

**Table 1.** Normal database of FDG-PET and IMP-SPECT

| Method    | No. | Sex (M/F) | Age  |
|-----------|-----|-----------|------|
| FDG-PET   | 37  | 23/11     | 59.0 |
| IMP-SPECT | 18  | 7/F       | 69.9 |

FDG-PET, 2-[<sup>18</sup>F]-fluoro-2-deoxy-D-glucose positron emission tomography; IMP-SPECT, *N*-isopropyl-*p*-<sup>123</sup>I iodoamphetamine single photon emission computed tomography

projection filter. Chang's attenuation correction and scattering correction using the triple energy window method were applied to the reconstructed images. The in-plane spatial resolution was 11.1 mm in FWHM. The final image slices were set parallel to the orbitomeatal line and were obtained at an interval of 3.44 mm through the entire brain. The rCBF images were quantitated according to the IMP-ARG method.<sup>18</sup>

#### Statistical images (3D-SSP)

The original FDG-PET and IMP-SPECT data were analyzed by an iSSP (SSP; Nihon Mediphysics) program, which was modified based on the NEUROSTAT program (Dr. Minoshima, Department of Radiology and Bioengineering, University of Washington, Seattle, WA, USA). After rotation and centering of the data set, the original data were realigned to the bicommissural line (AC-PC) and transformed into a stereotactic standard Talairach space. The cortical peak activity was projected onto the brain surface, and the peak value was projected back and assigned to the originating surface pixel. The extracted data sets were displayed on eight different angles, including the lateral, medial, superior, inferior, anterior, and posterior views. With the 3D-SSP programs in NEUROSTAT, the pixel values were normalized to the whole brain, thalamus, pons, and cerebellum (Fig. 1). The pixel values of an individual's image were compared with a normal database that originated at our institution (Table 1). The normal database was built as follows: 18 normal subjects (7 men, 11 women; mean age 69.9 years) for IMP-SPECT and 37 normal subjects (23 men, 14 women; mean age 59.0 years) for FDG-PET.

#### Statistical analysis/visual interpretation

Four nuclear medicine physicians randomly interpreted FDG-PET images of 14 AD patients and 7 NCs and IMP-SPECT images of 9 NCs. The interpretation of each image was performed at separate sessions, respectively. The statistical analyses were conducted using the following procedure. Two areas of the brain, the posterior cingulate gyri/precunei and parietotemporal region

in each image (which were determined as characteristic for AD in advance of the study) were evaluated.<sup>9</sup> The degree of reduction in glucose metabolism or rCBF were interpreted, and a score of 5 was assigned if the change was considered "apparent decrease," which is characteristic of AD. Accordingly, scores of 4 to 1 were assigned to the changes that were considered "probable decrease," "unclear," "probable not a decrease," and "apparently not a decrease," respectively. Then, the reliability was evaluated using five steps similar to the regional evaluation.

When laterality was observed in a visual interpretation, the evaluation of a regional change was performed as follows. The scores assigned to the right and left sides (of the brain) of FDG-PET were compared for each side of the brain, and the higher score was recorded as the case's FDG-PET score for the specific area of the brain. The side of the higher FDG-PET score was considered a regional finding, and the IMP-SPECT score of the same side was recorded as the case's IMP-SPECT score. If the right- and left-side FDG-PET scores were the same, the higher score in IMP-SPECT was recorded as the regional finding. In fact, there were no discrepancies between the FDG-PET scores and the IMP-SPECT images in terms of laterality.

Second, the ratings of four readers were pooled for each area and for each image. The ROC analysis was performed for the graphics presentation, and the area under the curve was calculated to express the diagnostic accuracy of each image and each area numerically. Confidence intervals (CIs) of the AUC were used to test the difference in overall diagnostic accuracy between the images, the upper and lower limits of which were calculated either by adding or subtracting the standard error of the AUC times 1.96 for 95% CI and 1.65 for 90% CI. Because our primary interest lies in the difference in the diagnostic ability of FDG-PET from IMP-SPECT, we did not compare the difference in the AUC between the areas of the brain and did not take multiplicity in comparison into account.  $P < 0.05$  was considered statistically significant. The statistical analyses were conducted using SPSS for Windows.

#### Correlation between FDG-PET and IMP-SPECT

To assess the agreement between two modalities, ROIs were set in the posterior cingulate gyri/precunei and parietotemporal region on 3D-SSP images, of which the pixel values were normalized to the whole brain bilaterally (Fig. 2). Then the mean Z-values of the right and left hemisphere were summed and used for analysis. We calculated kappa ( $\kappa$ ) statistics, the proportion of concordance, and the Spearman's correlation coefficients of

## Presentation of 3D-SSP in the same probalbe AD patinets

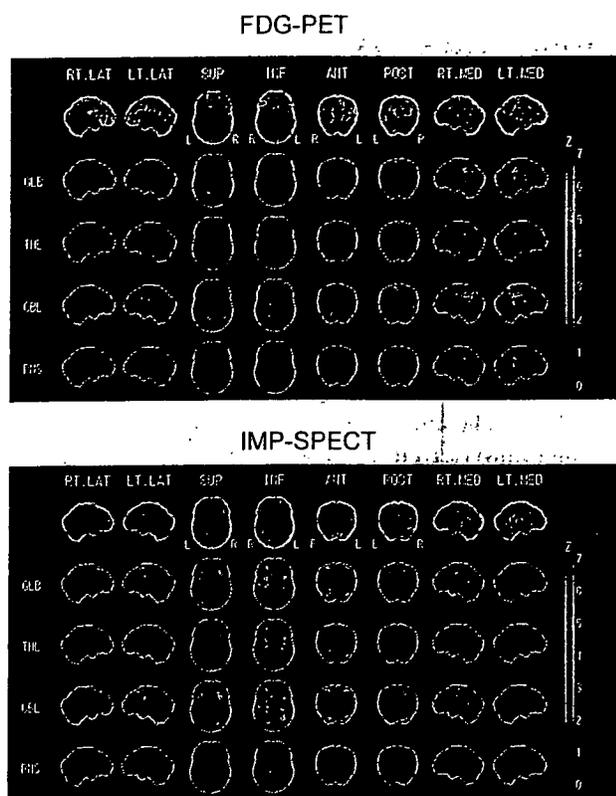
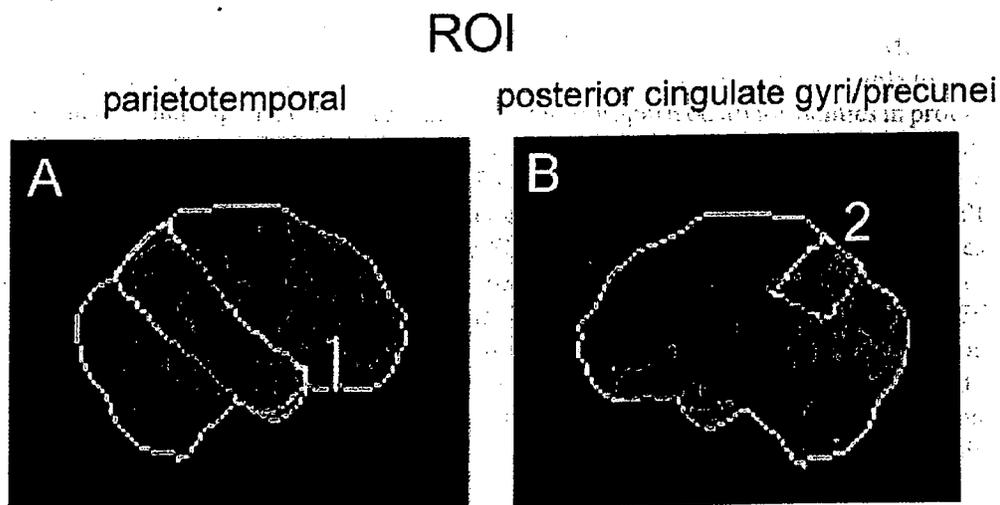


Fig. 1. Z-score images of three-dimensional stereotactic surface projections (3D-SSP) are shown from a representative Alzheimer's disease (AD) patient. These images demonstrate a decrease in the glucose metabolism and regional cerebral blood flow (rCBF). *Upper panel:* 2-[<sup>18</sup>F]-Fluoro-2-deoxy-D-glucose positron emission tomography (FDG-PET). *Lower panel:* N-Isopropyl-p-<sup>123</sup>I iodoamphetamine single photon emission computed tomography (IMP-SPECT). The extracted data sets are displayed on eight different angles, including the lateral (RT LAT, LT LAT), medial (RT MED, LT MED), superior (SUP), inferior (INF), anterior (ANT), and posterior (POST) views. With the 3D-SSP programs in NEUROSTAT, the pixel values are normalized to the whole brain (GLB), thalamus (THL), cerebellum (CBL), and pons (PNS)

Fig. 2. Regions of interest (ROI) were set in the posterior cingulate gyri/precune (B) and parietotemporal region (A) on 3D-SSP images of FDG-PET, in which the pixel values were normalized to the whole brain bilaterally



Z-values in each area (the posterior cingulate gyri/precune and parietotemporal region). For assessing  $\kappa$  statistics and the proportion of concordance, Z-values were first categorized according to their quartiles. Then the lowest quartile was considered as the "decrease" category. The  $\kappa$  statistic is defined as the agreement beyond chance divided by the amount of agreement possible by chance. As in most studies,  $\kappa > 0.75$  was taken to represent excellent agreement beyond chance, 0.40–0.75 to mean fair agreement, and  $< 0.40$  to mean poor agreement.<sup>19</sup>

## Results

Figure 3 shows the results of the clinical diagnosis of AD by 3D-SSP. The AUC of FDG-PET was  $0.952 \pm 0.023$ , and that of IMP-SPECT was  $0.935 \pm 0.026$ . There was no significant difference between FDG-PET and IMP-SPECT. Table 2 shows the results of each interpreter and the sum as the diagnosis ability. The sensitivity and specificity of each modality were 86%, and 97% for FDG-PET and 70%, and 100% for IMP-SPECT, respectively (Table 3).

