Original Research Article

Added Diagnostic Value of $^{11}$C-PiB-PET in Memory Clinic Patients with Uncertain Diagnosis

K.S. Frederiksen$^a$ S.G. Hasselbalch$^a$ A.-M. Hejl$^a$ I. Law$^b$
L. Højgaard$^b$ G. Waldemar$^a$

$^a$Memory Disorders Research Group, Department of Neurology, and $^b$Department of Clinical Physiology, Nuclear Medicine and PET, Copenhagen University Hospital, Rigshospitalet, Copenhagen, Denmark

Key Words
$^{11}$C-PiB-PET · Diagnosis · Added diagnostic value · Alzheimer’s disease · Neurodegenerative disease · Imaging

Abstract

Introduction: The added diagnostic value of $^{11}$C-PiB-PET for the assessment of the accumulation of cortical beta-amyloid in memory clinic patients with uncertain diagnosis remains undetermined. Methods: All patients who underwent PiB-PET at the Copenhagen Memory Clinic between March 2008 and November 2011 were included in this uncontrolled, retrospective study. The standard diagnostic evaluation program included physical and neurological examination, cognitive and functional assessment, a cranial CT or MRI, functional imaging and cerebrospinal fluid sampling. Based on anonymized case reports, three experienced clinicians reached a consensus diagnosis and rated their confidence in the diagnosis before and after disclosure of PiB-PET ratings. PiB-PET scans were rated as either positive or negative. Results: A total of 57 patients (17 females, 30 males; age 65.7 years, range 44.2–82.6) were included in the study. Twenty-seven had a positive PiB-PET scan. At the first diagnostic evaluation, 16 patients were given a clinical Alzheimer’s disease diagnosis (14 PiB positive). Of the 57 patients, 13 (23%) were diagnostically reclassified after PiB-PET ratings were disclosed. The clinicians’ overall confidence in their diagnosis increased in 28 (49%) patients. Conclusion: PiB-PET adds to the specialist clinical evaluation and other supplemental diagnostic investigations in the diagnostic classification of patients with uncertain diagnosis in a specialized memory clinic.

Copyright © 2012 S. Karger AG, Basel

Kristian Steen Frederiksen
Memory Disorders Research Group, Department of Neurology
Copenhagen University Hospital, Rigshospitalet, Blegdamsvej 9
DK–2100 Copenhagen (Denmark)
E-Mail kristian.steen.frederiksen@rh.regionh.dk
Introduction

The diagnosis of Alzheimer’s disease (AD) has previously been established within a probabilistic diagnostic framework with an assumed clinicopathological correlation where definite AD requires neuropathological confirmation [1]. In recent years, revised diagnostic criteria have been proposed [2–4]. These include incorporating biomarkers such as measurement of beta-amyloid (Aβ), phospho-tau (P-tau) and total-tau (T-tau) in cerebrospinal fluid (CSF), hippocampal volume on MRI, and cerebral metabolism on fluoro-deoxy-glucose positron emission tomography (FDG-PET), and defining a framework allowing for the diagnosis of AD prior to the dementia stage. The establishment of an AD diagnosis within this framework is, to a large degree, based on assumptions on and the detection of the underlying etiology [3, 5]. In this perspective, the Aβ-binding PET ligand Pittsburgh compound B (PiB) may improve diagnostic accuracy.

Several studies have reported sensitivity and specificity of PiB-PET for AD patients versus normal aging individuals [6, 7] or frontotemporal lobar degeneration (FTLD) patients [8–10]. This approach, for which the clinical AD diagnosis is ‘the gold standard’, does not specifically assess the diagnostic impact of PiB-PET in the clinical setting, where there is often diagnostic uncertainty and where ancillary investigations are available. Therefore, we sought to assess the value of PiB-PET when added to the specialist clinical evaluation and supplemental diagnostic investigations such as MRI and CT, other imaging modalities, and CSF-based fluid biomarkers. This was examined in patients from a specialized memory clinic, who presented atypical symptoms and disease course, and in whom there was diagnostic uncertainty. This is, to our knowledge, the first study to do so.

