Visual Versus Fully Automated Analyses of ^{18}F-FDG and Amyloid PET for Prediction of Dementia Due to Alzheimer Disease in Mild Cognitive Impairment

Timo Grimmer[1], Carolin Wutz[1], Panagiotis Alexopoulos[1], Alexander Drzezga[2], Stefan Förster[3], Hans Förstl[1], Oliver Goldhardt[1], Marion Ortner[1], Christian Sorg[1], and Alexander Kurz[1]

[1]Department of Psychiatry and Psychotherapy, Klinikum Rechts der Isar der Technischen Universität München, Munich, Germany; [2]Department of Nuclear Medicine, University Hospital of Cologne, Cologne, Germany; and [3]Department of Nuclear Medicine, Klinikum Rechts der Isar der Technischen Universität München, Munich, Germany

Biomarkers of Alzheimer disease (AD) can be imaged in vivo and can be used for diagnostic and prognostic purposes in people with cognitive decline and dementia. Indicators of amyloid deposition such as ^{11}C-Pittsburgh compound B (^{11}C-PiB) PET are primarily used to identify or rule out brain diseases that are associated with amyloid pathology but have also been deployed to forecast the clinical course. Indicators of neuronal metabolism including ^{18}F-FDG PET demonstrate the localization and severity of neuronal dysfunction and are valuable for differential diagnosis and for predicting the progression from mild cognitive impairment (MCI) to dementia. It is a matter of debate whether to analyze these images visually or using automated techniques. Therefore, we compared the usefulness of both imaging methods and both analyzing strategies to predict dementia due to AD. **Methods:** In MCI participants, a baseline examination, including clinical and imaging assessments, and a clinical follow-up examination after a planned interval of 24 mo were performed. **Results:** Of 28 MCI patients, 9 developed dementia due to AD, 2 developed frontotemporal dementia, and 1 developed moderate dementia of unknown etiology. The positive and negative predictive values and the accuracy of visual and fully automated analyses of ^{11}C-PiB for the prediction of progression to dementia due to AD were 0.50, 1.00, and 0.68, respectively, for the visual and 0.53, 1.00, and 0.71, respectively, for the automated analyses. Positive predictive value, negative predictive value, and accuracy of fully automated analyses of ^{18}F-FDG PET were 0.37, 0.78, and 0.50, respectively. Results of visual analyses were highly variable between raters but were superior to automated analyses. **Conclusion:** Both ^{18}F-FDG and ^{11}C-PiB imaging appear to be of limited use for predicting the progression from MCI to dementia due to AD in short-term follow-up, irrespective of the strategy of analysis. On the other hand, amyloid PET is extremely useful to rule out underlying AD. The findings of the present study favor a fully automated method of analysis for ^{11}C-PiB assessments and a visual analysis by experts for ^{18}F-FDG assessments.

Key Words: mild cognitive impairment; MCI; Alzheimer's disease; AD; positron emission tomography; PET; Pittsburgh compound B; PiB; fluoro-deoxy-D-glucose; FDG

Mild cognitive impairment (MCI) is a syndrome of impaired cognitive function with relatively preserved activities of daily living. It is heterogeneous in terms of etiology, clinical appearance, and prognosis (1). In many cases, MCI is progressive and represents a predementia stage of Alzheimer disease (AD). With regard to the novel treatment strategies that are aimed at slowing disease progression, early detection of AD is of utmost importance. In addition, information on the underlying cause and on the associated outcome is essential for affected individuals and their family members. However, neither the type of pathology nor the individual prognosis can be reliably established on the basis of clinical symptoms alone.

Biomarkers of amyloid pathology and of neurodegeneration have been suggested to improve the identification of underlying AD at the stage of MCI (2). The corresponding imaging biomarkers are amyloid imaging using PET, for example, with the amyloid-specific Pittsburgh compound B (^{11}C-PiB), and neuronal metabolism imaging using ^{18}F-FDG PET. Although amyloid imaging is usually rated dichotomously as positive or negative for AD, in ^{18}F-FDG PET the tracer distribution pattern not only provides information on abnormal brain metabolism but also the pattern of hypometabolism is associated with different underlying pathologies.