Table 2. Results of visual inspection of 3D-SSP by each interpreter

| Interpreter | FDG-PET |       | IMP-SPECT |       |
|-------------|---------|-------|-----------|-------|
|             | AUC     | SE    | AUC       | SE    |
| A           | 0.939   | 0.057 | 0.976     | 0.026 |
| B           | 0.954   | 0.047 | 0.96      | 0.041 |
| C           | 0.888   | 0.074 | 0.821     | 0.091 |
| D           | 0.995   | 0.01  | 1         | 0     |
| Sum         | 0.952   | 0.023 | 0.935     | 0.026 |

3D-SSP, three-dimensional stereotactic surface projections; AUC, area under the curve; SE, standard error

Fig. 3. Receiver operating characteristic (ROC) curves for clinical diagnosis of AD with visual inspection by experts. There was no significant difference between PET and SPECT. AUC, area under the curve

## Clinical diagnosis of Alzheimer disease by 3D SSP

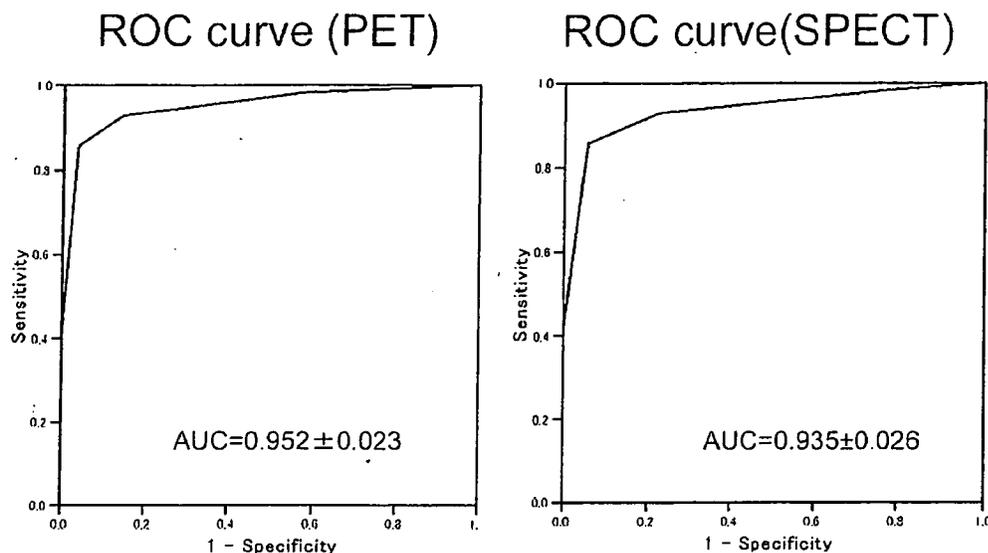


Table 3. Results of visual inspection of 3D-SSP by each interpreter: sensitivity and specificity

| Interpreter | FDG-PET         |                 | IMP-SPECT       |                 |
|-------------|-----------------|-----------------|-----------------|-----------------|
|             | Sensitivity (%) | Specificity (%) | Sensitivity (%) | Specificity (%) |
| A           | 93              | 100             | 64              | 100             |
| B           | 79              | 100             | 79              | 100             |
| C           | 86              | 86              | 64              | 100             |
| D           | 86              | 100             | 71              | 100             |
| Sum         | 344             | 386             | 278             | 400             |
| Average     | 86              | 97              | 70              | 100             |

Figure 4 shows the regional evaluation of the posterior cingulate gyri/precuneus and the parietotemporal region. The AUC of FDG-PET was  $0.935 \pm 0.028$ , and that of IMP-SPECT was  $0.807 \pm 0.046$  in the posterior cingulate gyri/precuneus. A significant difference between FDG-PET and IMP-SPECT was observed ( $P < 0.05$ ). The AUC of FDG-PET was  $0.871 \pm 0.038$ , and that of IMP-SPECT was  $0.802 \pm 0.046$  in the parietotemporal region. There was no significant difference between FDG-PET and IMP-SPECT detected for this region.

Interrater agreement among four nuclear medicine physicians was evaluated by intraclass correlation coefficient with the Spearman-Brown correction. For both PET and SPECT, the agreement in the ratings among four raters were high: PET 0.94, SPECT 0.96.

There was a high correlation of Z-values between FDG-PET and IMP-SPECT in the parietotemporal region (Spearman's  $r = 0.82$ ,  $P < 0.001$ ). The correlation

of the two modalities in the posterior cingulate gyri/precuneus region was lower than that in the parietotemporal region (Spearman's  $r = 0.63$ ,  $P < 0.016$ ). In terms of the proportion of concordance and  $\kappa$  statistics, there was a perfect agreement in the parietotemporal region (both 1.0,  $P < 0.001$ ). The proportion of concordance in the posterior cingulate gyri/precuneus region was comparable (0.71), but agreement beyond chance evaluated by the  $\kappa$  statistics was poor and not significantly greater than zero (0.15,  $P = 0.57$ ).

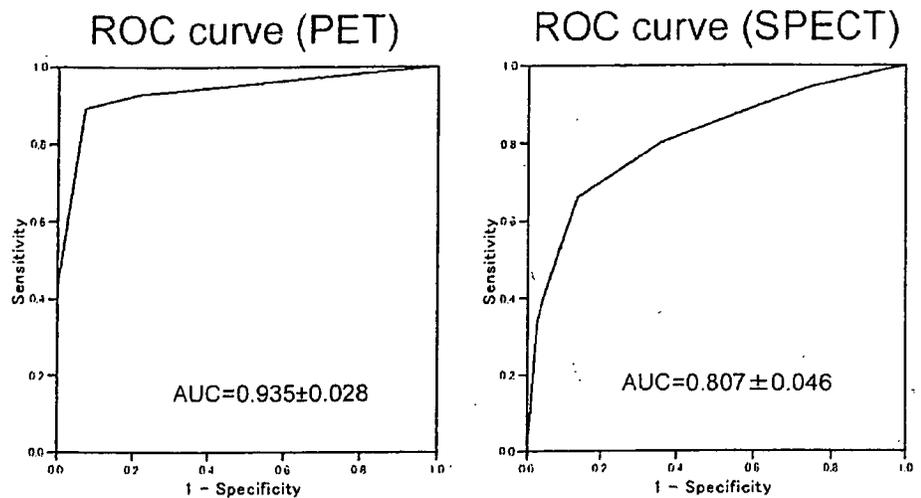
### Discussion

Voxel-based statistical mapping methods, such as 3D-SSP and SPM, have been reported to be useful when delineating AD individuals from normal subjects.<sup>9–16,20,21</sup> Honda et al showed that 3D-SSP enhanced the specificity of SPECT inspection by nuclear medicine physicians

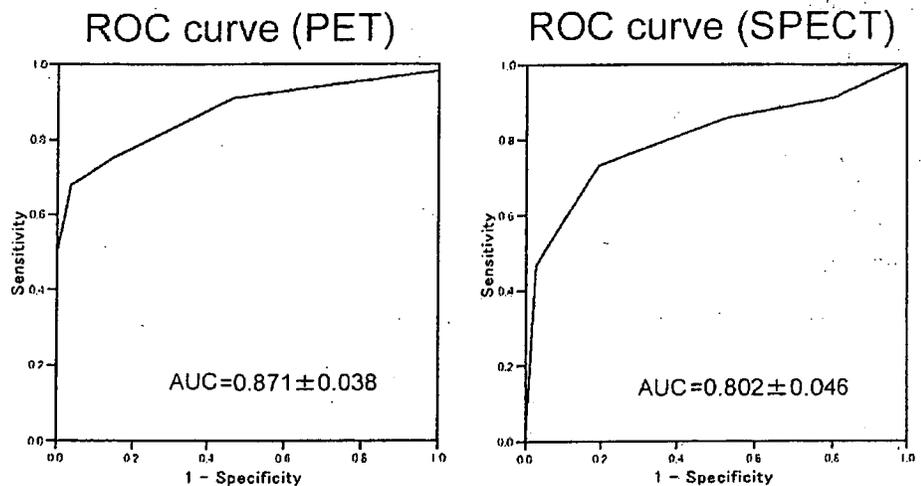
Fig. 4. ROCs for clinical diagnosis of AD with visual inspection of the posterior cingulate gyri/precuneus (A) and the parietotemporal region (B). For evaluation of the posterior cingulate gyri/precuneus, PET is significantly superior to SPECT

## Clinical diagnosis of Alzheimer disease

### A. Evaluation on the posterior cingulate gyri/precuneus



### B. Evaluation on the parietotemporal region



and the performance of the automated diagnosis, when focusing on the posterior cingulate gyri and precuneus, exceeded that of the physicians when evaluating SPECT with and without 3D-SSP.<sup>11</sup> These reports showed that voxel-based statistical mapping methods may add useful information when diagnosing AD. The usefulness of such statistical methods has become clear, and these methods are used frequently clinically.

A few studies have compared FDG-PET and rCBF SPECT directly in terms of diagnosing AD. Mielke et al. reported that both PET and SPECT are able to distinguish AD patients from controls. On the other hand, for differentiation between AD and vascular dementia (VD),

PET was shown to be superior to SPECT.<sup>22</sup> Messa et al. reported that both PET and SPECT were able to detect characteristic temporoparietal abnormalities in probable AD, whereas for the evaluation for other associative areas PET was superior to SPECT.<sup>23</sup> Silverman reported that higher diagnostic accuracy was obtained with PET than with SPECT, and the increased accuracy was estimated to be approximately 15%–20%.<sup>24</sup> Herholz et al. reported that correspondence between PET and SPECT was limited to the main finding of temporoparietal and posterior cingulate functional impairment in mild to moderate AD using SPM.<sup>25</sup> Although many studies have shown that PET is superior to SPECT in terms of diag-

nosing AD, it seems that a decisive difference has not been determined concerning the diagnostic ability of PET and SPECT as far as the posterior cingulate gyri/precuneus and the parietotemporal region are concerned, which are the characteristic regions for AD.

In the present study, our purpose was to compare directly the diagnostic ability of FDG-PET and IMP-SPECT using 3D-SSP and focusing on as the posterior cingulate gyri/precuneus and the parietotemporal region. Our results show that there was no significant difference between FDG-PET and IMP-SPECT in terms of diagnosing AD using 3D-SSP. However, there was a tendency that FDG-PET was slightly superior to IMP-SPECT, especially when evaluating the posterior cingulate gyri/precuneus. These results are consistent with those of previous studies and our clinical experience. We considered two possibilities to explain the findings. First, a possible reason why FDG-PET is superior to IMP-SPECT is that the metabolic change due to degeneration might slightly proceed to a decrease in rCBF. Second, PET has the advantage of greater sensitivity and greater spatial resolution. Both possibilities were plausible.