Methods

Patients

All patients who underwent PiB-PET imaging at the Copenhagen Memory Clinic, Department of Neurology, Copenhagen University Hospital, Rigshospitalet (Copenhagen, Denmark) between March 2008 and November 2011 were included in the study. The clinic is a multidisciplinary outpatient clinic specialized in neurology. The clinic offers diagnostic evaluation and treatment of cognitive disorders and dementia, and receives secondary and tertiary referrals from general practitioners, neurologists, psychiatrists and other hospitals. Patients may also be referred from other memory clinics for second opinion evaluations. Patients with rare (such as genetic disorders) or uncertain etiology are referred from other Danish counties, which is reflected in the wide age range of the patients in the clinic. A total of 568 patients with an average age of 65.9 years (range 16–95) were referred to the clinic in 2009. The most common dementia diagnosis was AD, but more rare conditions such as Huntington’s disease and FTLD were also diagnosed. A significant number of patients evaluated at the clinic did not fulfill the criteria for dementia and had cognitive symptoms due to other conditions.

The patients in this study underwent the standard dementia assessment program including a physical and neurological examination as well as cognitive [Mini-Mental State Examination (MMSE) and Addenbrooke’s Cognitive Examination (ACE)] and functional assessment [Functional Assessment Questionnaire-Activities of Daily Living (FAQ-ADL)]. Most were seen by a neuropsychologist for further cognitive assessment. Moreover, all subjects underwent structural scans (MR and/or CT). Functional imaging included 18F-FDG-PET, cerebral blood flow single-photon emission computed tomography (CBF-SPECT) and dopamine transporter SPECT (DAT-SPECT). Routine laboratory test screening was performed in
all patients, and diagnostic CSF samples were collected from a majority of patients and analyzed for protein, cells, oligoclonal bands, Aβ42, P-tau and T-tau (see table 1 and Results section).

Most patients in the present study had been followed up at the memory clinic for some time prior to referral to PiB-PET, but some were also referred as part of the initial diagnostic evaluation. All patients had been referred to PiB-PET to confirm or rule out AD. Specific reasons for referral were: atypical symptoms or disease course in relation to AD, differential diagnosis between AD and other neurodegenerative diseases/neurological conditions, differential diagnosis between AD and FTLD, lack of progression in a patient who had been given an AD diagnosis and differential diagnosis between AD and depression.

Probable AD was diagnosed according to the criteria proposed by McKhann et al. [4]. Other relevant specific dementia diagnoses were FTLD, mixed dementia [AD/vascular dementia (VaD)], cerebral amyloid angiography (CAA) and corticobasal degeneration (CBD). If patients with cognitive deficits did not meet the criteria for dementia, a diagnosis of prodromal AD according to the criteria of Dubois et al. [2], a diagnosis of amnestic mild cognitive impairment (aMCI) according to the revised criteria by Petersen [11] or a diagnosis of vascular cognitive impairment (VCI) could be given. When relevant, patients were classified as having subjective cognitive complaints (SCC) if no evidence for any objectively measurable cognitive impairment was found.

All participants provided written informed consent. Consent was obtained in accordance with the Declaration of Helsinki of 1975, and the study was approved by the Capital Region Ethics Committee.

### Diagnostic Rating

Anonymized case reports of each patient with all available clinical information from the first visit in the memory clinic up until the PiB-PET scan, including clinical notes from all visits in the memory clinic, results from blood and CSF samples, genetics and imaging (excluding PiB-PET), were presented in printed form to three experienced clinicians (A.-M.H., S.G.H., G.W.). Imaging results were presented as written reports, but if clinicians requested to inspect the scans themselves this was possible. The clinicians proceeded to reach a consensus diagnostic classification after which PiB-PET ratings were disclosed and the clinicians were asked to re-evaluate their initial diagnostic classification. Similarly, the clinicians were asked to rate their confidence in consensus with regard to (1) their confidence in the diagnosis they had reached and (2) their confidence in confirming or ruling
out AD. The process was carried out as a consensus exercise, where a decision was not taken until all clinicians were in agreement. This was done to best mimic the process in the clinical setting. The overall confidence was rated on a 3-level scale (low, medium and high confidence). For AD confidence, a 6-level scale was used with the following levels: clear AD/non-AD, probably AD/non-AD, and subtle AD/non-AD. If the clinicians were very confident, a ‘clear’ rating was given; for some confidence, a ‘probable’ rating was given, and for little confidence a ‘subtle’ rating was given. Both the diagnostic classification and rating of confidence were based on the assumptions that a negative PiB-PET excluded the diagnosis of AD with a high likelihood, whereas a positive PiB-PET was considered a prerequisite for AD in most cases.