There is currently no consensus as to whether amyloid and ^{18}F-FDG PET images should be analyzed visually or via fully automated procedures and which of these methods is better at determining whether a patient with MCI is likely to progress to dementia. Therefore, the aim of this study was to compare the results of 2 methods of analyzing amyloid PET images and ^{18}F-FDG PET images and to compare their usefulness in accurately predicting progression from MCI to dementia due to AD 2 y after the baseline evaluation. The first method of analysis consisted of an assessment by experienced specialists (visual analysis), whereas the second method used cutoff values that were determined using validated computer programs (fully automated analysis).

MATERIALS AND METHODS

Patients

Recruitment and inclusion criteria of the patient sample have been described elsewhere (3). Outpatients were recruited from the Centre for Cognitive Disorders, Department of Psychiatry and Psychotherapy, Klinikum Rechts der Isar, Technische Universität München, Munich, Germany. They had been referred by general practitioners, neurologists, psychiatrists, or other institutions, or were self-referred.

Patients had to fulfill standard diagnostic criteria for MCI (1) and were required to achieve a global score of 0.5 on the Clinical Dementia Rating scale (CDR) (4). Patients with other brain diseases such as normal-pressure hydrocephalus, brain tumors, or Parkinson disease; treatment with psychotropic drugs; addiction to alcohol or legal or illegal drugs; abnormal routine laboratory results; and symptoms indicative of other disorders that might lead to cognitive impairment, for example, sleep apnea, were excluded from the analyses.

All consecutively recruited patients who fulfilled these criteria; who underwent brain MRI, including 3-dimensional dataset, and ^{18}F-FDG and ^{11}C-PIB PET scans; and for whom clinical follow-up data after 24 mo were available were included in the analyses.

A study protocol was submitted to the ethic committee of the Faculty of Medicine of the Technische Universität München that raised no objections and was approved by radiation protection authorities. All patients provided written informed consent, and all clinical investigations were conducted in accordance with the principles of the Declaration of Helsinki, sixth revision.

Imaging

[11]C-PiB PET and ^{18}F-FDG PET scans followed standardized protocols (5,6). ^{11}C-PiB and ^{18}F-FDG images were coregistered to high-resolution 3-dimensional MRI scans and normalized to the Montreal Neurologic Institute space with the warping parameters of the MRI to obtain interindividually comparable images using standard procedures (7,8).

Automated Analyses

The fully automated method of ^{11}C-PiB image analysis was based on a reference tissue model (9): a cerebrum-to-vermis ratio was calculated (10) using the established threshold of above 1.4 for amyloid positivity (3).

For the fully automated image analysis of ^{18}F-FDG images a validated software program, the PMOD Alzheimer Discrimination Tool (PMOD Technologies Ltd.), was used (11). This software classifies ^{18}F-FDG images as either normal or abnormal for AD depending on the extent of voxels showing decreased tracer uptake in AD-typical regions

Visual Analyses

The visual analysis of ^{18}F-FDG and ^{11}C-PIB PET images was conducted by 2 independent and board-certified specialists in nuclear medicine who were masked to all clinical information, particularly the clinical outcome, and to the results of the automated analyses. As in daily clinical routine, the original color-coded axial images, surface projections, and surface projections of difference images compared with a normative dataset using the Neurostat routine (12) were provided. Images had to be classified as either typical of AD or not typical of AD. In addition, ^{18}F-FDG images that were rated as not typical of AD were subclassified in terms of whether they were typical of other neurodegenerative diseases (e.g., frontotemporal lobar degeneration or Lewy body dementia). To be able to compare the results of visual and fully automated analyses, visual analysis findings classified as typical of AD were categorized as positive for AD, whereas all other results were categorized as negative for AD.

A Pearson correlation analysis was performed for the ^{11}C-PiB PET assessment to provide an estimate of interrater reliability. In addition, a consensus rating was calculated for the 2 raters' visual analyses of the amyloid scans. A consensus estimate for the visual analysis of ^{18}F-FDG images could not be calculated because of the high level of discrepancy between the 2 raters' results.

Positive and negative predictive values (PPVs and NPVs, respectively) and the accuracy of predictions regarding conversion to dementia due to AD (hit rate) were calculated for both the visual and the fully automated analyses of ^{11}C-PIB and ^{18}F-FDG images. Furthermore, visual analysis data for ^{18}F-FDG images were used to calculate predictive values for conversion to dementia of any cause. A Pearson correlation analysis was performed as an estimate of agreement between visual and fully automated ratings.