We found that there was a high correlation of Z-values between FDG-PET and IMP-SPECT in the parietotemporal region; in contrast, there was not a good correlation in the posterior cingulate gyri/precuneus region. These results showed that there was the difference between the glucose metabolic and rCBF changes in the posterior cingulate gyri/precuneus region. We considered two possibilities regarding this problem. First, coupling the glucose metabolic change and rCBF change might depend on the degree of the degeneration. Therefore, in the posterior cingulate gyri/precuneus region, because the degenerative change is slight, the change in glucose metabolism might be greater than that of the rCBF change. Second, distributions of the glucose metabolism and rCBF were not the same in every area of the brain. Sakamoto and Ishii showed that there are great glucose metabolic and rCBF differences between the medial temporal lobe, the cerebellum, and other brain regions in the normal human brain.<sup>26</sup> If there was a difference between the glucose metabolism and rCBF in the posterior cingulate gyri/precuneus region, it was not impossible to understand this result. To evaluate the finding in the posterior cingulate gyri/precuneus region, further studies are needed from a radiological and/or pathological approach.

There are three restrictions to the present study. First, the normal controls and normal database differed, respectively, for FDG-PET and IMP-SPECT. Ideally, for each examination, the normal controls and database should be the same. However, prescribing nuclear medicine to healthy subjects twice in a short time is difficult.

Second, in the present study, we analyzed only moderately affected AD patients. Recently, Dobert et al. evaluated the diagnostic potential of FDG-PET and SPECT in terms of early detection and the differential diagnosis of early dementia in the same patients and showed that PET was the superior imaging method, especially for detecting early AD or mixed-type dementia.<sup>27</sup> If we analyze early or mild AD, the results may change. Third, owing to a relatively small number of subjects, the statistical power might have been limited in the present study. In an illustrative example dataset where a total of about 160 subjects were studied in which the correlation of the underlying bivariate binormal distribution is 0.75, statistical power to detect a 0.10 difference in the AUC of the two modalities was estimated as 0.75.<sup>28</sup> In the present study, a total of about 90 subjects were studied where the AUC difference of more than 0.1 had been expected, and the statistical power of the analysis would have been retained to a similar degree. Thus, the present finding of no difference could be interpreted as relevant. Ideally, further studies with larger sample sizes are needed to examine whether these two modalities are really equivalent in diagnosing AD given that we did not observe a statistically significant difference in the AUC between FDG-PET and IMP-SPECT.

Using 3D-SSP, our results show that FDG-PET and IMP-SPECT have similar diagnostic ability in moderately affected AD patients. However, there was the tendency that the ability of FDG-PET to detect probable AD was superior to that of IMP-SPECT. Specifically, for regional glucose metabolism or cerebral blood flow analysis of the posterior cingulate gyri/precuneus, PET was significantly superior to SPECT.

## References

1. Duara R, Grady C, Haxby J, Sundaram M, Cutler NR, Heston L, Moore A, et al. Positron emission tomography in Alzheimer's disease. *Neurology* 1986;3:879–87.
2. Heiss WD, Szelies B, Kessler J, Herholz K. Abnormalities of energy metabolism in Alzheimer's disease studied with PET. *Ann N Y Acad Sci* 1991;640:65–71.
3. Silverman DH, Small GW, Chang CY, Lu CS, Kung De Aburto MA, Chen W, et al. Positron emission tomography in evaluation of dementia: regional brain metabolism and long-term outcome. *JAMA* 2001;286:2120–7.
4. Bradley KM, O'Sullivan VT, Soper ND, Nagy Z, King EM, Smith AD, et al. Cerebral perfusion SPECT correlated with Braak pathological stage in Alzheimer's disease. *Brain* 2002; 125:1772–81.
5. Ishii K. Clinical application of positron emission tomography for diagnosis of dementia. *Ann Nucl Med* 2002;16:515–25.
6. Herholz K, Salmon E, Perani D, Baron JC, Holthoff V, Frolich L, et al. Discrimination between Alzheimer dementia and controls by automated analysis of multicenter FDG PET. *Neuroimage* 2002;17:302–16.

7. Herholz K. PET studies in dementia. *Ann Nucl Med* 2003;17: 79-89.
8. Hirao K, Ohnishi T, Hirata Y, Yamashita F, Mori T, Moriguchi Y, et al. The prediction of rapid conversion to Alzheimer's disease in mild cognitive impairment using regional cerebral blood flow SPECT. *Neuroimage* 2005; 28:1014-21.
9. Minoshima S, Frey KA, Koeppe RA, Foster NL, Kuhl DE. A diagnostic approach in Alzheimer's disease using three-dimensional stereotactic surface projections of fluorine-18-FDG PET. *J Nucl Med* 1995;36:1238-48.
10. Burdette JH, Minoshima S, Vander Borgh T, Tran DD, Kuhl DE. Alzheimer disease: improved visual interpretation of PET images by using three-dimensional stereotactic surface projections. *Radiology* 1996;198:837-43.
11. Honda N, Machida K, Matsumoto T, Matsuda H, Imabayashi E, Hashimoto J, et al. Three-dimensional stereotactic surface projection of brain perfusion SPECT improves diagnosis of Alzheimer's disease. *Ann Nucl Med* 2003;17: 641-8.
12. Tang BN, Minoshima S, George J, Robert A, Swine C, Laloux P, et al. Diagnosis of suspected Alzheimer's disease is improved by automated analysis of regional cerebral blood flow. *Eur J Nucl Med Mol Imaging* 2004;31:1487-94.
13. Kaneko K, Kuwabara Y, Sasaki M, Ogomori K, Ichimiya A, Koga H, et al. Posterior cingulate hypoperfusion in Alzheimer's disease, senile dementia of Alzheimer type, and other dementias evaluated by three-dimensional stereotactic surface projections using Tc-99m HMPAO SPECT. *Clin Nucl Med* 2004;29:362-6.
14. Imabayashi E, Matsuda H, Asada T, Ohnishi T, Sakamoto S, Nakano S, et al. Superiority of 3-dimensional stereotactic surface projection analysis over visual inspection in discrimination of patients with very early Alzheimer's disease from controls using brain perfusion SPECT. *J Nucl Med* 2004; 45:1450-7.
15. Kemp PM, Hoffmann SA, Holmes C, Bolt L, Ward T, Holmes RB, et al. The contribution of statistical parametric mapping in the assessment of precuneal and medial temporal lobe perfusion by <sup>99m</sup>Tc-HMPAO SPECT in mild Alzheimer's and Lewy body dementia. *Nucl Med Commun* 2005;26: 1099-106.
16. Kim EJ, Cho SS, Jeong Y, Park KC, Kang SJ, Kang E, et al. Glucose metabolism in early onset versus late onset Alzheimer's disease: an SPM analysis of 120 patients. *Brain* 2005;128:1790-801.
17. Ishii K, Kono AK, Sasaki H, Miyamoto N, Fukuda T, Sakamoto S, et al. Fully automatic diagnostic system for early- and late-onset mild Alzheimer's disease using FDG PET and 3D-SSP. *Eur J Nucl Med Mol Imaging* 2006;10:1-9.
18. Iida H, Itoh H, Nakazawa M, Hatazawa J, Nishimura H, Onishi Y, et al. Quantitative mapping of regional cerebral blood flow using iodine-123-IMP and SPECT. *J Nucl Med* 1994;35:2019-30.
19. Fleiss J. The measurement of interrater agreement. In: *Statistical methods for rates and proportions*. 2nd edition. New York: Wiley; 1981. p. 212-36.
20. Dougall NJ, Bruggink S, Ebmeier KP. Systematic review of the diagnostic accuracy of <sup>99m</sup>Tc-HMPAO-SPECT in dementia. *Am J Geriatr Psychiatry* 2004;12:554-70.
21. Huang C, Wahlund LO, Almkvist O, Elehu D, Svensson L, Jonsson T, et al. Voxel- and VOI-based analysis of SPECT CBF in relation to clinical and psychological heterogeneity of mild cognitive impairment. *Neuroimage* 2003;19:1137-44.
22. Mielke R, Pietrzyk U, Jacobs A, Fink GR, Ichimiya A, Kessler J, et al. HMPAO SPET and FDG PET in Alzheimer's disease and vascular dementia: comparison of perfusion and metabolic pattern. *Eur J Nucl Med* 1994;21:1052-60.
23. Messa C, Perani D, Lucignani G, Zenorini A, Zito F, Rizzo G, et al. High-resolution technetium-99m-HMPAO SPECT in patients with probable Alzheimer's disease: comparison with fluorine-18-FDG PET. *J Nucl Med* 1994;35:210-6.
24. Silverman DH. Brain <sup>18</sup>F-FDG PET in the diagnosis of neurodegenerative dementias: comparison with perfusion SPECT and with clinical evaluations lacking nuclear imaging. *J Nucl Med* 2004;45:594-607.
25. Herholz K, Schopphoff H, Schmidt M, Mielke R, Eschner W, Scheidhauer K, et al. Direct comparison of spatially normalized PET and SPECT scans in Alzheimer's disease. *J Nucl Med* 2002;43:21-6.
26. Sakamoto S, Ishii K. Low cerebral glucose extraction rates in the human medial temporal cortex and cerebellum. *J Neurol Sci* 2000;172:41-8.
27. Dobernt N, Pantel J, Frolich L, Hamscho N, Menzel C, Grunwald F. Diagnostic value of FDG-PET and HMPAO-SPET in patients with mild dementia and mild cognitive impairment: metabolic index and perfusion index. *Dement Geriatr Cogn Disord* 2005;20:63-70.
28. Zhou XH, Obuchowski NA, McClish DK. *Statistical methods in diagnostic medicine*. New York: Wiley; 2002.

