Pitfalls of Voxel-Based Amyloid PET Analyses for Diagnosis of Alzheimer’s Disease: Artifacts due to Non-Specific Uptake in the White Matter and the Skull

Akira Arai,¹ Tomohiro Kaneta,¹ Nobuyuki Okamura,² Manabu Tashiro,¹ Ren Iwata,⁴ Kentaro Takanami,¹ Hiroshi Fukuda,⁵ Shoki Takahashi,¹ Kazuhiko Yanai,² Yukitsuka Kudo⁶ and Hiroyuki Arai⁷

¹Department of Diagnostic Radiology, Tohoku University School of Medicine, Sendai, Miyagi, Japan
²Department of Pharmacology, Tohoku University School of Medicine, Sendai, Miyagi, Japan
³Cyclotron Nuclear Medicine, Cyclotron and Radioisotope Center, Tohoku University, Sendai, Miyagi, Japan
⁴Division of Radiopharmaceutical Chemistry, Cyclotron and Radioisotope Center, Tohoku University, Sendai, Miyagi, Japan
⁵Nuclear Medicine and Radiology, Institute of Development, Aging and Cancer, Tohoku University, Sendai, Miyagi, Japan
⁶Innovation New Biomedical Engineering Center, Tohoku University, Sendai, Miyagi, Japan
⁷Department of Geriatrics and Gerontology, Institute of Development, Aging and Cancer, Tohoku University, Sendai, Miyagi, Japan

Two methods are commonly used in brain image voxel-based analyses widely used for dementia work-ups: 3-dimensional stereotactic surface projections (3D-SSP) and statistical parametric mapping (SPM). The methods calculate the Z-scores of the cortical voxels that represent the significance of differences compared to a database of brain images with normal findings, and visualize them as surface brain maps. The methods are considered useful in amyloid positron emission tomography (PET) analyses to detect small amounts of amyloid-β deposits in early-stage Alzheimer’s disease (AD), but are not fully validated.

We analyzed the ¹¹C-labeled 2-(2-[2-dimethylaminothiazol-5-yl]ethenyl)-6-(2-[fluoro]ethoxy)benzoxazole (BF-227) amyloid PET imaging of 56 subjects (20 individuals with mild cognitive impairment [MCI], 19 AD patients, and 17 non-demented [ND] volunteers) with 3D-SSP and the easy Z-score imaging system (eZIS) that is an SPM-based method. To clarify these methods’ limitations, we visually compared Z-score maps output from the two methods and investigated the causes of discrepancies between them. Discrepancies were found in 27 subjects (9 MCI, 13 AD, and 5 ND). Relatively high white matter uptake was considered to cause higher Z-scores on 3D-SSP in 4 subjects (1 MCI and 3 ND). Meanwhile, in 17 subjects (6 MCI, 9 AD, and 2 ND), Z-score overestimation on eZIS corresponded with high skull uptake and disappeared after removing the skull uptake (“scalping”). Our results suggest that non-specific uptakes in the white matter and skull account for errors in voxel-based amyloid PET analyses. Thus, diagnoses based on 3D-SSP data require checking white matter uptake, and “scalping” is recommended before eZIS analysis.

Keywords: Alzheimer’s; amyloid; 3-dimensional stereotactic surface projections; positron emission tomography; statistical parametric mapping

Received November 18, 2013; revised and accepted September 17, 2014. Published online October 11, 2014; doi: 10.1620/tjem.234.175.
Correspondence: Akira Arai, M.D., Ph.D., Department of Diagnostic Radiology, Tohoku University School of Medicine, 1-1 Seiryomachi, Aoba-ku, Sendai, Miyagi 980-8574, Japan.
e-mail: a-arai@rad.med.tohoku.ac.jp