**PET Scans**

PET was conducted using a Siemens Biograph 40 or Biograph 64 PET/CT scanner. Patients underwent a $^{11}$C-PiB-PET scan to measure the Aβ burden: a bolus injection of $^{11}$C-PiB-PET with an average activity of 440 MBq (range 169–720) was given and a PET scan was acquired 40–70 min after injection as 6 × 5 min frames. The scans were read on a Siemens Syngo workstation using the software TrueD in the Rainbow color scheme and after coregistration to either a recent MRI scan, a diagnostic cerebral CT scan or a low-dose CT scan.

The PiB-PET scans were visually rated by one neuro-PET expert (I.L.) who has extensive experience in rating PiB-PET scans. The neuro-PET expert was blinded to all demographic and clinical data of the individual patients. The scans were rated either positive if two or more cortical regions out of eight (frontal, parietal, temporal or occipital right and left lobes) showed an increased uptake relative to the white matter, or negative if the uptake in all cortical region was below that of the unspecific activity in the white matter. In cases where one unilateral cortical region showed an increased uptake, the scan was still considered negative, but clinicians were informed that the uptake was elevated in one region. Previous studies have used a similar dichotomization in rating PiB-PET scans [9, 12]. See figure 1 for examples of scans.

**Analysis of Aβ42, T-tau and P-tau in CSF**

CSF was sampled in nonabsorbant polypropylene test tubes and sent by ordinary mail to the Statens Serum Institut for analysis. Quantification of CSF levels of Aβ1–42 [Innotest – amyloid (1–42)], T-tau [Innotest – hTau Ag] and P-tau [Innotest – Phospho-tau (181P)] was carried out using ELISA from Innogenetics (Ghent, Belgium).

**Statistical Analysis**

Data are presented as mean ± standard deviation (SD) or number of patients, if not otherwise specified. Primary outcome measures were the change in diagnosis, change in confidence of the diagnostic classification and change in confidence of whether the patient did or did not have AD. To quantify the added diagnostic value of PiB-PET, we applied the ‘numbers needed to test’ (NNT) as described by others [13]. To calculate the NNT, we used the following formula: NNT = 1/[(Pa/Ta) – (Pb/Tb)], where Pa is the number of patients given the correct diagnoses and Ta the total number of subjects after PiB-PET, Pb is the number of patients given the correct diagnoses and Tb the total number of subjects before PiB-PET. Akin to ‘numbers needed to treat’, NNT estimate the number of patients who need to undergo the diagnostic procedure for 1 patient to change diagnosis. Where relevant, categorical variables were compared with Fisher’s exact test. The significance level was set at p = 0.05 (two-tailed). Statistical analyses were carried out using Intercooled Stata 9.2 for Macintosh (Stata Corporation, USA).
Results

