Theoretical Impact of Florbetapir (\(^{18}\)F) Amyloid Imaging on Diagnosis of Alzheimer Dementia and Detection of Preclinical Cortical Amyloid

Thomas G. Beach, MD, PhD, Julie A. Schneider, MD, Lucia I. Sue, BS, Geidy Serrano, PhD, Brittany N. Dugger, PhD, Sarah E. Monsell, MS, and Walter Kukull, PhD

Abstract

In 2012, florbetapir (\(^{18}\)F) (Amyvid) received US Food and Drug Administration approval as a diagnostic agent for detecting neuritic (β-amyloid) plaques in living patients. Although such approval is specifically not extended to the use of florbetapir as a single definitive diagnostic test for Alzheimer disease dementia (ADD), it is of considerable importance to examine its potential in this regard. To estimate the ability of florbetapir amyloid imaging to detect specified densities of postmortem-identified neuritic plaques, we used the data of Clark et al [Clark CM, Pontecorvo MJ, Beach TG, et al. Cerebral PET with florbetapir compared with neuropathology at autopsy for detection of neuritic amyloid-beta plaques: A prospective cohort study. Lancet Neurol 2012;11:669–78]. We then used the data of Beach et al [Beach TG, Monsell SE, Phillips LE, et al. Accuracy of the clinical diagnosis of Alzheimer disease at National Institute on Aging Alzheimer Disease Centers, 2005–2010. J Neuropathol Exp Neurol 2012;71:266–73], derived from the National Alzheimer’s Coordinating Center, to estimate the fraction of subjects who would have been called florbetapir-positive and, among these, the fraction of subjects who would also meet neuropathologic criteria for the presence of ADD.

The accuracy of a positive florbetapir β-amyloid scan for the detection of neuropathologically defined ADD is estimated at between 69% and 95% sensitivity and between 83% and 89% specificity. From the same National Alzheimer’s Coordinating Center data set, 144 subjects were recorded as having normal cognition. Among these, 84 (58%) had at least sparse neuritic plaques at autopsy and, among these, florbetapir imaging was estimated to detect 47 (56%). These findings suggest that amyloid imaging may significantly improve the clinical identification of ADD.

Key Words: Alzheimer disease, Amyloid imaging, Autopsy, Dementia, Diagnosis, Neuropathology, Sensitivity, Specificity, Therapy.

INTRODUCTION

The relative inaccuracy of the clinical diagnosis of Alzheimer disease dementia (ADD) may be a major impediment to clinical trials of candidate therapeutic agents (1). A recent study found that sensitivity ranged from 70.9% to 87.3% whereas specificity ranged from 44.3% to 70.8%, depending on the confidence levels of clinical and neuropathologic criteria (2). With the use of “clinically probable” ADD as clinical diagnosis (3) and the combination of moderate or frequent cortical Consortium to Establish a Registry for Alzheimer’s Disease (CERAD) neuritic plaque densities with Braak Neurofibrillary Stages III to VI as neuropathologic definition (4) (the most commonly accepted clinical and neuropathologic criteria across the period examined), sensitivity and specificity were both only approximately 71%. For estimated drug effect sizes lower than 50%, such diagnostic inaccuracy would require at least a doubling of subject number to achieve sufficient statistical power for an adequate clinical trial. If the diagnostic inaccuracy is ignored, the clinical trial would likely be seriously underpowered. Improving diagnostic accuracy would allow clinical trials with smaller subject numbers at lower cost, thereby expediting the discovery of new efficacious agents.

Current data suggest that early treatment of Alzheimer disease (AD) may be more likely to prevent irreversible brain damage, including synaptic loss and neuronal loss. In fact, it may be best to begin such preventive treatment at the preclinical stage of AD, when pathophysiologic disease processes have commenced but the disease is not yet clinically manifest (5). This preclinical stage is thought to extend across 1 to 2 decades.
or more (6), with neocortical β-amyloid biochemical and histologic accumulation thought to be the first AD-specific alteration (7–11). Therefore, it may be critical to identify cortical β-amyloid deposition as early as possible to facilitate early prevention trials.