Introduction
Alzheimer’s disease (AD) is the most common form of dementia. The cause and progression of AD are not well understood; however, studies indicate that the disease is associated with plaques and tangles in the brain. Amyloid-β (Aβ) is the main component of these plaques, and the non-invasive detection of Aβ using positron emission tomography (PET) has been developed in recent years. Radioligands that target Aβ deposits have been widely used as amyloid PET tracers, including Pittsburgh compound B (PiB) (Klunk et al. 2004; Nordberg 2004), [¹⁸F]AV-45 (flortetapir) (Choi et al. 2009), and [¹¹C]-labeled 2-(2-[2-dimethylaminothiazol-5-yl]ethenyl)-6-(2-[fluoro]ethoxy)benzoxazole (BF-227) (Kudo 2006; Kudo et al. 2007). For amyloid PET imaging analysis, visual evaluation is an easy and use-
ful method; however, a voxel-by-voxel comparison of tracer distribution in the brain with a normal database (NDB) developed from images with normal findings is thought to be a more sensitive and robust method to detect amyloid burden in the brain.

Voxel-based analyses are commonly performed for brain perfusion single-photon emission computed tomography (SPECT) and $[^{18}F]$fluorodeoxyglucose (FDG) PET in dementia work-ups. Two methods are widely used for voxel-based analyses in clinical diagnosis and research: 3-dimensional stereotactic surface projections (3D-SSP) (Minoshima et al. 1994, 1995) and statistical parametric mapping (SPM) (Friston et al. 1995). These two methods have much in common; both register an individual brain image to a standard brain coordinate system through a template image (anatomical standardization [AS]) to normalize the size and shape of the brain among subjects. Next, a voxel-by-voxel statistical comparison of the individual brain image and NDB is performed in the common coordinate system. However, several reports revealed that individual brain perfusion SPECT analyses using 3D-SSP and SPM-based methods sometimes show different results under certain conditions (Ishii et al. 2001; Onishi et al. 2011a, b; Yamamoto and Onoguchi 2011). The different characteristics and pitfalls of each method have been revealed. Voxel-based amyloid PET analyses have also been reported using these methods (Aalto et al. 2009; Shao et al. 2010; Shin et al. 2010; Kaneta et al. 2011). Previously, AS errors in 3D-SSP amyloid PET analysis were reported (Kaneta et al. 2011); however, the limitations of voxel-based-methods specific to amyloid PET have not been thoroughly investigated. We believe that it is important to determine the differences in the natures of the methods and to clarify the limitations of each method to avoid the risk of a dementia misdiagnosis. In this study, we performed a head-to-head comparison of results analyzed by 3D-SSP and SPM-based methods and investigated the limitations and their causes for each method by focusing on the differences in results.

Methods

Subjects

The present study enrolled 20 individuals with amnestic mild cognitive impairment (MCI), 19 patients with AD, and 17 non-demented (ND) subjects, which is the same distribution as in a previous study by our colleagues (Kaneta et al. 2011). The demographic characteristics of the subjects are shown in Table 1. The MCI and probable AD diagnoses followed MCI clinical criteria ( Petersen et al. 1999) and National Institute of Stroke-Alzheimer’s Disease and Related Disorders Association criteria (McKhann et al. 1984), respectively. The ND group was recruited from among volunteers; none were receiving centrally acting medication, had cognitive impairment, or had cerebrovascular lesions identified on magnetic resonance imaging (MRI). Cognitive performance was evaluated with Mini-Mental State Examination (MMSE) scores. The study protocol was approved on June 20, 2005 by the Committee on Clinical Investigation and the Advisory Committee on Radioactive Substances at Tohoku University School of Medicine, and has therefore been performed in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki and all subsequent revisions. After describing the study to the patients and subjects, written informed consent was obtained from all participants.