Fifty-seven patients were included in the study. Prior to disclosure of PiB-PET ratings, the following diagnoses were given to the patients: 16 probable AD, 13 with neurodegenerative disorder of undetermined etiology, 6 depression, 4 FTLD, 3 CAA, 3 prodromal AD, 3 with SCC, 2 with cognitive deficits due to cerebrovascular insults, 2 with neurological disorders not associated with neurodegeneration, 1 aMCI, 1 VCI, 1 CBD, 1 mixed dementia (VaD/AD) and 1 post-traumatic stress disorder. Prior to PiB-PET, all patients in the study had undergone extensive supplemental diagnostic testing, including 35 patients who had undergone CSF sampling, 42 MRI and 42 CT, and 50 FDG-PET. Demographical data and further clinical characteristics are displayed in table 1 and CSF biomarker profiles in table 2.
A total of 13 patients were reclassified with regard to diagnosis following disclosure of PiB-PET ratings. Seven patients diagnosed with a neurodegenerative disorder of undetermined etiology were reclassified, making it the group in which most patients were reclassified (4 patients reclassified to FTLD, 1 patient to mixed dementia (AD/VaD), 1 patient to prodromal AD and 1 patient to probable AD). This was followed by the probable AD group with 2 patients reclassified (1 reclassified to neurodegenerative disorder of undetermined etiology and 1 to FTLD) and 1 prodromal AD (to aMCI). Additionally, 3 patients were reclassified from the CBD, depression and SCC groups, respectively, to probable AD (table 3). This led to a NNT of 4.2 (SEM ±7.5) for probable AD/prodromal AD and a NNT of 4.4 (SEM ±11.2) for all diagnoses. There was no difference in the distribution of PiB-positive scans between patients who were reclassified and patients who were not (positive 6 vs. negative 7, p > 0.05). Concerning ancillary investigations, 7 patients (53.8%) who were reclassified had CSF biomarkers available versus 28 patients (63.6%) in the group of patients not reclassified, and 12 reclassified patients (92.3%) had undergone FDG-PET versus 38 patients (86.3%) not reclassified (table 2).
The clinicians’ overall confidence in their diagnosis increased for 28 patients and remained unchanged for 27 patients following disclosure of PiB-PET ratings. Prior to disclosure of PiB-PET ratings, 20 ‘high’, 28 ‘medium’ and 9 ‘low’ ratings were given, which changed to 44 ‘high’, 11 ‘medium’ and 2 ‘low’ ratings following disclosure. In the group of patients not diagnostically reclassified, 14 were given a ‘high’ rating prior to the scan, which increased to 34 after the scan (fig. 2). There was a significant effect of the result of the PiB-PET on the number of patients in whom overall confidence changed (positive 18, negative 9; p < 0.05).

Regarding the clinicians’ ratings of their confidence in diagnosing or excluding AD, an increase was observed in 42 patients, a decrease in 2 patients, and for 13 patients it remained unchanged. There was no difference between PiB-positive and PiB-negative patients as to change in confidence in diagnosing or excluding AD (positive 20, negative 20; p > 0.05).

Discussion

We examined the added value of PiB-PET to the diagnostic assessment program for patients with an uncertain diagnosis in a memory clinic setting. This is, to our knowledge, the first study to examine the added diagnostic value of PiB-PET. PiB-PET led to a high number of diagnostic reclassifications. Furthermore, clinicians reported a higher confidence in their diagnosis following PiB-PET, also in patients who were not diagnostically reclassified. Our findings suggest that PiB-PET adds value to the standard dementia assessment program in a specialized memory clinic setting.

Previous studies have reported sensitivity and specificity of PiB-PET in discriminating AD from other dementias and normal aging [6–10]. These studies were performed in patients with a relatively clear clinical diagnosis. However, our primary aim was to examine the impact of PiB-PET on the diagnostic evaluation when added to the clinical evaluation by a specialist and supplemental investigations such as MRI, CT, FDG-PET, CSF biomarkers and neuropsychological examination in patients with an uncertain diagnosis. We found that 23% of the patients were diagnostically reclassified after disclosure of the PiB-PET results. A previous study has examined the added value of CSF biomarkers in a similar manner and found that diagnostic reclassifications occurred in 10% [14] in a population of patients from a less specialized setting. This indicates that PiB-PET may have a higher impact on diagnosis than CSF biomarkers, but head-to-head comparisons in the same patients are required to support
this hypothesis, as indeed would a comparison with other diagnostic markers [15]. Moreover, 35 patients in our study, including 7 patients who were reclassified, had undergone CSF sampling prior to PiB-PET, in addition to a high rate of multimodal imaging (see table 3). This highlights the positive added value of PiB-PET even in patients in whom extensive diagnostic testing has been performed, although our findings are not able to provide data on which ancillary tests may have the highest added value. The higher rate may also reflect the fact that PiB-PET rating yields a dichotomous outcome, which is in contrast to CSF measurement, where a panel of three biomarkers (Aβ42, T-tau, P-tau) is usually reported. Thus, CSF biomarker profiles may be ambiguous and introduce uncertainty concerning the diagnostic interpretation. However, it should be stated that cutoff values yielding dichotomous outcomes for the three AD biomarkers are widely available, and that the possibility of ambiguity is a function of the additional information imbedded in the AD CSF biomarker panel, which both informs about brain amyloid as well as neurodegeneration. We calculated the NNT and found that approximately 4–5 patients need to undergo PiB-PET to change the diagnosis of 1 patient. Geroldi et al. [13], in a study of the added diagnostic value of neuropsychological assessment and neuroimaging reported a NNT 8.9 for all diagnoses with lowest numbers in diseases other than AD. However, a direct comparison may not be relevant due to the different levels of invasiveness as well as the fact that the patients undergoing PiB-PET are likely to be very different from those undergoing neuropsychological assessment and neuroimaging only.