Imaging β-amyloid in living subjects offers the promise of assisting with both of these objectives. The separation of demented subjects with or without significant cortical β-amyloid deposits should allow the rejection of the diagnosis of ADD in subjects with a negative β-amyloid scan, and the identification of cortical β-amyloid deposits in cognitively normal older subjects should allow the selection of subjects for prevention trials who are in the preclinical stage of AD. A quantitative estimate of the potentialities and limitations of β-amyloid imaging with respect to these objectives, using neuropathologically defined ADD as gold standard, has been lacking and is the subject of the present communication. To estimate the ability of florbetapir amyloid imaging to detect specific densities of postmortem-identified neuritic plaques, we used data derived from a published study of 59 subjects who underwent antemortem amyloid imaging with florbetapir, followed by death, autopsy, and histopathologic evaluation (12). We then used data from an autopsy study of 919 subjects, derived from the National Alzheimer’s Coordinating Center (NACC) (2), to estimate the fraction of subjects who would have been florbetapir-positive and, among these, the fraction that would also meet neuropathologic criteria for the presence of ADD.

**MATERIALS AND METHODS**

**Florbetapir Imaging Subjects**

Subjects were derived from those described in a previous publication, where details of recruitment, imaging, tissue processing, and analytic methodology were described (12, 13). Briefly, patients near the end of their lives were recruited from hospice, long-term care, and community healthcare facilities for florbetapir positron emission tomography (PET) scanning. The inclusion and exclusion criteria were as follows: 1) a physician’s assessment that the individual was likely to die within 6 months of study; 2) the absence of any known destructive lesion in the brain (e.g., stroke or tumor); and 3) the individual’s willingness to have florbetapir imaging followed by brain necropsy at the time of death. A flow chart (Fig. 1 in the original publication [13]) schematically illustrates subject recruitment and analysis. One hundred fifty-two subjects were recruited and from the decision of the least accurate of the 5 readers (Reader 5).

**NACC Subjects**

Data were derived from an NACC database search used in a prior publication (2). Subject data consisted of stipulated data elements from the NACC Uniform Data Set, obtained with the assistance of NACC personnel. The NACC Uniform Data Set has been collected since September 2005 from more than 30 National Institute on Aging (NIA) Alzheimer’s Disease Centers (ADCs) located throughout the United States. Most ADCs are at university medical centers in urban settings. Research subjects are generally recruited from the practices of participating neurologists with some additional community-based recruitment. The initial data pull included all 1,198 subjects who had at least 1 Uniform Data Set-compliant clinical assessment and then had died and were autopsied before September 2010. From this, 144 subjects were considered cognitively normal during life, whereas 919 subjects had been diagnosed as having dementia; 135 subjects were excluded after having been diagnosed as having mild cognitive impairment or because critical data fields were either not filled out or were marked “missing” or “not done.” The clinical diagnosis of “dementia” or “cognitively normal” was that given at the last assessment during life.

**Florbetapir PET Imaging Methods**

The details of imaging methods have been previously described (12). Briefly, each subject underwent a 10-minute PET scan at 50 minutes after receiving an intravenous bolus of 370 MBq (10 mCi) florbetapir (18F). Acquired PET scans were reconstructed either by iterative reconstruction with a postreconstruction Gaussian filter or by row action maximal likelihood algorithms to a 128 × 128 matrix with a zoom of 2.0 to 2.33. Florbetapir (18F) PET images were assessed visually as either positive or negative by 5 board-certified nuclear medicine physicians blind to each other’s readings and to all clinical and neuropathologic data. For each reader, an intense level of tracer uptake in any single cortical region or a significant signal in any 2 cortical regions was sufficient to classify the entire scan as positive. For the present study, scan classification as positive or negative was determined in 2 ways: from the majority decision of the 5 readers and from the decision of the least accurate of the 5 readers (Reader 5).