BF-227 PET procedure

All subjects underwent PET with BF-227. BF-227 and its N-demethylated derivative (a precursor of BF-227) were custom-synthesized by Tanabe R&D Service Co., Ltd. (Osaka, Japan). BF-227 was synthesized from a precursor by N-methylation in dimethyl sulfoxide using $[^{11}C]$methyl triflate. BF-227 PET was performed using a PET SET-2400W scanner (Shimadzu Inc., Kyoto, Japan) with a spatial resolution of 4 mm (transaxial) and 4.5 mm (axial) at full-width half-maximum in the center of the field-of-view. For attenuation correction, a transmission scan was performed using $^{40}$Ge/Ga sources for 7 min. The BF-227 PET scan was performed for 60 min after an intravenous injection of 211-366 MBq of BF-227 with the subjects’ eyes closed. The amyloid PET images used in this study were obtained using data acquired 40-60 min after the injection of BF-227.

Voxel-based analyses and evaluations of individual results

Amyloid PET images with BF-227 were analyzed with 3D-SSP and the easy Z-score imaging system (eZIS) (Matsuda et al. 2007), which is a software program that uses SPM version 2002 (SPM2) for AS. The analysis with 3D-SSP was performed according to the modified method reported by our colleagues (Kaneta et al. 2011) to avoid errors caused by using an FDG template for the AS of amyloid PET images. For AS in eZIS analyses, we used a BF-227 PET template developed from anatomically standardized and normalized BF-227 PET images. NDBs for voxel-based comparisons with normal subjects were created from BF-227 PET images of all 17 ND subjects. Individual results were visualized as surface Z-score maps within a range of $Z = 1-5$ normalized to the cerebellum according to the following formula: $Z$-score = (individual value – NDB mean)/NDB standard deviation. A significant abnormality was defined as $Z \geq 2$.

Table 1. Demographic detail of the subjects in this study.

<table>
<thead>
<tr>
<th></th>
<th>Number</th>
<th>Gender</th>
<th>Age</th>
<th>MMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>ND</td>
<td>17</td>
<td>M/F = 7/10</td>
<td>67.0 ± 4.1</td>
<td>29.9 ± 0.3</td>
</tr>
<tr>
<td>MCI</td>
<td>20</td>
<td>M/F = 10/10</td>
<td>76.6 ± 4.7</td>
<td>25.5 ± 2.3</td>
</tr>
<tr>
<td>AD</td>
<td>19</td>
<td>M/F = 5/14</td>
<td>73.7 ± 7.0</td>
<td>20.0 ± 3.5</td>
</tr>
</tbody>
</table>

ND, non-demented; MCI, mild cognitive impairment; AD, Alzheimer’s disease; M, male; F, female; MMSE, Mini-Mental State Examination.

Values are expressed as the number of subjects or mean ± s.d.
Individual Z-score maps created by the 3D-SSP and eZIS methods were visually compared, with an emphasis on the severity and extent of abnormalities. We initially focused on the results of the ND subjects to detect errors or problems in the voxel-based statistical mapping of amyloid PET data. Afterward, we expanded the evaluation to the AD and MCI subjects. Then, we focused on differences in results from the two methods and examined the original PET images to investigate the causes of the analytical errors.

**Results**

**Comparisons of individual results in the ND group**

All ND subjects showed low uptake in their cortices on the original BF-227 PET images, which indicated no or few Aβ depositions. Fig. 1 shows a representative example of a ND subject’s original BF-227 PET image and the surface Z-score maps output from the 3D-SSP and eZIS methods. Neither method detected a high Z-score area ($Z \geq 2$) indicating a significant difference from the NDB. The results from the voxel-based methods were consistent with the normal findings on the original images.

However, significant abnormalities were demonstrated in 53% (9/17) of the ND subjects on 3D-SSP and in 41% (7/17) of the ND subjects on eZIS, despite normal findings on the original PET images. Moreover, obvious discrepancy between the 3D-SSP and eZIS findings was seen in 29% (5/17) of the ND subjects in the severity and extent of the abnormalities. We classified these types of discrepancy into 3 patterns: pattern 1, a broader and stronger abnormality on 3D-SSP than on eZIS; pattern 2, a broader and stronger abnormality on eZIS than on 3D-SSP; and pattern 3, clearly different abnormality distributions on 3D-SSP and eZIS.