# MRIによるAlzheimer病の診断

—画像統計解析手法を用いて

Diagnosis of Alzheimer's disease using statistical analysis of MRI



松田 博史

Hiroshi MATSUDA

埼玉医科大学病院核医学科

◎Alzheimer 病における MRI の役割は、従来は脳腫瘍などの治療可能な脳疾患を除外するためであった。しかし、最近の画像統計解析手法の発展とともに、MRI は Alzheimer 病の早期診断ばかりでなく、疾患の進行および治療効果をモニタリングする可能性も認められるようになってきた。MRI を用いた容積測定のために画像統計解析手法でよく用いられるものは、灰白質濃度を評価できる voxel-based morphometry (VBM) である。さらに、萎縮率を画像化できる tensor-based morphometry (TBM) も注目されている。

**Key word** : Alzheimer病, MRI, SPM, voxel-based morphometry (VBM), tensor-based morphometry (TBM)

先進国では平均寿命が延びるとともに、Alzheimer 病発症率が上昇し、その社会経済的影響が高まっている。Alzheimer 病の病理学的発生機序に関する知識がここ数年蓄積され、Alzheimer 病における神経細胞死の遅延または予防を目的とする特定の治療戦略開発が可能になってきた。しかし、このためには下記の事柄が必要となる。すなわち、①非 Alzheimer 病型認知症では Alzheimer 病に特異的な治療は効果がみられないため、精度の高い Alzheimer 病診断が行えること、②治療効果を高めるため、ごく早期の段階で Alzheimer 病が診断できること、③治療効果のモニタリングを行うため、信頼性が高く、患者にとって意義のある診断法を確立することである。

現在、上記要項のすべてを満たす理想的なバイオマーカーは存在しないが、神経画像診断および脳脊髄液/血液バイオマーカーを組み合わせれば重要な補足情報が得られ、より正確かつ迅速な Alzheimer 病診断に役立つというエビデンスが得られるようになってきた。近年、アメリカを中心に、Alzheimer 病神経画像先導研究 (Alzheimer's Disease Neuroimaging Initiative: ADNI; <http://adni-info.org/>) とよばれる大規模臨床観察研究が

はじめられており、上記バイオマーカーをどのように組み合わせれば Alzheimer 病を診断し治療効果をモニタリングするうえでもっとも強力となるかについて検討されている。

Alzheimer 病の進行または治療効果をモニタリングするため、通常は ADAS-Cog などの神経心理検査が用いられている。しかし、神経心理検査はたしかに機能障害など疾患進行の重要な側面を反映するものではあるが、数々の限界もある。最大の限界は試験・再試験の信頼度が比較的 low、患者の気分、その他の疾患あるいはおもな生活上の出来事などの存在、その他の治療の副作用、試験反復による学習効果など、認知機能にその他の要因の影響が反映されるおそれが高いことである。神経心理検査の信頼性が比較的 low ことから、臨床試験での統計学的検出力が低下し、各患者の認知機能変化の解釈が複雑化する。神経心理検査の別の限界は、Alzheimer 病の神経変性過程に直接作用する治療と認知機能を強化するが、神経変性過程に作用するものではない対症的な治療とを識別できないことである。加えて治療により、Alzheimer 病の進行は緩和されても認知機能への効果は遅れて出現するため、無効と誤認されるおそ

れもある。

神経画像診断のなかでも MRI はおもに構造的画像診断である。構造的画像診断は Alzheimer 病の臨床症状に近似する脳腫瘍や慢性硬膜下血腫などの治療可能な脳疾患を除外するために用いられてきた。しかし、ここ数年、Alzheimer 病の臨床診断の精度改善ばかりでなく、疾患の進行および治療効果をモニタリングする可能性も広く認められるようになってきた。治験成績を示すマーカーとしての適性について検出力分析を実施したところ、構造的画像診断は、現在実施されている認知機能および臨床成績の測定よりもかなり小規模の患者集団で、かつ短期の観察期間で可能となることがわかった。Jack ら<sup>1)</sup>は、1 年間の MRI 試験において 90% の検出力で 25% の治療効果を検出するには各群に 69 名の患者が必要であるとした。それに比べ、神経心理検査を用いて 90% の検出力で 25% の治療効果を検出するには最低 1,277 名の患者が必要である。

MRI による構造的画像診断においては、従来は視覚評価や恣意性の高い関心領域の設定法に頼っていたが、現在では画像統計解析手法によるコンピュータ診断法が普及しつつある。本稿ではこのコンピュータ診断法を中心に述べる。

### 画像統計解析手法の発展

視覚評価による MRI の萎縮評価は定量性に劣る。このため、手動による関心領域の設定が、高

サイド  
メモ

#### Statistical parametric mapping (SPM)

Wellcome Department of Imaging Neuroscience, University College London の Friston らにより開発されている画像統計解析ソフトウェア。フリーウェアであり、<http://www.fil.ion.ucl.ac.uk/spm/> からダウンロードできる。Windows, Linux, Unix, MacOS で動作する Matlab 上のスクリプトである。1990 年代より開発が進んでおり、現在の最新バージョンは 2005 年版。MRI, PET, SPECT のほか、機能的 MRI や脳波、脳磁図の解析も可能である。全世界の神経画像研究者が用いている定番ソフトウェアである。VSRAD では SPM 2002 年版の一部を許可のもとに使用している。

速三次元収集法で得られた MRI において盛んに行われてきた。しかし、この関心領域設定法は数十スライスに施行しなければならず長時間を要するため、ルーチン検査とはなりえなかった。最近では MRI 画像の信号強度の不均一性を補正したうえで、脳全体から灰白質、白質、脳脊髄液成分を自動的に分けることが、statistical parametric mapping (SPM; 「サイドメモ」参照)<sup>2)</sup>などを用いてできるようになった。

このように分けられた灰白質成分の画像に対して平滑化を行うことで、灰白質濃度を PET や SPECT と同様の画像とし、さらに線形変換と非線形変換を組み合わせて Talairach の標準脳に形態変換したうえで、仮説に基づくことなく全脳領域の画像のボクセル単位での画像統計解析を行う voxel-based morphometry (VBM)<sup>3)</sup>が盛んとなってきている。平滑化操作は脳機能局在の個人差をより少なくするとともに、信号対雑音比を向上させ、さらに画像の計数率分布を統計処理するために正規分布に近づける意味がある。この VBM 手法は、従来の関心領域設定による容積測定よりも精度が高いと報告されている<sup>4,5)</sup>。

著者らは、この VBM を簡便に行うことができるソフトウェア (voxel-based specific regional analysis system for Alzheimer's disease: VSRAD) を開発した<sup>6)</sup> (図 1~3)。この VBM による解析法の欠点は、MRI 装置の磁場の不均一性が、得られる画像の信号強度の不均一性につながり、正確に灰白質濃度を評価できないおそれがあることである。このことは異なる MRI 装置を用いる際に誤差を生むことになり、正常 MRI データベースの共有化が困難になる。この信号強度の不均一性をより正確に補正する N3 とよばれるソフトウェアが開発されている<sup>7)</sup>。

画像統計解析手法の欠点としては、脳容積の絶対値を計測することができないこと、脳萎縮が強い場合に解剖学的標準化が不十分のため萎縮部位の位置情報が不正確になるおそれがあることなどがあげられる。後者の欠点に対しては MRI において標準脳のテンプレートに完璧に合致するように非線形の精密な形態変換を行い、その縮小と拡大率を示す変換パラメータ画像を用いて診断する方

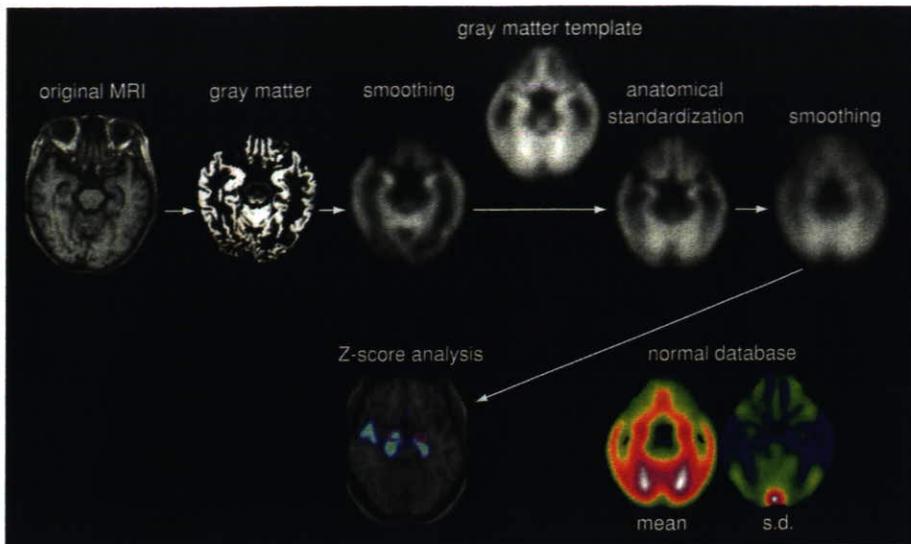


図1 Voxel-based morphometry (VBM) の概念

元の MRI (original MRI) から灰白質画像 (gray matter) を抽出し、1 回目の平滑化 (smoothing) を行う。その後、灰白質テンプレート (gray matter template) を用いて解剖学的標準化 (anatomical standardization) を行う。さらに、2 回目の平滑化 (smoothing) を行い、同様に処理した健常人の正常画像データベースの平均 (mean) と標準偏差 (s.d.) 画像を用いて、ボクセルごとに正常平均からの隔たりを示す Z スコアマップを表示する (Z-score analysis)。

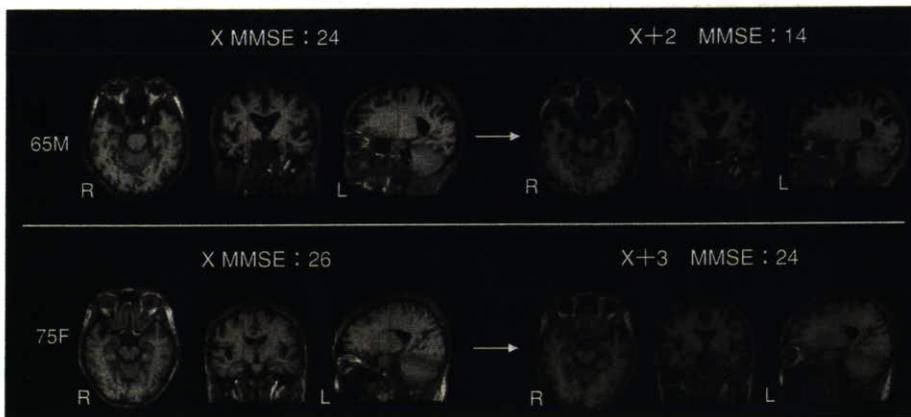


図2 軽度認知障害の段階の Alzheimer 病患者の縦断的 MRI

上段は MMSE 24 点の 65 歳男性、1 年後に Alzheimer 病に移行し、2 年後には MMSE 14 点に低下。内側側頭部をはじめ、脳が最初から全体的に萎縮し、萎縮が進行している。下段は MMSE 26 点の 75 歳女性、5 年後に Alzheimer 病に移行したが、移行していない 3 年後では MMSE 24 点。内側側頭部をはじめ、脳萎縮は軽度であり、萎縮の進行も遅い。