**Neuritic Plaque Density Quantification**

At autopsy, brains from subjects imaged with florbetapir were processed with methods previously described (12). Brains were fixed whole in 10% neutral-buffered formalin for 2 weeks before dissection. One set of tissue blocks was taken from the same cortical regions of interest as were used for imaging. These blocks were processed and embedded in paraffin and stained with a modified Bielschowsky silver method, as recommended by CERAD for the estimation of neuritic plaque density (14). Neuritic plaque density scores were obtained from these sections by assigning values (none, sparse, moderate, and frequent) according to published CERAD templates (14). Identical methods are used by neuropathologists contributing to the NACC.

**Analysis Strategy**

The goal of this analysis was, first, to ascertain the fraction of florbetapir study subjects at each of the 4 histologically defined CERAD neuritic plaque densities that were classified as positive by florbetapir imaging (12). As mentioned previously, 2 estimates were used: 1 for the majority read of the 5 readers.
used in the original study and 1 for the least accurate of the 5 individual readers (Reader 5 of the original study). These estimates were then applied to NACC data to indicate the fraction of demented subjects who would be correctly identified by a positive florbetapir scan as having neuropathologically confirmed ADD, with the latter defined as subjects having moderate or frequent cortical CERAD neuritic plaque densities and Braak Neurofibrillary Stages III to VI, roughly equivalent to NIA Reagan “intermediate” or “high” levels of probability (4). In addition, the fractions of subjects detected by florbetapir imaging at each of the defined CERAD neuritic plaque densities were used to estimate the fraction of cognitively normal subjects with at least sparse neuritic plaques that would have been identified by florbetapir imaging. For this calculation, only the majority-read data were used. Sensitivity and specificity were calculated with no adjustments made for age, sex, or other subject characteristics. Groups were compared with analysis of variance and Kruskal-Wallis analysis of variance. For all tests, the significance level was set at p < 0.05.

RESULTS

Florbetapir and NACC Subject Characteristics

Of 59 florbetapir study subjects autopsied after imaging, 46 had an autopsy within 12 months (mean, 3.8 months) of their florbetapir scan, whereas the other 13 died with a mean elapsed scan-death duration of 16.3 months (Table 1). Forty-three subjects were demented, and 16 were nondemented. For NACC subjects, the mean interval between the last clinical assessment and death was 10.8 months. The subgroups of subjects (florbetapir and NACC, demented and nondemented) differed significantly in age, but all subgroups were predominantly elderly, with mean ages ranging between 76.9 and 86.4 years. For both subject sets, demented subjects had significantly greater neuritic plaque densities and Braak neurofibrillary stages than nondemented subjects, but there were no significant differences in these measures between the florbetapir group and the NACC group.

Ability of Florbetapir to Detect Defined CERAD Neuritic Plaque Densities

From the majority reads published by Clark et al (12), florbetapir scans were classified as negative in all 15 subjects found after death to have no cortical neuritic plaques and in all 5 subjects with sparse neuritic plaque densities. For subjects with moderate densities of neuritic plaques at autopsy, 10 of 13 (77%) were classified as florbetapir-positive during life, whereas all 26 subjects with frequent neuritic plaques were florbetapir-positive. For the least accurate of the individual readers (Reader 5), 13 of 15 (86.7%) subjects with no postmortem cortical neuritic plaques were classified as negative; all 5 subjects with sparse neuritic plaques were classified as negative. Among those with moderate neuritic plaques, 9 of 13 (69.2%) were classified as positive, whereas for those with frequent neuritic plaques, 18 of 26 (69.2%) were classified as positive.

Sensitivity and Specificity of Florbetapir Imaging for Neuropathologically Defined ADD

These calculations are based on the use of florbetapir scan as “test” for neuropathologically confirmed ADD, with the “gold standard” being the combination of moderate or frequent CERAD neuritic plaques and Braak Neurofibrillary Stages III to VI, as determined at autopsy. Because the subjects who underwent florbetapir scan are an entirely different set from those who underwent autopsy, the results can only be an approximate and theoretical estimate. The fraction of NACC subjects from the study by Beach et al (2) who would be correctly diagnosed as having ADD by a florbetapir-positive scan was calculated based on the fraction of subjects with moderate or frequent CERAD neuritic plaque density who would be expected to be florbetapir-positive based on the data of Clark et al (12) and on the fraction of these who would also have Braak Neurofibrillary Stages III to VI based on the data of Beach et al (2). The contingency tables used for the calculation of sensitivity and specificity are shown in Table 2. For the majority read, sensitivity was thus calculated to be 95.1% and specificity was calculated to be 89.4%. For the least accurate reader, sensitivity was 69.1% and specificity was 83.0%.