Fig. 2 shows 3 representative examples of discrepancy in the results between the two methods. Case 1 was a 62-year-old ND woman in whom the abnormality appeared broader and more severe on 3D-SSP than on eZIS ("pattern 1 discrepancy"; Fig. 2a). Case 2 was a 66-year-old ND woman in whom eZIS demonstrated a global abnormality while 3D-SSP showed a normal pattern ("pattern 2 discrepancy"; Fig. 2b). Case 3 was a 67-year-old ND man in whom 3D-SSP demonstrated a broader abnormality than eZIS. Further, eZIS showed a remarkable abnormality in the medial occipital region ("pattern 3 discrepancy"; Fig. 2c).

**Individual comparisons of results in patients with AD and MCI**

3D-SSP analyses detected abnormal findings in 90% (18/20) of the MCI subjects and in 79% (15/19) of the AD subjects. Meanwhile, eZIS analyses detected abnormal findings in 65% (13/20) of the MCI subjects and 84% (16/19) of the AD subjects. Obvious discrepancy between 3D-SSP and eZIS was seen in 45% (9/20) of the MCI subjects and in 68% (13/19) of the AD subjects (Table 2). Among all subjects in the AD, MCI, and ND groups, 48% (27/56) showed such discrepancy.

All 3 patterns of discrepancy mentioned above were also found in subjects with AD and MCI (Table 2). Only "pattern 1 discrepancy" was more frequently seen in the ND group than in the AD group.

**Relationships between the discrepancy patterns and the original PET images**

An examination of the original PET images from Case 1 ("pattern 1 discrepancy") showed relatively high uptake in the white matter compared to that in the cortex (Fig. 2a). Among the 6 subjects showing "pattern 1 discrepancy", 4 subjects (3 ND subjects and 1 MCI subject) demonstrated relatively high white matter uptake. Further, Case 3 ("pattern 3 discrepancy") also demonstrated relatively high white matter uptake (Fig. 2c).

Meanwhile, Case 2 ("pattern 2 discrepancy") demonstrated a broad, high uptake in the skull (Fig. 2b, arrows). In addition, Case 3 with "pattern 3 discrepancy" demonstrated high uptake in part in the skull, which corresponded to the high Z-score area on eZIS (Fig. 2c, arrows). Such
high uptake in the skull corresponded to the high Z-score area on eZIS in 17 subjects (2 ND subjects, 6 MCI subjects, and 9 AD subjects): 9 subjects (1 ND subject, 3 MCI subjects, and 5 AD subjects) out of 12 cases with “pattern 2 discrepancy” and in 8 subjects (1 ND subject, 3 MCI subjects, and 4 AD subjects) out of 9 cases with “pattern 3 discrepancy”.

Discussion
In this study, we observed obvious differences in individual results between the 3D-SSP and SPM-based amyloid PET analysis methods. Previous studies using voxel-based methods for amyloid PET have shown similar results to group analyses using 3D-SSP and SPM-based methods.
Pitfalls of Voxel-Based Amyloid PET Analysis

(Aalto et al. 2009; Shao et al. 2010; Shin et al. 2010; Kaneta et al. 2011), but did not focus on differences in the individual results between the two methods. To the best of our knowledge, this report is the first to perform a head-to-head comparison between the results of these two methods for amyloid PET analyses, and to demonstrate the differences in the results described here. We believe that our results could contribute to individual diagnoses of dementia in clinical settings, and the future development of methods of analysis.