法 (tensor-based morphometry: TBM) が発表されている<sup>8)</sup>(図 4, 5)。TBM は信号強度の不均一性に対しては影響を受けにくい、画像の幾何学的歪みには大きく影響を受ける。この画像歪みに対しても補正法が発表されている<sup>9)</sup>。

### Alzheimer 病における MRI の画像統計解析結果

Alzheimer 病では早期に側頭葉内側部が侵され、その後、大脳皮質に進展することが病理学的研究で明らかとなっている。嗅内野皮質は海馬傍回前方の主要部分をなしている。嗅内野皮質は移行型

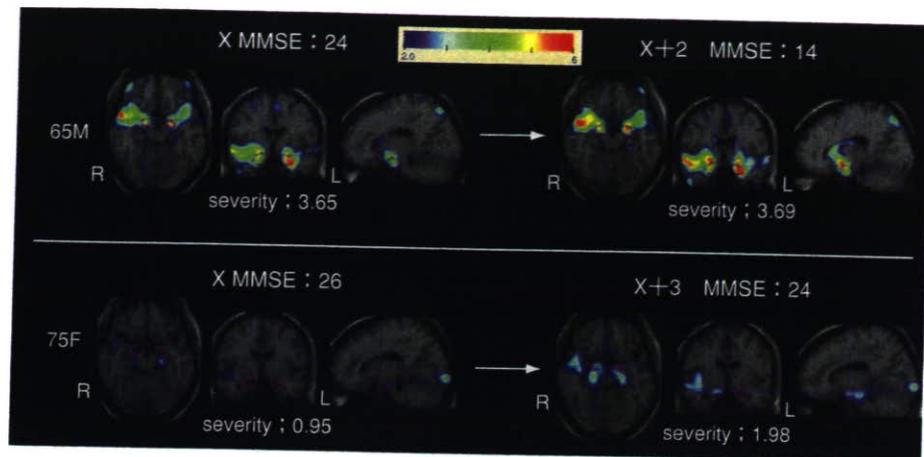


図3 VSRADを用いた軽度認知障害の段階のAlzheimer病患者の縦断的VBM解析  
 図2と同一症例。上段のAlzheimer病急速移行例では、最初から海馬傍回を中心とする内側側頭部に高度の萎縮を示すZスコア(萎縮度で3.65)がみられ、2年後もほぼ同じ値(萎縮度で3.69)を示す。下段のAlzheimer病緩徐移行例では、最初は海馬傍回を中心とする内側側頭部の萎縮は軽度(萎縮度で0.95)であるが、3年後には萎縮が進行している(萎縮度で1.98)。

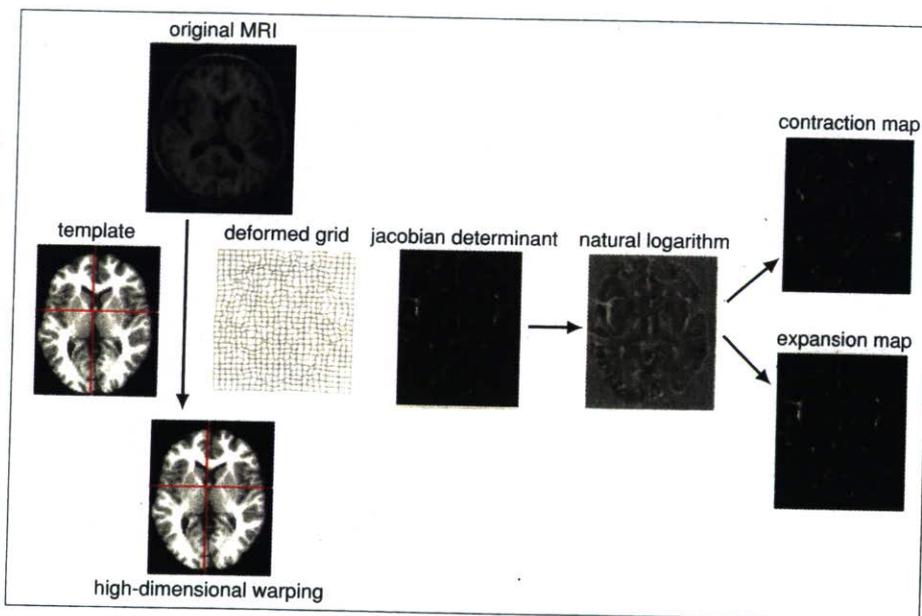


図4 Tensor-based morphometry(TBM)の概念  
 元のMRI(original MRI)をtemplateに完璧に合うように精密に非線形の形態変換を行う(high-dimensional warping)。その形態変換パラメータ画像(deformed grid)から作製されたjacobian determinantを用いて統計解析を行う。具体的には、形態変換パラメータ画像の自然対数(natural logarithm)を取り元の画像の縮小を補正したパラメータ画像(contraction map)と拡大を補正したパラメータ画像(expansion map)に分けて解析する。

の6層構造を示し、第II層はAlzheimer病ではもっとも早期に障害され、軽度認知障害の段階においてすでに40%程度の神経細胞数の減少がみられるとされている<sup>10,11)</sup>。MRIによるVBMでは

この部位の萎縮が軽度認知障害の段階ですでに検出されている<sup>6,12-15)</sup>(図3)。軽度認知障害患者は、Alzheimer病患者と比較して頭頂葉皮質と後帯状回の萎縮が乏しいとされている<sup>14)</sup>。また、Alz-

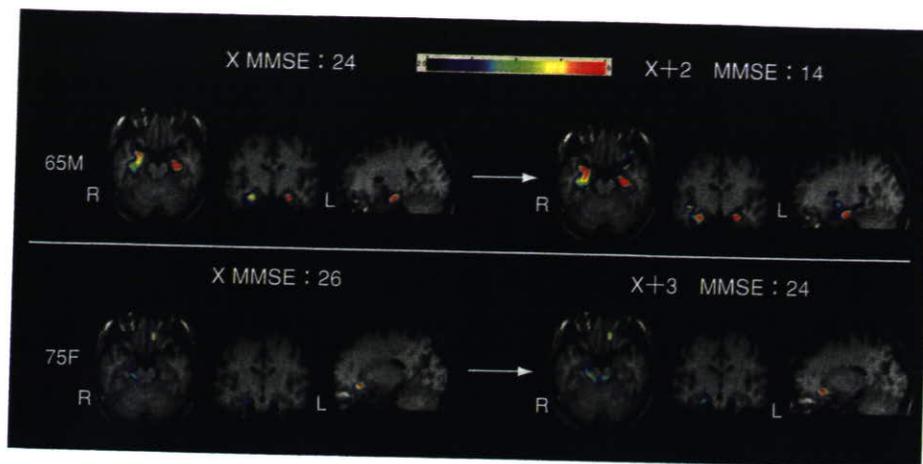


図5 軽度認知障害の段階の Alzheimer 病患者の縦断的 TBM 解析

図2と同一症例。上段の Alzheimer 病急速移行例では、最初から海馬傍回を中心とする内側側頭部に高度の萎縮がみられ、2年後、萎縮の進行がみられる。下段の Alzheimer 病緩徐移行例では、最初は海馬傍回の萎縮が右側にみられ、3年後には萎縮が進行している。VBM 解析と同様の結果であるが、TBM 解析のほうが萎縮部位が限局しており、解剖学的位置精度が高い。

Alzheimer 病患者は内側側頭部のみではなく、側頭葉新皮質にも萎縮が広がっているとされる<sup>14)</sup>。Alzheimer 病の若年発症例は高齢発症例に比べ、内側側頭部の萎縮はあまりめだたず、側頭・頭頂葉皮質に萎縮がめだつ傾向がある<sup>16)</sup>。

Alzheimer 病の前駆期とされている健忘型の軽度認知障害を有する患者は、年 10~15%の割合で Alzheimer 病に移行するといわれている。しかし、すべての症例が Alzheimer 病に移行するとは限らず、さらに経過観察中に軽度認知障害の診断基準を満たさなくなることもあることが指摘されている。この軽度認知障害患者の予後に関して、Alzheimer 病に移行するかどうか、また、その移行が早い遅いを判定することは、患者の治療のみならず、介護者にとっても重要である。MRI によるこの判定に関して、すでにいくつかの報告がみられる。海馬よりも嗅内野皮質の萎縮が軽度認知障害の段階でめだつ例は Alzheimer 病への早期の移行が危惧される<sup>17)</sup>。VBM を用いた検討では、軽度認知障害から Alzheimer 病への移行例は非移行例に比べ、軽度認知障害の段階において前頭葉皮質や内側側頭部、後帯状回~楔前部の萎縮が強いと報告されている<sup>18)</sup>。縦断的検討において、移行例は非移行例よりも海馬の萎縮の進行が早いとされている<sup>19)</sup>。

## おわりに

このように Alzheimer 病の MRI を用いた画像診断では画像統計解析手法の応用は必須となっている。VBM や TBM のほかにも、大脳皮質の厚さを計測する手法<sup>20)</sup>などが発表されている。前述の ADNI 研究では数種類の解析法が検討されており、いずれの解析法がもっとも有用であるかが判明するものと思われる。安価で施行が容易な MRI は、今後ますます Alzheimer 病の診断に用いられていくものと考えられる。

## 文献

- 1) Jack, C. R. Jr. et al. : *Neurology*, **62** : 591-600, 2004.
- 2) Friston, K. J. : *Human Brain Function*, 1st ed. (ed. by Frackowiak, R. S. J. et al.). Academic Press, San Diego, 1997, pp.25-41.
- 3) Ashburner, J. and Friston, K. J. : *Neuroimage*, **11** : 805-821, 2000.
- 4) Good, C. D. et al. : *Neuroimage*, **17** : 29-46, 2002.
- 5) Testa, C. et al. : *J. Magn. Reson. Imaging*, **19** : 274-282, 2004.
- 6) Hirata, Y. et al. : *Neurosci. Lett.*, **382** : 269-274, 2005.
- 7) Sled, J. G. et al. : *IEEE Trans. Med. Imaging*, **17** : 87-97, 1998.
- 8) Ohnishi, T. et al. : *Brain*, **129** : 399-410, 2006.
- 9) Wang, D. et al. : *Magn. Reson. Imaging*, **22** : 1223-1232, 2004.
- 10) Gomez-Isla, T. et al. : *J. Neurosci.*, **16** : 4491-500, 1996.
- 11) Petersen, R. C. et al. : *Arch. Neurol.*, **63** : 665-672,

- 2006.
- 12) Baron, J. C. et al. : *Neuroimage*, **14** : 298-309, 2001.
  - 13) Pennanen, C. et al. : *J. Neurol. Neurosurg. Psychiatry*, **76** : 11-14, 2005.
  - 14) Chetelat, G. et al. : *Neuroimage*, **27** : 934-946, 2005.
  - 15) Karas, G. B. et al. : *Neuroimage*, **23** : 708-716, 2004.
  - 16) Frisoni, G. B. et al. : *J. Neurol. Neurosurg. Psychiatry*, **76** : 112-114, 2005.
  - 17) Killiany, R. J. et al. : *Neurology*, **58** : 1188-1196, 2002.
  - 18) Bozzali, M. et al. : *Neurology*, **67** : 453-460, 2006.
  - 19) Jack, C. R. et al. : *Neurology*, **55** : 484-489, 2000.
  - 20) Singh, V. et al. : *Brain*, **129** : 2885-2893, 2006.