Subjects With Positive Florbetapir Scan But Without Neuropathologically Defined ADD

From the majority-read florbetapir detection rates calculated for differing CERAD neuritic plaque densities (12) and NACC data (2), among 520 subjects expected to be positive on florbetapir imaging, 32 (5.2%) would be expected not to meet the neuropathologic definition of ADD, as defined for this study (Table 2), for having a Braak neurofibrillary stage lower than III. Such subjects would be “false-positives” if a florbetapir scan were used as the sole criterion for the diagnosis of ADD. The diagnostic composition of such subjects can be inferred from the total number of subjects in the NACC demented subject set (n = 919) who have moderate or frequent CERAD neuritic plaque densities but Braak neurofibrillary stage lower than III. Thirty-eight of

<table>
<thead>
<tr>
<th>Clinical Diagnosis</th>
<th>Age, Mean ± SD, y</th>
<th>Sex (Male/Female)</th>
<th>Neuritic Plaque Density, Median (Range)</th>
<th>Braak Neurofibrillary Stage, Median (Range)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amyloid imaging subjects</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Demented (n = 43)</td>
<td>81.2 ± 11.1</td>
<td>34/25</td>
<td>3 (0–3)</td>
<td>5 (2–6)</td>
</tr>
<tr>
<td>Nondemented (n = 16)</td>
<td>76.9 ± 16.0</td>
<td>11/5</td>
<td>0 (0–3)</td>
<td>2 (0–4)</td>
</tr>
<tr>
<td>NACC subjects</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Demented (n = 919)</td>
<td>79.0 ± 11.4</td>
<td>551/368</td>
<td>3 (0–3)</td>
<td>5 (0–6)</td>
</tr>
<tr>
<td>Nondemented (n = 144)</td>
<td>86.4 ± 8.8</td>
<td>65/79</td>
<td>1 (0–3)</td>
<td>2 (0–6)</td>
</tr>
</tbody>
</table>

TABLE 1. Amyloid Imaging and NACC Patient Characteristics
Based on the results of 2 studies (2, 12), we have calculated the theoretical sensitivity and specificity of a positive brain β-amyloid imaging scan with florbetapir (18F), when used as the sole diagnostic criterion for the presence of neuropathologically defined ADD. We have also calculated the theoretical fraction of nondemented subjects with cortical β-amyloid that would be detected with a florbetapir scan. For the definition of neuropathologically confirmed ADD,

TABLE 3. Primary Neuropathologic Diagnoses for 38 NACC Subjects With Moderate or Frequent Neuritic Plaque Densities But Did Not Meet the Neuropathologic Definition of AD

<table>
<thead>
<tr>
<th>Primary Neuropathologic Findings</th>
<th>No. Subjects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary neuropathologic diagnosis of AD despite a low level of AD histopathology</td>
<td>9</td>
</tr>
<tr>
<td>FTLD</td>
<td>12</td>
</tr>
<tr>
<td>Lewy body disease</td>
<td>11</td>
</tr>
<tr>
<td>Hippocampal sclerosis</td>
<td>1</td>
</tr>
<tr>
<td>Progressive supranuclear palsy</td>
<td>1</td>
</tr>
<tr>
<td>AD pathology present but insufficient for diagnosis</td>
<td>2</td>
</tr>
<tr>
<td>“Extreme depopulation of substantia nigra”</td>
<td>1</td>
</tr>
<tr>
<td>Normal brain</td>
<td>1</td>
</tr>
</tbody>
</table>