Our results suggest that abnormalities detected on 3D-SSP and eZIS amyloid PET analyses might not always represent increased uptake in the cortex, but are sometimes influenced by the uptake of contiguous tissues such as the white matter and skull, although the original target was the cortical region. The cases that demonstrated stronger abnormalities on 3D-SSP than on eZIS (“pattern 1 discrepancy”) showed a tendency to have relatively high white matter uptake. In the AS of PET images, the white matter uptake may be partially included in the cortical region because of local anatomical variability between individuals and the low spatial resolution of PET. Based on our results, even a small amount of relatively high white matter uptake in regions near the surface may impact 3D-SSP results, because the maximum counts in this region are extracted as cortical uptake in the 3D-SSP analysis, while the mean cortical Z-score values are projected in the eZIS analysis. Many amyloid PET tracers, such as PiB, AV-45, and BF-227, sometimes show non-specific, high uptake in the white matter. Because ND and MCI subjects often demonstrate higher uptake in the white matter than in the cortex, the false extraction of white matter uptake occurs frequently in these groups, and may increase the false-positive rate in 3D-SSP. This problem is critical in the detection of amyloid burden in the brain, especially for normal cognitive subjects. This pitfall of 3D-SSP should be recognized and

Fig. 3. Anatomically standardized BF-227 PET images and the results of eZIS with or without the removal of the skull uptake (“scalping”). Shown are the data of Case 2, a 66-year-old ND woman. (a-c) An anatomically standardized image of the original BF-227 PET image (a) and the results of eZIS (b, c). (a) The white line represents the contour of the standard brain. Some portions of the skull uptake are seen in the white line (arrows). (b) A high Z-score area is observed at the peripheral region of the brain (arrowheads). (c) The abnormalities are seen globally on the surface map (arrows). (d-f) An anatomically standardized image after “scalping” (d) and the results of eZIS (e, f). (d) The edge of the brain matches the white line. (e, f) The high Z-score area at the peripheral region of the brain and the global abnormalities on the surface map disappeared after “scalping”.

Table 2. 27 cases showing significant discrepancies in results between 3D-SSP and eZIS.

<table>
<thead>
<tr>
<th>Discrepancy pattern</th>
<th>ND (N = 17)</th>
<th>MCI (N = 20)</th>
<th>AD (N = 19)</th>
<th>Total (N = 56)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pattern 1</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>Pattern 2</td>
<td>1</td>
<td>4</td>
<td>7</td>
<td>12</td>
</tr>
<tr>
<td>Pattern 3</td>
<td>1</td>
<td>3</td>
<td>5</td>
<td>9</td>
</tr>
<tr>
<td>Total</td>
<td>5</td>
<td>9</td>
<td>13</td>
<td>27</td>
</tr>
</tbody>
</table>

Pattern 1: a broader and stronger abnormality demonstrated on 3D-SSP than on eZIS.
Pattern 2: a broader and stronger abnormality demonstrated on eZIS than on 3D-SSP.
Pattern 3: a clearly different distribution of the abnormality on 3D-SSP and eZIS.
the original PET images re-examined, even if the 3D-SSP results were positive. Recently, a new method was proposed that uses MRI to identify cortical voxels and measures cortical uptake at the same locations on a co-registered amyloid PET image (Frey et al. 2012). This may be a solution for removing non-specific high uptake in the white matter.