The participants' 2-y outcome was determined on a syndrome level by the global CDR score (0, normal; 0.5, MCI; >0.5, dementia). The CDR rating was performed by a clinician who was masked for the baseline MCI subtype and for PET findings. In patients who had deteriorated to dementia, the etiology was specified using current standard criteria (13,14).

RESULTS

Twenty-eight patients were included in the study. At the time of the first assessment, the average age of participants was 67.1 y, with all of the participants meeting a global CDR score of 0.5. Participants' demographic data are presented in Table 1. Participants were scheduled to be reassessed after a 2-y follow-up period. Mean follow-up time was 31.2 ± 7.8 mo. Of a total of 28 patients, 9 developed dementia due to AD, 2 developed frontotemporal dementia, and 1 developed moderate dementia of unknown etiology.

The results of visual analyses including consensus ratings are presented in Supplemental Tables 1 and 2 (supplemental materials are available at http://jnm.snmjournals.org). The 2 independent raters had an interrater reliability (r) of 0.861 (P < 0.001) for the ^{11}C-PiB PET assessment and 0.268 (P = 0.168) for the ^{18}F-FDG PET assessment. Fully automated analyses results are presented in Supplemental Tables 3 and 4. PPV, NPV, and accuracy of conversion to dementia due to AD are shown in Table 2. NPVs of progression to dementia due to AD of ^{11}C-PIB PET were substantially higher (100%) than of ^{18}F-FDG PET (65%–78%), irrespective of the strategy of analysis. PPVs were generally low in predicting dementia due to AD at short-term follow-up, irrespective of imaging modality and analyzing strategy (20%–52%).

Visual analysis–based predictions of ^{18}F-FDG images of conversion to dementia of any cause resulted in moderate accuracy rates (64%–68%). Visual expert ratings are presented in Supplemental Table 5. Predictive values and accuracy are given in Table 3.

TABLE 1
Patient (n = 28) Characteristics

Characteristic	n or mean ± SD
Age (y) at baseline	67.1 ± 7.3 (50–78)
Sex	Male, 14/female, 14
CDR global at baseline	0.50 ± 0.00 (0.5–0.5)
CDR global at follow-up	0.88 ± 0.70 (0.0–3.0)
Follow-up interval in mo	31.2 ± 7.8 (24–53)
CDR global at follow-up ≥ 1	12
AD yes	9
AD no	3 (2 frontotemporal dementia, 1 moderate dementia of unknown etiology)

Data are n or mean ± SD, with ranges in parentheses.

TABLE 2
PPV, NPV, and Accuracy of Prediction of Dementia
Due to AD

Analysis	PPV in %	NPV in %	Accuracy in %
[11]C-PiB			
Rater 1	50.0	100.0	67.9
Rater 2	56.0	100.0	75.0
Consensus rating	50.0	100.0	67.9
Calculated cutoff values	52.9	100.0	71.4
^{18}F-FDG			
Rater 1	20.0	65.2	53.6
Rater 2	50.0	77.8	67.9
Calculated cutoff values	36.8	77.7	50.0

TABLE 3
Prediction of Dementia of Any Cause with Visual Rating of
^{18}F-FDG PET

Rater	PPV in %	NPV in %	Accuracy
1	58.8	81.8	67.9
2	56.3	75.0	64.3

DISCUSSION

The aim of this study was to determine how visual analyses and fully automated analyses of ^{11}C-PIB and ^{18}F-FDG imaging data compare in terms of their ability to predict conversion to dementia due to AD in patients with MCI.

The maximum PPV for prediction of AD dementia achieved by any of the analyses performed was almost random (56%), with slightly superior PPVs for ^{11}C-PiB, irrespective of analyzing strategy. This suggests that neither amyloid nor ^{18}F-FDG assessments lend themselves to predicting a conversion from MCI to dementia due to AD within a short follow-up interval of 2 y. Therefore, physicians might be well advised to exert caution with regard to the type of information provided to patients based on neuroimaging findings at the time of imaging.

