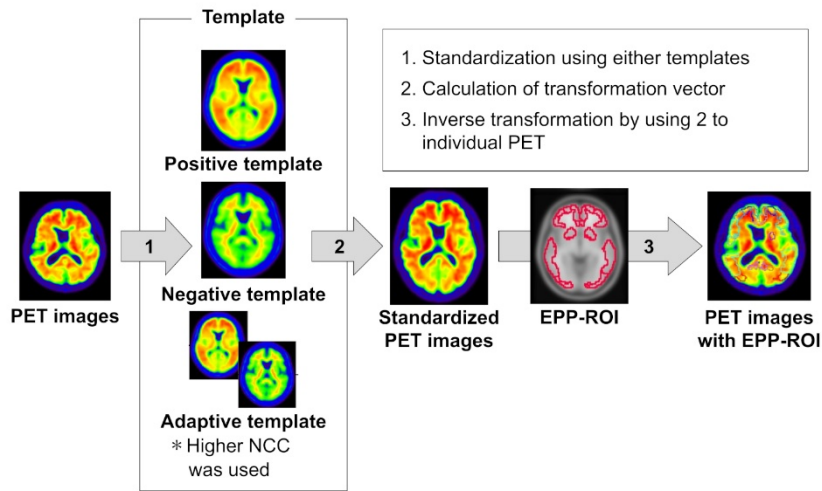
FIGURE 2. The result of NCC. Comparison of NCC between each diagnosis was performed by the

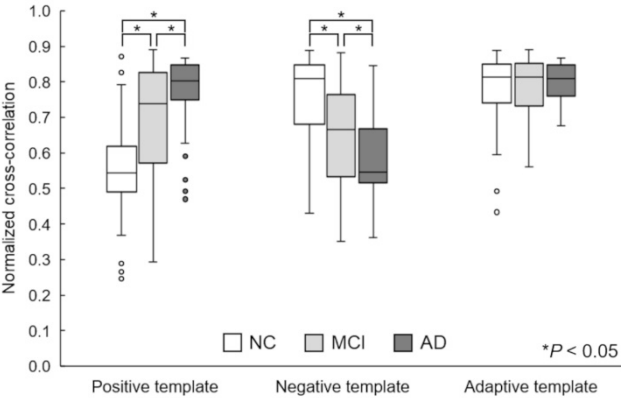Steel-Dwass test. The asterisk indicates that $P < 0.05$.

FIGURE 3. The result of mcSUVR. Comparison of mcSUVR between each diagnosis was performed

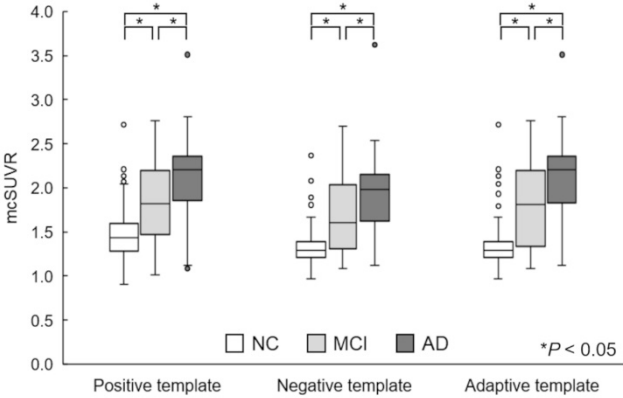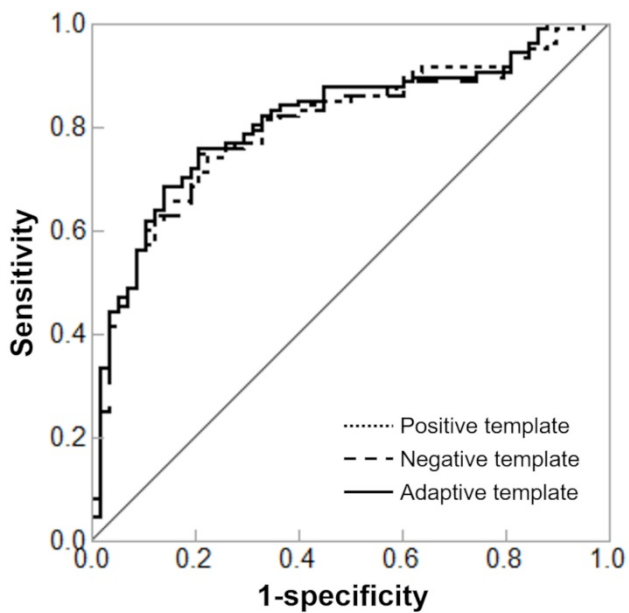by the Steel-Dwass test. The asterisk indicates that $P < 0.05$.

FIGURE 4. The ROC curve of mcSUVR**.** The ROC curve of each template is presented. The AUC of

the positive template-based method, negative template-based method, and the adaptive template-based

method were 0.806, 0.801, and 0.815, respectively.

GRAPHICAL ABSTRACT