\* \* \*

# ADNI と神経心理評価

## ADNI and neuropsychological battery

新潟リハビリテーション大学院大学教授

Morihiro Sugishita 杉下守弘

### Summary

21世紀を迎え、認知症患者が急増し、認知症患者と健常高齢者の判別が大きな問題になっている。このため、米国ではアルツハイマー病に関する長期縦断的プロジェクト(ADNI)が始まっており、本邦でも同じ趣旨のプロジェクト(J-ADNI)が始まる。

米国のADNIでどのような神経心理検査が用いられているか、それらは、軽度認知障害(MCI)、認知症、健常高齢者を定義するためにどのように使われているかを解説した。さらに、MCIの下位分類であるAmnesicやNon Amnesicの判別が神経心理検査とどのように関連しているかも論じた。また、これらの神経心理検査を日本で使用する際の問題点に言及し、その例としてMMSEを取り上げた。

おわりに、2007年6月に行われた「アルツハイマー病の神経心理検査」に関する国際会議の内容を紹介した。アルツハイマー病の初期、記憶障害が軽度のときの変化を捉えられるテストが求められていること、テストとしてADAS-COGが有用であること、Amnesic MCIでは記憶の次に言語が障害されることが多いことなどが論ぜられた。

### Key words

- ADNI
- 神経心理検査
- MMSE
- 認知症
- 記憶 ●MCI

### はじめに

人間の精神能力を測る試みに大きな進歩があったのは、20世紀初頭、知的障害児の測定を契機にしたものであった。20世紀になると、小学校に児童が進学するようになった。そして、小学校で学校教育についていけない児童(知的障害児)と、学校教育についていける児童(健常児童)を選別することに重大な関心が寄せられるようになった。フランスの心理学者ビネに、小学校に入学しても学業に耐えられない知的障害児と、学業に耐えられる児童を選別するテストを研究するようにパリ市の教育委員会から依頼があった。この研究の結果、ビネの知能検査ができあがり、その後さまざまな知能検査が作成され、この方面の発展につながった。

21世紀になり、先進国では高齢者の数が急増した。それに伴い、認知症患者数が増大し、いかに認知症を予防し、治療するかに重大な関心がもたれるようになった。認知症患者と健常高齢者の判別の問題は知的障害児と健常児童の判別問題と同様に、人間の精神能力を測る試みにさらなる発展をもたらす可能性が大である。

日本では、2007年から米国のADNI(Alzheimer's Disease Neurological Initiative)に倣い、J-ADNI「将来のアルツハイマー病予防・治療薬開発に向けた臨床治験デザイン策定のための多施設共同前向き観察研究」が始まる。このプロジェクトでは、脳画像やバイオマーカーに

重点を置いている。しかし、認知症や軽度認知障害の判定は、臨床症状や神経心理テストなどによってなされるので、認知症患者や健常高齢者の精神能力の測定、すなわち神経心理検査にも配慮がなされているのはいうまでもない。

本稿では、①認知症の神経心理検査について、②ADNIにおける神経心理検査について、③日本でADNIと同じ神経心理検査を行う場合の問題点、④2007年6月に米国ワシントンで行われた国際会議「アルツハイマー病およびMCIのための臨床試験における認知評価」の内容などについて論じたい。

### 認知症の神経心理検査について

ADNIと神経心理検査という本題に入る前に、少し大所高所に立って、認知症の神経心理検査について考えてみたい。認知症の症状を捉えるには、記憶、特にエピソード記憶の障害が初期に現れるのでエピソード記憶を検査することが重要といわれてきた。知能テストでは短期記憶の検査しかないため、知能検査を行うだけでは認知症を捉えるには不十分であった。このため、ウェクスラー記憶検査-改訂版(Wechsler Memory Scale-Revised; WMS-R)などの記憶検査が認知症の検査として用いられてきた。最近では認知症を捉えるには記憶以外に言語、実行機能、視空間などの領域を調べることも必要といわれている<sup>1)</sup>。したがって、言語、実行機能、視空間などの領域を調べる検査も行われるようになった。しかし、現在使用されている認知症のテストにはいくつか共通した問題点がある。

第1はテストの問題数が少なすぎることで、計るべき領域について、それぞれ1問か2問しかない。これでは認知症の症状を捉えるには不十分である。たとえば、Mini-Mental State Examination (MMSE)<sup>2)</sup>を取り上げてみると、記銘で1題、注意/計算で1問で、全体でも30問と非常に少ない。したがって、認知症を捉えるのに十分とはいえない。

認知症のテストで、問題数が多いテストとしては、Alzheimer's Disease Assessment Scale-Cognitive (ADAS-

COG)<sup>3)</sup>がある。しかし、このテストが認知症を測定するのに十分かどうかは今後の課題であろう。

第2の問題は、記憶のテスト自体の問題である。記憶のテストのほとんどは、多大の意志というか、努力というかそういうものを要求している。たとえば、ADAS-COGを例にとると、記憶を測るために10語からなるリストを患者に音読させ、覚えてもらう。そして、努力すれば、どれだけたくさん覚えられるかをみるために3回繰り返すテストを行っている。この種のテストは、記憶を測定していると同時に努力あるいは意志の強さを測っているのもであって、記憶自体を測っているとは思えないのである。言い換えれば、あまり努力を要さない記憶のテストを作製することが望ましいといえよう。Petersen<sup>4)</sup>は記憶のテストとして、WMS-Rの論理的記憶<sup>5)</sup>を推奨している。このテストは、患者が短い話を聞き、聞き終わった後で、どのような内容の話であったか思い出して話すというものである。また、一度思い出して話した30分後にもう一度思い出して話すことも求められる。このテストは10語のリストを覚えるように3回求められるテストよりは自然で、努力を要しない。しかし、今後は「さらに努力を要しない記憶」のテストを作製すべきであろう。

以上に述べたように、認知症の神経心理テストには問題があり、よりよいテストの作製は緊急な課題である。

早期AD研究から神経心理検査は脳画像測定法よりバラツキが大きいという結果が出ている。この結果は神経心理評価が役に立たない証拠として受け取られがちである。しかし、それは現存の認知症の神経心理検査レベルが低いためと思われる。

### ADNIにおける神経心理検査について

ADNIでは、使用されている神経心理評価は15ある(表1)。このうち、健常者、軽度認知障害(mild cognitive impairment; MCI)、および認知症の定義のために、MMSE、Clinical Dementia Rating (CDR)、Logical Memory II および Geriatric Depression Scale (GDS)、Func-

表1 ADNIで使用されている15種の神経心理評価

1. Mini-Mental State Examination [MMSE] (Folstein et al, 1975)<sup>2)</sup>
2. Alzheimer's Disease Assessment Scale-Cognitive [ADAS-COG] (Rosen et al, 1984)<sup>3)</sup>
3. Logical Memory II [Delayed Paragraph Recall] (Wechsler, 1987)<sup>4)</sup>
4. Boston Naming Test (Kaplan, Goodglass, Weiraub, 1983)<sup>5)</sup>
5. Category Fluency (Butters et al, 1987)<sup>6)</sup>
6. Clock Drawing Test (Goodglass & Kaplan, 1983)<sup>7)</sup>
7. Digit Span (Wechsler, 1981)<sup>8)</sup>
8. American National Adult Reading Test [ANART] (Nelson & O'Connell, 1978)<sup>9)</sup>
9. Rey Auditory Learning Test (Rey, 1964)<sup>10)</sup>
10. Trail Making Test : Parts A and B (Reitan, 1958)<sup>11)</sup>
11. Digit Symbol Substitution Test (Wechsler, 1981)
12. Clinical Dementia Rating [CDR] (Berg, 1988)<sup>12)</sup> (Morris, 1993)<sup>13)</sup>
13. Functional Activities Questionnaire [FAQ] (Pfeffer et al, 1982)<sup>14)</sup>
14. Neuropsychiatric Inventory Q [NPI-Q] (Kaufers et al, 2000)<sup>15)</sup>
15. Geriatric Depression Scale [GDS] (Sheikh & Yesavage, 1986)<sup>16)</sup>

tional Activities Questionnaire (FAQ)の5つが使用される。たとえば、MCIの定義は次のようになる。本人や近親者が認知機能の低下を訴えたことがあり、MMSEが24~30、CDRが0.5、FAQで本質的に正常な機能的活動であれば、MCIと考えている。MCIをさらに分類するのに Logical Memory IIが使用される。Logical Memory IIの cutoff scoreより低く、記憶障害があることがはっきりすれば、Amnesic MCIであり、cutoff scoreより高く記憶障害がないと考えられるなら、Non Amnesic MCIが考えられる。Cutoff scoreは米国のADNIでは決まっているが、日本では現在、検討中である。

認知症の検査で、記憶以外に問題になる領域として、言語、注意/実行機能、視空間の3領域がある。これらの評価は次の6つの検査で行われる。言語は Boston Naming Testと Category Fluencyの2つの検査で検討する。注意/実行機能は Trail Making Test : Parts A and Bと Digit Symbol Substitutionの2つの検査で、視空間は Clock Drawing Testおよび補助として ADAS-COGを用いて検討する。これらのテストについては障害ありとなしを決める cutoff scoreが米国でもわが国でも決められておらず、今後の課題である。