A possible explanation for the low PPVs could be that persons with identical pathologies may vary greatly with regard to factors that determine their resilience to AD. This could make it challenging to predict the clinical course of the disease. The cognitive reserve hypothesis suggests that elderly people who have significant amyloid deposition but do not show signs of cognitive impairment may have more efficient mechanisms of resistance or compensation that protect them against cognitive decline (15,16).

However, the fact that all of the patients with MCI who converted to dementia due to AD had a positive amyloid scan at their initial assessment indicates that a positive amyloid scan should be regarded as a relevant risk factor for conversion to dementia due to AD. Furthermore, it would justify initiating a treatment that is aimed at slowing disease progression by reducing amyloid burden.

NPVs obtained for ^{11}C-PIB images, on the other hand, were excellent (100%), of both visual and fully automated analyses, indicating that a patient is not at risk of converting to dementia due to AD during the following 2 y. Although the NPV of ^{18}F-FDG images after visual analysis is as good as the NPV of ^{18}F-FDG images after fully automated analysis, NPVs for ^{18}F-FDG images were markedly lower (65%–78%) than those obtained for ^{11}C-PIB images. An amyloid scan is therefore clearly preferable to an ^{18}F-FDG scan if the purpose of scanning is to confirm that a patient is

not at risk of converting to dementia due to AD. Patients with MCI and a negative amyloid scan finding should undergo further investigations to assess other causes of cognitive impairment such as other forms of neurodegeneration, normal-pressure hydrocephalus, or depression.

[11]C-PIB image accuracy rates after visual analysis (68%) were similar to those achieved after fully automated analysis (71%). However, as the quality of visual analysis is heavily dependent on the experience of individual raters, the fully automated method of analysis should be considered as the preferred option.

Results were different for ^{18}F-FDG images, with accuracy rates after visual analysis higher (54%–68%) than those achieved after fully automated analysis (50%), thus suggesting that for ^{18}F-FDG images, visual analysis by experts should be regarded as the preferred option.

When PPV, NPV, and accuracy using visual analysis were calculated in relation to the ability of ^{18}F-FDG imaging to predict conversion to dementia of any cause, predictive values obtained were much better than those obtained in relation to the ability of ^{18}F-FDG imaging to predict conversion to dementia due to AD (Tables 2 and 3); in the case of rater 2, this also included a better accuracy. This finding underlines the strengths of visual analysis of ^{18}F-FDG imaging to enable differential diagnosis within 1 scan.

Results obtained in relation to the predictive values using fully automated analysis may have been influenced by the fact that conservative threshold values were selected for the ^{11}C-PIB scans (cerebrum-to-vermis ratio $\geq$ 1.4 (3,17)) and for the ^{18}F-FDG scans (cutoff, 11,089 voxels). However, applying a cutoff of 8,116 voxels for ^{18}F-FDG scans to our sample, as proposed in a recent study aiming at discriminating patients with MCI who later converted to AD from healthy controls (18), resulted in a comparable NPV (77.8%), whereas PPV and accuracy even decreased (31.8% and 39.3%).

The discriminating power of visual or fully automated analysis of ^{18}F-FDG imaging improves in research scenarios, for example, to differentiate MCI subjects who converted to AD dementia from normal control subjects (18,19). However, this does not closely resemble clinical routine.

One major limitation of the current study is the small sample size. Therefore, our findings should be confirmed in a larger group of patients, although a PPV of 36.8% and NPV of 77.7% of the fully automated analysis of ^{18}F-FDG images in our study are comparable to a previous study that included a larger cohort. When a calculated cutoff to differentiate 85 MCI subjects who progressed to AD dementia during a follow-up of 2 y from those who did not was used, PPV was 41% and NPV was 79% (20). Another limitation is the relatively short follow-up interval; however, it reflects typical scenarios in clinical routine. However, a longer follow-up time would have presumably increased the PPVs.

CONCLUSION

The results of this study show that amyloid PET scans and ^{18}F-FDG scans are of limited benefit in terms of their ability to predict conversion to dementia due to AD in patients with MCI at short-term follow-up, but amyloid PET is extremely useful to rule out a risk of underlying AD. This study favors a fully automated method of analysis for ^{11}C-PiB assessments and a visual analysis by experts for ^{18}F-FDG assessments.