Nearly half of the patients who were reclassified had an initial diagnosis of a neurodegenerative disorder of undetermined etiology, typically because clinicians were unable to assign a higher probability to one of two diagnoses. This was especially the case with the diagnosis of AD versus FTLD, a common clinical challenge, in which PiB-PET has been reported to have a high specificity and sensitivity for AD [8–10].

The clinicians’ overall confidence in their diagnosis increased in nearly half of all patients following disclosure of PiB-PET ratings. Even in patients in whom PiB-PET did not lead to a reclassification concerning the diagnosis, there was an added value of PiB-PET in the form of a higher confidence of the clinicians in the diagnosis. We found that the clinicians’ overall confidence increased significantly more for patients with a positive PiB-PET. It may be speculated that a negative PiB-PET scan would only be useful in cases where AD and one other specific diagnosis are considered (e.g. AD and FTLD), whereas a positive PiB-PET would be useful also if several PiB-negative disease entities are considered. This would explain the higher increase in confidence following disclosure of positive PiB scans.

Clinicians also rated their confidence in confirming or ruling out AD. Here, we also found a large increase in confidence, which exceeded the increase in the overall confidence. This may reflect a higher use of negative PiB-PET scans in terms of ruling out AD, which is also supported by the difference in the effects of positive versus negative PiB-PET scans on the overall confidence and confidence in confirming and ruling out AD. We evaluated PiB-PET by visual rating rather than quantitative assessment. This was based on several observations. Firstly, for a quantitative assessment it is necessary to implement an image-processing pipeline, which may be difficult to apply in clinical practice. An MRI would also most likely be a prerequisite. In contrast, for visual rating little infrastructure is necessary; in addition, this method is not time-consuming and is fairly simple due to the dichotomous nature of PiB-PET. Secondly, PiB-PET-measured Aβ deposition may reach a plateau where patients can be characterized as PiB positive relatively early in AD, and little evolution is observed over time [16]. Therefore, additional grading of PiB retention beyond the dichotomous ‘positive or negative’ may be superfluous in a clinical context. Thirdly, visual rating performs as well as or better than quantitative assessment with regard to sensitivity and specificity [9, 10].
The basis upon which clinicians classified patients concerning diagnosis and confidence following disclosure of PiB-PET ratings was that a positive PiB-PET scan carries a very high likelihood for the person to harbor AD pathology. Conversely, a negative PiB-PET scan carries a high likelihood that a person does not have AD, despite the possible presence of cognitive deficits and dementia.

The current knowledge in line with the amyloid hypothesis supports these assumptions. Although the pathophysiological cascade leading from healthy aging to dementia due to AD is not fully understood, the prevailing evidence shows a prominent role of Aβ deposition in allo- and neocortical areas [17]. However, some observations from PiB studies may seem to counter these assumptions. For example, it has been shown that between 10–30% of patients with a clinical diagnosis of probable AD have no increase in PiB retention [6–8, 18]. Similar numbers have been reported when the clinical diagnostic criteria for AD have been validated against neuropathological confirmation of AD [19]. Therefore, PiB-negative AD patients in most cases probably represent diagnostically misclassified patients rather than an absence of Aβ or failure of PiB to bind to Aβ, although this cannot be ruled out [20]. Furthermore, elderly persons without dementia or cognitive impairment may show signs of cerebral Aβ on PiB-PET scans. The prevalence of this type of PiB positivity may well be age related indicating that the diagnostic value of PiB-PET may decrease with age as the risk of misclassification of patients with positive scans increases [6, 17, 21, 22]. However, a recent study in dominant inherited AD has shown that PiB-PET was positive in mutation-positive individuals up to 15 years prior to expected symptom onset. Thus, PiB-PET positivity in healthy elderly subjects may reflect AD in a very early phase [23].