## 日本で ADNI と同じ神経心理検査を行う場合の問題点

日本で ADNI と同様のプロジェクトを行う場合、ADNI が行われた米国で使用したテストと同じものを使用することになる。使用するといっても英語で書かれているテストを使うわけにはいかないので、日本語でできた、できれば日本の実情に合わせた、いわゆる日本語版を使用することとなる。ところが、日本には現在、日本語版のないテストがある。Boston Naming Test や Rey Auditory Learning がそうである。これらは作製しなければならないであろう。

第2の問題は、いくつも日本語版があるテストの場合である。たとえば、MMSE や CDR がそうである。このような場合どのテストを使うべきか検討しなければならない。

第3の問題は、日本語版に不適切な部分が含まれている場合である。たとえば、文を英語から日本語に訳した場合、注意しないと、英語文より日本語文のほうが長くなる。この文を記憶テストに使う場合、長いほうが記憶するのが難しくなるという結果を引き起こし日本語版は英語版より難しいことになる。

ここで、第2、第3の問題に関連する例としてMMSEを取り上げてみる。

MMSEにはいくつかの日本語版がある。その中で、はじめに作られたのが森版である。その後作製され、現在、最も広く用いられているのが北村版である。

北村版は最も広く用いられているだけあってよくできている。森版については、北村版をはじめとする他の版と異なっていて問題があると考えている人がいるが、よくできたテストである。たとえば森版は、見当識のところでは何曜日か尋ねる項目がない。北村版やその他の版では曜日を尋ねる項目がある。しかし原版をみると、曜日を尋ねる項目はないのである。また、森版は冒頭で「今日は何曜日ですか」と聞いている。一方、北村版は「今年は何年ですか」と冒頭で聞いている。原版をみると、はじめに日付を聞いて、答えられなかった部分を尋ねていくテストであることがわかる。森版がおかしいわけではない。

北村版では、「ここは何地方ですか」という項目がある。一方、森版にはこのような項目はない。原版をみると、「state(州, 国)」があり、これを北村版は「地方」に置き換えたようである。カナダ版ではstateを国(カナダ)に置き換えており、日本語版でもstateを国(日本)に置き換えて、「ここは何という国ですか」という項目にするほうが適切であろう。

言語のところには、3段階の命令というのがある。森版では「大きな方の紙を取り、半分に折って、床に置く」となっている。原版には「大きな方の」という部分はないので問題ではある。しかし、北村版のように「右手にこの紙を持ってきてください。それを半分に折りたたんでください。机の上に置いてください。」はあまりにも長すぎる。記憶のテスト材料は長いと難しくなる。英語では「Take a paper in your right hand, fold it in half and put it on the floor」であり、短い命令文である。森版のように短くないと検査が難しくなってしまう。

両方の版について問題であるのは、復唱に使う「No ifs, ands or buts」という文である。この文を森版は「ちりも積もれば山となる」としており、北村版は「みんなで力を合わせて綱を引きます」としている。本来は「も

し、そして、しかしはだめ」とか「だけど、やっぱり、でもはだめ」というような訳のほうがよいであろう。

このように日本語のどれをJ-ADNIで使用するのかが決めるのはかなり困難が伴う。

できあがっているMMSE日本語訳について検討し、どこに問題があるか指摘することは容易である。しかし、自分で作成するとなると非常に難しい。表2に、筆者のMMSEの試案を作ってみただけでも、なかなかよいものを作ることは難しい。この試案では、曜日を尋ねる項目や地方を聞く項目など誤答が多い項目がない。また、3段階命令も短くなっている。したがって、人によっては北村版より2、3点得点が低くなる可能性がある。総得点が30点であるから、この影響は大きい。MCIはMMSEの得点が24~30でなければならない。北村版では本来MCIのものが認知症(20~26)に入ってしまう場合がでてくる。ここにMMSEを例として挙げたが、J-ADNIで他に14のテストを使用するわけで、これらをどうするかは大きな課題である。

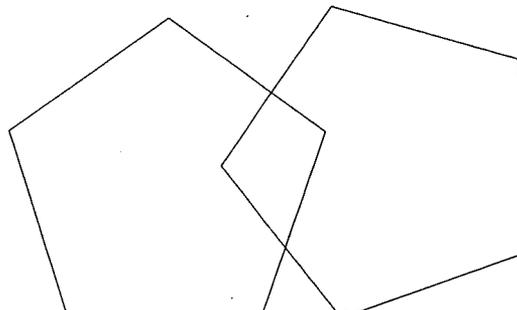
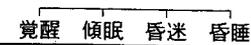
### 国際会議「アルツハイマー病およびMCIのための臨床試験における認知評価」について

2007年6月19日、ワシントンD.C.で「アルツハイマー病およびMCIのための臨床試験における認知評価」という国際会議が開かれた。「ADNIと神経心理検査」についての最新の情報13題が提供されたので、この会議の内容を紹介する。

13題の演題のうち4題(3, 5, 11, 12)が製薬会社の関係者の演題であった。残りの9題のうち、1題(13)が政府機関FDA(米国食品医薬品局)の演題で、他の8題が大学からの発表である。大学や臨床の場からの発表が主流とはいえ、製薬会社の発表が増加しているのがみてとれる。

参加者も39名中十数名が製薬会社関係者であり、FDAのDr. Katzが製薬会社のデータも十分考慮するように述べるなど、この領域での製薬会社の学問的貢献が増大しているのがわかる。

表2 Mini-Mental State Examination (MMSE)

|                                   |  |   |
|-----------------------------------|--|---|
|                                   |  | 患者名: _____  |
|                                   |  | 検査者: _____  |
|                                   |  | 日付: _____   |
| 見当識                               |  |   |
| 5 ( ) 点                           | ①「今日の日付をご存知ですか」<br>②「何年かわかりますか」<br>③「季節は何かわかりますか」<br>④「何月かわかりますか」<br>⑤「何日かわかりますか」  |   |
| 5 ( ) 点                           | ①「病院の名前をご存知ですか」<br>②「ここは何階ですか」<br>③「ここは何市(区)ですか」<br>④「ここは何県ですか」<br>⑤「ここは何という国ですか」  |   |
| 記 録                               |  |   |
| 3 ( ) 点                           | 3物品の命名: 記憶を検査してよいか患者様に尋ねる(「記憶を検査してよろしいですか」)。それから3つの無関係な物品の名前(たとえば、「ボール、海、草」)を1つ1秒でいう。検査者が3つの名前をいった後で、3つの名前すべてをいうように患者様に求める(「今いった3つの名前を全部いって下さい」)。各々の正答に1点を与える。被験者が3つすべてを憶えるまで繰り返す(6回まで)。繰り返した回数を数え、記録する。 _____ 回   |   |
| 注意と計算                             |  |   |
| 5 ( ) 点                           | 連続7(100から7を引き算していくように患者様に求める。5回引算をしたら止める[93, 86, 79, 72, 65])。 (「100から7を引き算してください」、「また、7を引いてください」) 各々の正答に1点を与える。(患者がこの課題をできなかつたり、しようとしなないなら、代わりに、「世界一(セカイイチ)」を逆唱させる。(「セカイイチ」をうしろからいってください) 得点は正しい順序で答えられた音節の数である(たとえば、チイカセ=5点、チカイセ=3点)。  |   |
| 想 起                               |  |   |
| 3 ( ) 点                           | 前にいった3つの物品の名前をいえるかどうか患者様に尋ねる(「前にいった3つの物の名前をいってください」)。各正答に1点を与える。   |   |
| 言 語                               |  |   |
| 9 ( ) 点                           | <p>呼称: 腕時計をみせて「これは何ですか」と尋ねる。(1点)<br/>鉛筆をみせて「これは何ですか」と尋ねる。(1点)<br/>復唱: 検査に続いて文章を繰り返すように患者様に求める(「文章をいいますから繰り返してください」、「だけど、やっぱり、でも、はだめ」)。患者様には1回しかいわせない。(1点)<br/>3段階命令: 紙を机の上に置いて、命令をいう。「紙を右手に取って、それを半分に折り、床に置いてください」正しく行われた部分1つにつき1点を与える。(3点)<br/>読んで従う: 「目を閉じてください」という文を患者様がはっきりみえるように十分大きな字で印刷する。それを読み、それが意味する内容を実行するように患者様に求める(「これを読んで、この通りにやってください」)。(1点)<br/>文を書く: 患者様に何も書いてない紙を渡し、検査者のために文を1つ書くように求める(「文章を書いてください」)。文をいって、患者様に書き取りさせてはいけない。文は自発的に書けなければならない。主語、動詞を含み、理にかなったものでなければならない。正しい文法と句読法は必要でない。(1点)<br/>図の模写: 一辺が約2.5cmの重複する五角形を紙に書き、それをあるがままに、正確に模写するよう患者様に求める(「この図形を書いてください」)。10すべての辺があり、2つの辺は交差していれば1点を与える。振線や回転は無視される。(1点)</p> |   |
| _____ 総得点                         |    |   |
| 意 識                               |  |   |
| 患者様の意識レベルを左の覚醒から右の昏睡までの連続に沿って評価する |  |   |
|                                   |  |  |

製薬会社では近年専門家を雇用し、自分のところで研究をしていく状況にあることが感ぜられた。

### 1) "Background and meeting objectives" Dr. Dennis Selkoe

冒頭の演題では、会議の背景と目的が述べられた。その中で、エピソード記憶障害に焦点をあてるべきであることが論ぜられた。認知症はエピソード記憶障害が前面にでてくるとというのが一般的であり、常識であった。特にアルツハイマー病(Alzheimer's disease; AD)はエピソード記憶障害が前景にたち、数も多い。ところが、レビー小体型認知症や前頭側頭葉型認知症など、エピソード記憶障害が必ずしも前面に立たない認知症が強調されるようになった。患者数の多いADの症状である記憶障害に焦点をあてることは治験を進めるにあたって重要と思われる。ADNIにおいて、MCIも従来の Amnestic MCIだけでなく、Non Amnestic MCIを考慮せざるをえなくなってきており、これは2003年のヨーロッパの会議や2005年の米国の会議(petersen, 2006)から明らかである。しかし、治験を進めるうえでは Amnestic MCIに絞ることが必要で、ADNIもそれには言及している。