DISCLOSURE

REFERENCES

1. Winblad B, Palmer K, Kivipelto M, et al. Mild cognitive impairment: beyond controversies, towards a consensus: report of the International Working Group on Mild Cognitive Impairment. *J Intern Med.* 2004;256:240–246.
2. Albert MS, DeKosky S, Dickson D, et al. The diagnosis of mild cognitive impairment due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement.* 2011;7:270–279.
3. Grimmer T, Wutz C, Drzezga A, et al. The usefulness of amyloid imaging in predicting the clinical outcome after two years in subjects with mild cognitive impairment. *Curr Alzheimer Res.* 2013;10:82–85.
4. Morris JC. The Clinical Dementia Rating (CDR): current version and scoring rules. *Neurology.* 1993;43:2412–2414.
5. Lopresti B, Klunk W, Mathis C, et al. Simplified quantification of Pittsburgh compound B amyloid imaging PET studies: a comparative analysis. *J Nucl Med.* 2005;46:1959–1972.
6. Grimmer T, Diehl J, Drzezga A, Förstl H, Kurz A. Region-specific decline of cerebral glucose metabolism in patients with frontotemporal dementia: a prospective ^{18}F-FDG-PET study. *Dement Geriatr Cogn Disord.* 2004;18:32–36.
7. Grimmer T, Riemenschneider M, Forstl H, et al. Beta amyloid in Alzheimer's disease: increased deposition in brain is reflected in reduced concentration in cerebrospinal fluid. *Biol Psychiatry.* 2009;65:927–934.
8. Grimmer T, Tholen S, Yousefi B, et al. Progression of cerebral amyloid load is associated with the apolipoprotein E epsilon4 genotype in Alzheimer's disease. *Biol Psychiatry.* 2010;68:879–884.
9. Ziolko SK, Weissfeld L, Klunk W, et al. Evaluation of voxel-based methods for the statistical analysis of PIB PET amyloid imaging studies in Alzheimer's disease. *Neuroimage.* 2006;33:94–102.
10. Loening AM, Gambhir SS. AMIDE: a free software tool for multimodality medical image analysis. *Mol Imaging.* 2003;2:131–137.
11. Herholz K, Salmon E, Perani D, et al. Discrimination between Alzheimer dementia and controls by automated analysis of multicenter FDG PET. *Neuroimage.* 2002;17:302–316.
12. Minoshima S, Frey K, Koeppe R, Foster N, Kuhl D. A diagnostic approach in Alzheimer's disease using three-dimensional sterotactic surface projections of fluorine-18-FDG PET. *J Nucl Med.* 1995;36:1238–1248.
13. The Lund and Manchester Groups. Clinical and neuropathological criteria for frontotemporal dementia. *J Neurol Neurosurg Psychiatry.* 1994;57:416–418.
14. McKhann GM, Knopman D, Chertkow H, et al. The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement.* 2011;7:263–269.
15. Roe CM, Mintun M, D'Angelo G, Xiong C, Grant E, Morris J. Alzheimer disease and cognitive reserve: variation of education effect with carbon 11–labeled Pittsburgh compound B uptake. *Arch Neurol.* 2008;65:1467–1471.
16. Alexander GE, Furey M, Grady C, et al. Association of premorbid intellectual function with cerebral metabolism in Alzheimer's disease: implications for the cognitive reserve hypothesis. *Am J Psychiatry.* 1997;154:165–172.
17. Jack CR, Lowe V, Senjem M, et al. ^{11}C-PiB and structural MRI provide complementary information in imaging of AD and amnestic MCI. *Brain.* 2008;131:665–680.
18. Pagani M, De Carli F, Morbelli S, et al. Volume of interest-based [^{18}F]fluorodeoxyglucose PET discriminates MCI converting to Alzheimer's disease from healthy controls: a European Alzheimer's Disease Consortium (EADC) study. *Neuroimage Clin.* 2015;7:34–42.
19. Morbelli S, Brugnolo A, Bossert I, et al. Visual versus semi-quantitative analysis of ^{18}F-FDG-PET in amnestic MCI: an European Alzheimer's Disease Consortium (EADC) project. *J Alzheimers Dis.* 2015;44:815–826.
20. Landau SM, Harvey D, Madison CM, et al. Comparing predictors of conversion and decline in mild cognitive impairment. *Neurology.* 2010;75:230–238.