Subjects had moderate or frequent CERAD neuritic plaque densities but did not meet the neuropathologic definition of AD because of a Braak neurofibrillary stage lower than III (based on majority-read data).
we used the combination of moderate or frequent CERAD neuritic plaque density and Braak Neurofibrillary Stages III to VI. This is equivalent to an intermediate or high probability under the NIA Reagan classification system (4), which was generally accepted as the neuropathologic “gold standard” for the diagnosis of dementia caused by AD between 1997 and 2012. Updated neuropathologic criteria were published in 2012 under the sponsorship of the NIA and the Alzheimers’ Association (NIA Alzheimers’ Association criteria) (15), but we were not able to apply these because the NACC data used in the present study were collected between 2005 and 2010, before the NIA Alzheimers’ Association criteria were published. Furthermore, the contributing NIA ADCs were not using the Thal-amyloid phasing system (16), which is required as part of the new criteria. To meet their respective intermediate or high levels, both the NIA Reagan classification system and the NIA Alzheimers’ Association criteria stipulate a Braak neurofibrillary stage higher than II, however. Therefore, the case proportion determined in the present study to be “false-positive” on florbetapir imaging—because of the presence of moderate or frequent CERAD neuritic plaque densities but a Braak neurofibrillary stage lower than III—would be the same under both sets of criteria.

Neuritic plaques were first defined based on the ultrastructural criteria developed by Wisniewski and Terry (17), but our use of the term is necessarily that of Mirra et al (14), as both of the publications on which the present study is based used the CERAD definition of neuritic plaques. Although not informative with respect to molecular composition, the CERAD definition of neuritic plaques as particular morphologic entities has the great advantage of being sufficiently unambiguous to result in relatively high interobserver agreement (18, 19); moreover, a wealth of studies have assessed its clinicopathologic significance. The adoption of the CERAD definition of neuritic plaques by the NIA ADCs enabled the accumulation of statistically large autopsied subject numbers who had all been assessed with relatively equivalent methods. This led to the confirmed realization that neuritic plaques have a significant association with dementia and cognitive impairment, whereas diffuse plaques generally do not (20). The revised NIA Alzheimers’ Association recommendations for the neuropathologic assessment of AD include the Thal-Braak staging of β-amyloid plaques (defined as any plaque type that is immunoreactive to β-amyloid) as 1 arm of the tripartite histologic assessment (15, 21); the accumulation of sufficient postmortem data will allow for a test of whether brain regional plaque distribution is also an important predictor of clinical status. At the present time, however, immunohistochemical staining for β-amyloid has been judged to be poorly suited for the subtyping of plaques or for the grading of plaque density (22). The 2 publications from the Avid study included analyses of quantitative florbetapir imaging compared with quantitative β-amyloid immunohistochemistry and Bielschowsky silver method—obtained neuritic plaque density estimates and found both staining methods to have strong and significant correlations with each other and with florbetapir uptake (12, 13).

The diagnostic capability of a positive florbetapir β-amyloid PET scan for the presence of neuropathologically confirmed ADD (as defined in this study) is estimated here at between 69% and 95% sensitivity and between 83% and 89% specificity. The scans classified by a majority read of a panel of 5 readers may represent an approximation of the upper limit of attainable accuracy, whereas the single read by the least accurate reader may be a reasonable approximation of the lower limits of accuracy. Based on data from the single most accurate reader (Reader 4), the calculated sensitivity and specificity were equivalent to the majority read (calculations not shown). Either set of figures is, however, a significant improvement from the 71% sensitivity and 71% specificity obtained in the setting of academic cognitive neurology practices, using clinical examination and standard diagnostic modalities (but excluding β-amyloid imaging and other experimental biomarker studies) such as those in the NIA ADCs (2). The use of florbetapir imaging or other comparable β-amyloid imaging agents would seem, therefore, to improve diagnostic accuracy significantly and hence subject selection for ADD clinical trials, thereby allowing smaller subject numbers and lower trial costs. A more accurate diagnosis would also limit undesirable clinical consequences that may result from the misclassification of demented subjects (23).