This study has some limitations. It was an uncontrolled study, and a randomized design would have given a better estimation of the impact of PiB-PET. This would also have enabled head-to-head comparisons with other diagnostic strategies such as CSF biomarkers [24]. Although the aim was not to validate PiB-PET against AD pathology but rather to examine the practical translation of PiB-PET into clinical practice, postmortem verifications of diagnoses or follow-up data would have enabled the validation of the clinicians’ diagnosis, also regarding PiB-negative patients [25]. Moreover, we did not investigate whether referral to PiB-PET was used reasonably in the first instance or whether other diagnostic strategies such as CSF sampling should have been applied in patients in whom this had not been carried out. This may influence the findings of the study, resulting in either a higher or a lower change rate of diagnosis and confidence. The extensive diagnostic investigations which the patients in this study had undergone serve to prove the usefulness of PiB-PET in patients in whom other diagnostic tools may have been insufficient. This also indicates that the cohort of patients was not typical of patients in less specialized memory clinics but rather represented atypical cases. Furthermore, concerning the generalizability of our results, it should be kept in mind that PiB-PET is not available in less specialized centers, and that readers in less specialized centers may not read scans as well as specialized readers. However, the advancement of 18F-labelled ligands [26–28] and the recent approval of one such ligand, florbetapir, for clinical use in the United States will enable the implementation of amyloid imaging outside highly specialized centers. Therefore, our results should be validated with these new ligands, and efforts in this direction are under way [29, 30]. The introduction of amyloid imaging in less specialized centers may also result in an increased value of PiB-PET because patients may have had undergone less extensive diagnostic testing compared to patients in highly specialized centers. Lastly, as outlined in the Introduction, the basis for the clinicians to re-evaluate their diagnosis following disclosure of PiB-PET readings was that the presence of cerebral Aβ is highly associated with AD; in other words, the clinicians believed that the amyloid hypothesis is valid. This may have led to a confirmation bias. However, the fact that some patients with
a positive scan were not diagnosed as having AD indicates that clinicians were not so inclined.

The strengths of our study are the relatively large number of patients and the extensive diagnostic investigations undertaken prior to PiB-PET. Moreover, the design of the study enabled the assessment of PiB-PET as an integral part of the diagnostic process by which the clinician reached a diagnosis. By including the clinicians’ rating of their confidence in the diagnosis, we were able to assess the diagnostic evaluation undertaken in more detail. AD is a probabilistic diagnosis, and therefore it may be assumed that the clinicians’ confidence is a valid measure of the certainty of a diagnosis and will influence the clinicians’ treatment choices.

In conclusion, our study supports the hypothesis that PiB-PET has an impact on the final clinical diagnosis when added to MR/CT, FDG-PET, neuropsychological examination and clinical evaluation performed by a specialist in a specialized memory clinic. The higher confidence in diagnosis achieved following PiB-PET may indicate a more certain diagnosis and may facilitate intensified treatment and psychosocial support. Further studies are needed to validate our findings as well as examine whether they are relevant for less specialized centers. Studies examining 18F-labelled ligands are under way and will assist in this endeavor. Importantly, post-PiB-PET scan diagnoses need to be validated through longitudinal follow-up and neuropathology. Furthermore, the impact on treatment and cost-effectiveness, which are the consequences of the diagnostic reclassifications following scans, also need to be evaluated.

Disclosure Statement

The authors have no conflict of interest to declare.

References


Avid Radiopharmaceuticals: Potential of Florbetapir F 18 PET to Inform Clinical Diagnosis and Management of Patients with Progressive Cognitive Decline. clinicaltrials.org/show/NCT01703702.

Avid Radiopharmaceuticals: Effectiveness of Florbetapir (18F) PET Imaging in Changing Patient Management and the Relationship Between Scan Status and Cognitive Decline. clinicaltrials.org/show/NCT01400425.