On the other hand, cases with stronger abnormalities on eZIS tended to show high skull uptake. Skull uptake is considered non-specific, as no autopsy case diagnosed as AD showed Aβ deposits in the bone marrow (Skodras et al. 1993). To clarify the relationship between high uptake in the skull and strong abnormalities on eZIS, we inspected the images after AS by eZIS (SPM2). On the AS image from Case 2, the skull uptake was partially included inside the brain area in the standard coordinate system (Fig. 3a), which suggests an AS failure. On the axial Z-score image, a high Z-score area was observed at the peripheral zone of the brain where the skull uptake was misregistered (Fig. 3b, c). Thus, we hypothesized that such mis-registered skull uptakes located in the brain area resulted in strong abnormalities on eZIS. In a voxel-based analysis, the removal of radioactivities outside of the brain (“scalping”) is sometimes performed to avoid effects on the analysis results. However, in clinical practice, eZIS analysis has been commonly performed without “scalping”. Although the eZIS program includes a “masking” procedure to remove extra-brain tissue activities, mis-registered skull uptake cannot be removed, as this procedure is performed after AS. Therefore, we manually removed it first (“scalping”), and then performed the eZIS analyses again to clarify the effects of high uptake in the skull. As of result of “scalping”, the AS errors were corrected (Fig. 3d) and the strong abnormalities on eZIS disappeared (Fig. 3e, f). These AS errors were seen in 25% (14/56) of all subjects. Similar errors have been reported in the analysis of T1-weighted MRI using SPM2 and were eliminated by the removal of signals from non-brain tissues (Fein et al. 2006). On the other hand, 3D-SSP is relatively insulated from the influence of such tracer distributions because it uses a number of landmarks for AS. Thus, the routine use of “scalping” for SPM-based methods might be helpful in reducing analytic errors. Essentially, “scalping” should be performed for all tracers that show uptake in the skull or skin.

The artifacts revealed in the present study have not been considered in previous studies using a voxel-based amyloid PET analysis. As described above, false surface projection and overestimation of the cortical uptake frequently occurs in normal subjects in 3D-SSP analyses. A previous study by our colleagues showed that 3D-SSP was useful for detecting differences among groups (Kaneta et al. 2011), but the presence of overestimated data in the ND group may have reduced the differences from the AD and MCI groups. In another study, most AD patients and MCI-to-AD converters showed high Z-scores in the temporal cortices in the eZIS results (Shao et al. 2010). Some of these results might have been influenced by artifacts resulted from high skull uptake. However, similar findings in which 79% (15/19) of AD patients showed significant abnormalities in the temporal cortices were also observed in our results, even after “scalping”.

“Pattern 1 discrepancy” in 1 MCI subject and 1 AD subject was not explained by artifacts on 3D-SSP resulted from high white matter uptake. Moreover, “scalping” before eZIS analysis did not resolve “pattern 2 discrepancy” in both subjects (1 MCI subject and 2 AD subjects) and “pattern 3 discrepancy” in 1 AD subject. The causes of discrepancy in these 6 subjects remained unclear, but the combined effects of other factors were suspected. In analyses of perfusion SPECT and FDG-PET, different factors are known to cause discrepancies in the results of 3D-SSP and SPM-based methods. For example, a high count resulting from local noise can be directly projected onto the surface map in a 3D-SSP analysis, whereas a smoothing process improves the signal-to-noise ratio in an eZIS analysis. It has also been reported that SPM is more sensitive to artifacts derived from localized cortical atrophy than 3D-SSP (Ishii et al. 2001). Attention should be paid to these factors in amyloid PET analyses.

In this study, we used relatively severe diagnostic criteria to detect subtle findings (i.e., a significant abnormality was defined as Z ≥ 2). Generally, in a clinical setting, a significant abnormality might be defined as a Z-score around 5. Thus, the positive rates for abnormalities or discrepancies in the results must be higher in the present study than those found in clinical settings.

In conclusion, our study demonstrated obvious differences in the results of 3D-SSP and eZIS for the voxel-based analysis of amyloid PET. These discrepancies may be mainly caused by non-specific uptake in the white matter and the skull. Thus, when making a diagnosis based on 3D-SSP results, examining the white matter uptake on the original PET images is necessary. When making a diagnosis based on eZIS results, performing “scalping” before the analyses is recommended. These matters will be helpful not only in a clinical setting, but also in developing new methodologies for voxel-based amyloid PET analyses.

Acknowledgments

We appreciate the technical assistance provided by Seiichi Watanuki of CYRIC in Tohoku University (Sendai, Japan) and Satoshi Minoshima of Department of Radiology in University of Washington (WA, USA).

Conflict of Interest

The authors declare no conflict of interest.

References

Pitfalls of Voxel-Based Amyloid PET Analysis

1660.