It is important to recognize, however, that the sensitivity and specificity figures calculated here are not generalizable to all settings or populations. We specifically limited our study to subjects who had already been clinically diagnosed as having dementia. The florbetapir study had a relatively small sample size, and this may have affected the accuracy of the estimate of the ability of florbetapir to detect defined neuritic plaque density levels. Both the subjects in the Avid study and those in the NACC database are mostly derived from specialized dementia clinics, and it is possible that even the control subjects may have been enriched in individuals at higher risk for AD. Both sets of subjects were predominantly composed of those with end-stage dementia and thus are not representative of the usual community population of elderly subjects or the usual dementia clinic population. It is likely that the predictive power of a positive florbetapir scan for ADD would have been much lower if the setting had been a random sample of both nondemented and demented elderly subjects from a community-based cohort owing to the presence of significant cortical β-amyloid in many nondemented subjects (24). Studies will need to be performed in more socially and economically diverse populations to determine whether this might be true. Nonetheless, the present data suggest that, in the setting of an established dementia, the sensitivity and specificity of a florbetapir scan for the presence of clinically significant AD are likely to be significantly more accurate than what it would be using standard neurologic practices alone. This has important implications for both clinical trials and medical practice.

Our calculations also show that florbetapir imaging may detect 56% of nondemented subjects with cortical neuritic plaques. This is encouraging but demonstrates that the earliest stages of what might be termed “preclinical AD” (when there may be only sparse densities of cortical neuritic plaques) are still expected to be mostly or entirely florbetapir-negative, at least when using equivalent scan analysis methods as were used in the florbetapir trial. Only more advanced preclinical AD subjects—predominantly those who harbor moderate or frequent neuritic plaques—are likely to
be identified. It is possible, however, that if the objective was to identify sparse densities of cortical neuritic plaques, other more sensitive ligands or other more sensitive qualitative or quantitative scan analysis methods could be developed to do this. In the florbetapir study, it is also possible that the necessary interval between scanning and death resulted in an underestimate of the ability of florbetapir to detect sparse plaques owing to the progression of β-amyloid deposition during this interval. A longer scan-death interval had, however, only a slight effect on florbetapir sensitivity, as previously reported (12).

Although β-amyloid imaging is thus expected to increase clinical diagnostic accuracy for the presence of ADD, much heterogeneity within this diagnosis is still generally undetectable during life and might be expected to affect clinical trial response rates (25). At present, there are no imaging agents for neurofibrillary tangles or pathologic tau protein aggregates; therefore, we cannot place any particular living ADD subject into their Braak neurofibrillary stage. It is likely that subjects in Braak Neurofibrillary Stage III or IV, when tangles are limited to the limbic system, will be easier to treat than subjects in Braak Neurofibrillary Stage V or VI, when tangles have spread throughout the neocortex. In many or even most subjects with ADD, 1 or more additional major neuropathologic diagnoses are present, such as vascular dementia, dementia with Lewy bodies, hippocampal sclerosis dementia, and progressive supranuclear palsy. Nevertheless, the increased diagnostic accuracy offered by florbetapir imaging is likely to have a beneficial impact on the efficiency of ADD clinical trials. Increased diagnostic accuracy might also result in improved treatment of AD and non-AD dementias (23).

An important consideration is that the US Food and Drug Administration has specifically directed that florbetapir be used only as a diagnostic aid along with a standard clinical diagnostic work-up, rather than as the sole diagnostic agent for ADD. These results are not intended as a recommendation that florbetapir scanning be used as the sole diagnostic test for ADD. Nevertheless, owing to the potentially large improvements in diagnostic accuracy and clinical trial efficacy that might be obtained with florbetapir and similarly effective β-amyloid imaging agents, we felt it important to explore this potential in the current study.

ACKNOWLEDGMENTS

We are thankful to Michael Pontecorvo and Daniel Skovronsky of Avid Radiopharmaceuticals for helpful advice and criticism.

REFERENCES


© 2014 American Association of Neuropathologists, Inc.