また、ADについては、糖尿病のように、early response と later response を分けて考えて検討すべきであるという。ADの early response として、髄液の生化学的マーカーについての研究が言及された。さらに、ADの初期、記憶障害が軽度の時の変化をはっきりと捉えられるテストは何かという観点が必要と述べた。

### 2) "Designing cognitive tools for clinical trials: Lessons from the past" Dr. Mary Sano

過去の経験から、現在使われているテストのうち MMSE が AD および MCI を捉えるのに有効であることが論ぜられた。

### 3) "The Alzheimer's Disease Assessment Scale: A current perspective on it's use in clinical trials" Dr. Richard Mohs

Mohs らの開発したテスト ADAS-COG についての最近の知見が述べられた。ADAS-COG で18ヵ月間認知症の症状を評価したところ、認知症の重症度が異なる場合、記憶障害進行の程度が異なることを捉えられるという。

また、ADAS-COG 中の記憶テストの得点が悪い症例は言語テストの得点の変化が大きいという。軽度の症例で記憶の得点が一番変化するとは限らないという結果が報告された。

### 4) "Measuring neuropsychological changes in preclinical Alzheimer's disease" Dr. David Salmon

California Verbal Learning Test (CVLT) が MMSE より preclinical AD を検出するのに優れていることが論ぜられた。また、認知非対称(cognitive asymmetry)という概念が言及された。記憶機能が障害されているかどうかではなく、他の機能との対比で、非対称として捉えるほうがよいという考えである。

### 5) "Clinical experience with the neuropsychological changes in preclinical Alzheimer's disease" Dr. Michael Grundman

ADの初期に治療を行ったときの効果と、評価するのに適した有効なテストとして次に述べるテストバッテリーを紹介した。

記憶テストは WMS-R の Visual Paired Associates, WMS-R の Verbal Paired Associates および Rey Auditory Verbal Learning Test の3つである。実行機能のテストとして、WMS-R の Digit Span, Controlled Word Association Test および Category Naming Test (animals) である。これらのテストバッテリーが有効であるというデータが示された。これらのテストバッテリーは ADNI と同じものが3つあるが、他の3つは異なっている。これらのテストバッテリーは側頭葉の機能を測っているのではないかという質問があった。

### 6) "Assessment of cognitive decline in AD and MCI with performance-based composite measures" Dr. Robert Wilson

神経心理テストは同じテストを2回同じ被験者に行えば、2度目の成績は1度目の成績よりよくなることが多い。いわゆる学習効果である。通常、ADへの薬の効果とみる場合は薬剤を投与する前とその後の少なくとも2度は検査をする。したがって、同じテストを薬剤投与前後に用いると、成績がよくなった場合、薬物の効果と学習効果が区別できなくなる。ADの神経心理検査でこ

の学習効果があるのかどうか検討した結果、ADではごくわずかしが学習効果がないと報告された。

また、ADの認知機能の減退は、罹病率と死亡率、特に後者と相関がよいという。

7) "Cognitive assessment in second and prime prevention" Dr. Mary Sano

The Resource Use Inventory (RUI)という経済的利益を含んだ治療効果をみる評定法の使用経験が論ぜられた。9問からなる簡単な評定法である。ADの評定法には簡易なものが多いが、これくらい簡易だと、それなりに意義があると思われる。

8) "Cognitive changes in the MCI spectrum" Dr. Ron Petersen

Amnesic MCIが経過とともに何が障害されてくるかは臨床的に興味のある問題である。この問題につき、Petersenは記憶の次は言語が障害されることが多いという。また、視空間機能は保たれるものが多いと述べていた。

9) "Advantage and disadvantage of computer-based cognitive assessment in clinical trials for AD and MCI" Dr. Peter Snyder

紙と鉛筆を用い、人間が問題を出題して施行する通常のテストではなく、コンピュータが出題する形式のテストの長所と短所がテーマとなった。ADの場合、出題者が何を求めている課題なのか理解しにくいという点が問題であろう。

迷路学習テストがMCIの評価で重要であると述べ、具体的にはGroton Maze Learning Test (GMLT)に言及していた。

10) "Comparison of computerized and paper pencil neuropsychological tests in AD and MCI" Dr. Marilyn Albert

鉛筆と紙によるテストとして、ADAS-COG, MMSE, WMS-R, CVLT, Trail Making Testの5種、コンピュータに基づくテストとして、CANTAB Paired Associates Learning Test (PAL)とCANTAB Graded Naming Testを使用し、両群を比較した結果が発表された。心理テストはなるべくAD患者からよい得点を得られるように

しむけなければならない。その点、人間がテストを行うほうが有利であることは免れないと思われる。

11) "Use of cognitive biomarker in early phase clinical trials" Dr. Julie Chandler & Dr. Lyn Harper Mozley

Cognitive biomarkers (認知バイオマーカー)とは、認知機能に関連のあるバイオマーカーのことで、MRI, PET, 髄液、血液などが含まれる。この方面の研究が論ぜられた。

12) "Distribution and type of neuropathology define the cognitive correlates of progression to AD in Amnesic MCI" Dr. Roger Lane

Amnesic MCIにおいて、ADへ進行する場合に関連する認知要素を限定する神経病理について論ぜられた。

13) "FDA perspective : Cognitive testing in clinical trials for AD and MCI" Dr. Russell Katz

最後の演題は日本の厚生労働省にあたる米国の保健社会福祉局に属するFDAのDr. Katzの演題である。

ADおよびMCIの治験における神経心理テストについて、FDAの立場を述べた。12の疑問を提起する形式で講演が進んだ。

第1に臨床状態と病理的状态を考えたとき、ADやMCIの臨床状態は十分に把握されているのかという問題提起がされた。ADやMCIの患者のperformanceは時の経過とともに減退していく。そこへ活発な治療を行うとperformanceの減退の程度が弱まる。そこで治療を撤回すると、performanceの減退がまた強まるというモデルでADの問題を扱っていきたいという。

第2にADやMCIにはADAS-COGがよいテストであると考えていることが述べられた。それだけでなく、他のテストは受け入れられるのかという疑問が呈された。

第3に、ADAS-COG以外のテストを受け入れるには3つの条件が必要であるとして、①横断的な臨床状態と関連している証拠、②縦断的な臨床状態に関連した変化、③少なくとも6ヶ月の間の縦断的なデータを挙げた。

第4にコンピュータに基づいたテストあるいは紙と鉛筆で行うテストのどちらを用いるのかを考えなければな

らないことが述べられた。

第5に、異なるレベルの症状には異なるテストを用いるべきではないかという問いが出された。全体的に評価できるテストが必要であることや、記憶、視空間といった一領域のテストの結果から何がいえるのか、などが論ぜられた。

第6に、世界共通で使用できる Universal Cognitive Test をつくるのはどうかということが論ぜられた。この呼びかけに応じて、このようなテストを作製するグループが形成され、8月から活動を開始している。ファイザー株式会社の Dr. Holly Posner が窓口になっている。

第7に、AD に効果のある薬物は一般的な機能に関連した変化を生じないのかという問題提起がなされた。

第8、第9はテストにあまり関連がないと思われるので省略する。

第10に「治療効果をみるのに、ランダムな開始あるいは中止」の代わりに、'natural history staggered start design' がよいのではないかという議論がなされた。

第11に、さまざまな関連グループ間のコミュニケーションの増加が論ぜられた。

第12に、治療計画の研修会(workshop)を行う必要が説かれた。例としてパーキンソン病におけるさまざまな変容を示すための実験計画に関する研修会が現在予定されており、類似の研究会をADとMCIについて行うべきであることが論ぜられた。

以上がワシントンD.C.で開かれた国際会議の概要である。最後に述べたように、FDAが「ADとMCIの治療計画における神経心理テスト」に積極的な関心を寄せていることが印象的であった。

## VI おわりに

認知症について、その脳画像、脳機能画像、バイオマーカー、遺伝子、脳病理などは多くの人に関心をもち、また実際に研究が進んでいる。一方、認知症の症状を捉えるための神経心理検査の作成となると、関心をもつ人は少ない。しかし、認知症をバイオマーカー、画像などで定義できない現状では、神経心理検査を認知症研究の基

礎といわざるをえない。今後、ADなどの認知症が、バイオマーカーや画像などで定義できるようになったとき、神経心理検査はその存在意義を失うのであろうか。その場合でも、やはり病状はどのようなものであるかということを経験心理検査で捉える必要はあると思われる。神経心理テストもよいものを作成すれば、病状把握や病状変化の精度が向上するのであるから、少しでもこの方面の研究を推進しようとする意志が必要であろう。

## 文献

- 1) Petersen RC : Mild cognitive impairment ; Current research and clinical implications. *Semin Neurol* 27 : 22-31, 2007
- 2) Folstein MF, Folstein SE, McHugh PR : "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* 12 : 189-198, 1975
- 3) Rosen WG, Mohs RC, Davis KL : A new rating scale for Alzheimer's disease. *Am J Psychiatry* 141 : 1356-1364, 1984
- 4) Wechsler D : WMS-R Wechsler Memory Scale-Revised Manual, New York, The Psychological Corporation, Harcourt Brace Jovanovich Inc, 1987
- 5) Kaplan EF, Goodglass H, Weintraub S : Boston Naming Test. Philadelphia, Lea & Febiger, 1983
- 6) Butters N, Granholm E, Salmon DP, et al : Episodic and semantic memory ; A comparison of amnesic and demented patients. *J Clin Exp Neuropsychol* 9 : 479-497, 1987
- 7) Goodglass H, Kaplan E : The Assessment of Aphasia and Other Neurological Disorders. Philadelphia, Lea & Febiger, 1983
- 8) Wechsler D : Manual for the Wechsler Adult Intelligence Scale (rev. ed.), New York, The Psychological Corporation, Harcourt Brace Janvanovich, 1981
- 9) Nelson HE, O'Connell A : Dementia ; The estimation of premorbid intelligence levels using the New Adult Reading Test. *Cortex* 14 : 234-244, 1978
- 10) Rey A : L'examen clinique en psychologie, Paris, Presses Universitaires de France, 1964
- 11) Reitan RM : Validity of the trail making test as an indicator of organic brain damage. *Percept Mot Skills* 8 : 271-276, 1958
- 12) Berg L : Clinical Dementia Rating(CDR). *Psychopharmacol Bull* 24 : 637-639, 1988
- 13) Morris JC : The Clinical Dementia Rating(CDR) ; Cur-