## Original Investigation

# Phase 3 Trial of Flutemetamol Labeled With Radioactive Fluorine 18 Imaging and Neuritic Plaque Density

Craig Curtis, MD; Jose E. Gamez, MD; Upinder Singh, MD; Carl H. Sadowsky, MD; Teresa Villena, MD;
Marwan N. Sabbagh, MD; Thomas G. Beach, MD, PhD; Ranjan Duara, MD; Adam S. Fleisher, MD, MAS;
Kirk A. Frey, MD, PhD; Zuzana Walker, MD; Arvinder Hunjan, MRCPysch; Clive Holmes, MRCPsych, PhD;
Yavir M. Escovar, MD; Carla X. Vera, CCRC; Marc E. Agronin, MD; Joel Ross, MD, CMD, CPI; Andrea Bozoki, MD;
Mary Akinola, MB, BS; Jiong Shi, MD, PhD; Rik Vandenberghe, MD, PhD; Milos D. Ikonomovic, MD;
Paul F. Sherwin, MD, PhD; Igor D. Grachev, MD, PhD; Gillian Farrar, PhD; Adrian P. L. Smith, BSc, PhD;
Christopher J. Buckley, BSc, PhD; Richard McLain, MS; Stephen Salloway, MD, MS

**IMPORTANCE** In vivo imaging of brain β-amyloid, a hallmark of Alzheimer disease, may assist in the clinical assessment of suspected Alzheimer disease.

**OBJECTIVE** To determine the sensitivity and specificity of positron emission tomography imaging with flutemetamol injection labeled with radioactive fluorine 18 to detect β-amyloid in the brain using neuropathologically determined neuritic plaque levels as the standard of truth.

**DESIGN, SETTING, AND PARTICIPANTS** Open-label multicenter imaging study that took place at dementia clinics, memory centers, and hospice centers in the United States and England from June 22, 2010, to November 23, 2011. Participants included terminally ill patients who were 55 years or older with a life expectancy of less than 1 year.

**INTERVENTIONS** Flutemetamol injection labeled with radioactive fluorine 18 (Vizamyl; GE Healthcare) administration followed by positron emission tomography imaging and subsequent brain donation.

**MAIN OUTCOMES AND MEASURES** Sensitivity and specificity of flutemetamol injection labeled with radioactive fluorine 18 positron emission tomography imaging for brain β-amyloid. Images were reviewed without and with computed tomography scans and classified as positive or negative for β-amyloid by 5 readers who were blind to patient information. In patients who died, neuropathologically determined neuritic plaque levels were used to confirm scan interpretations and determine sensitivity and specificity.

**RESULTS** Of 176 patients with evaluable images, 68 patients (38%) died during the study, were autopsied, and had neuritic plaque levels determined; 25 brains (37%) were β-amyloid negative; and 43 brains (63%) were β-amyloid positive. Imaging was performed a mean of 3.5 months (range, 0 to 13 months) before death. Sensitivity without computed tomography was 81% to 93% (median, 88%). Median specificity was 88%, with 4 of 5 of the readers having specificity greater than 80%. When scans were interpreted with computed tomography images, sensitivity and specificity improved for most readers but the differences were not significant. The area under the receiver operating curve was 0.90. There were no clinically meaningful findings in safety parameters.

**CONCLUSIONS AND RELEVANCE** This study showed that flutemetamol injection labeled with radioactive fluorine 18 was safe and had high sensitivity and specificity in an end-of-life population. In vivo detection of brain β-amyloid plaque density may increase diagnostic accuracy in cognitively impaired patients.

**Author Affiliations:** Author affiliations are listed at the end of this article.

**Corresponding Author:** Stephen Salloway, MD, MS, Department of Neurology and the Memory and Aging Program, Butler Hospital, Department of Neurology and Psychiatry, The Warren Alpert Medical School of Brown University, 345 Blackstone Blvd, Providence, RI 02906 (ssalloway@butler.org).

The neuropathological hallmarks of Alzheimer disease (AD), β-amyloid (Aβ) plaques and neurofibrillary tangles have long been used to definitively diagnose the disease post mortem[1] but the inability to detect these hallmarks in life renders clinical diagnosis of AD imperfect. For example, Beach et al[2] reported 71% to 87% sensitivity and 44% to 71% specificity for clinical diagnosis of AD based on 919 cases in the National Alzheimer Coordinating Center database and Salloway et al[3] found that among patients diagnosed as having probable AD, 6.5% of apolipoprotein E ε4 carriers and 36.1% of ε4 noncarriers were amyloid negative by positron emission tomography (PET) amyloid imaging. Because Aβ is characteristic of AD pathology, its detection in vivo may support a clinical diagnosis of AD while its absence would make AD unlikely.[4-7] Historically, Aβ plaques have been detected post mortem using stains with high affinity for amyloid aggregates (eg, thioflavin S/T, Congo red) or dystrophic neurites in neuritic plaques (Bielschowsky silver stain). Recently, PET imaging agents (eg, Pittsburgh compound B [PiB], flutemetamol, florbetapir, florbetaben, and AZD4694) were developed to detect cerebral Aβ in vivo[8-13] and 3 drugs (florbetapir[14] [amyvid; Eli Lilly], flutemetamol [vizamyl; GE Healthcare], and florbetaben [neuraceq; Piramal]) are now approved in the United States and Europe.

Pittsburgh compound B and flutemetamol, which are structurally similar to thioflavin T and to one another, bind to Aβ with high affinity[15,16] and clear rapidly from normal brain tissue.[17] Prior flutemetamol studies showed good brain uptake and radiation dosimetry similar to other radiopharmaceuticals in clinical use,[17] test-retest variability for image quantitation (standard uptake value ratio, 1%-4%),[11] differentiation between healthy participants and patients with AD,[11,17] and the ability to detect brain Aβ.[18-20]

In this study, we compared interpretations of [18F]flutemetamol PET images with brain Aβ levels (determined post mortem as the standard of truth) to determine PET sensitivity and specificity.

## Methods

### Participants and Study Design

This was a phase 3 multicenter PET study of flutemetamol injection labeled with radioactive fluorine 18 ([18F]flutemetamol) for detecting brain Aβ. Institutional review boards or ethics committees at the following institutions approved the study protocol before initiation: Compass Research, Galiz Research, Las Vegas Radiology, Premier Research Institute, Banner Sun Health Research Institute, Mt Sinai Medical Center, Wien Center for Alzheimer's Disease, Warren Alpert Medical School of Brown University, Banner Alzheimer's Institute, University of Michigan, Moorgreen Hospital, VERITAS Research, St Margaret's Hospital, Miami Jewish Health Systems, Memory Enhancement Center, Oxford Radcliffe Hospitals, Michigan State University, Exodon LLC, and Barrows Neurological Institute. All participants or their legal representatives provided prior written informed consent. Consecutive eligible participants were 55 years or older, terminally ill with a life expectancy of 1 year or less, and had gen-eral health adequate to undergo study procedures. Participants were ineligible if they were pregnant/lactating, had known/suspected structural brain abnormalities, contraindication(s) for PET, known/suspected hypersensitivity/allergy to [18F]flutemetamol injection (or any component), or had participated in any clinical study using an investigational product within 30 days of signing consent.

The primary study objective was to determine the sensitivity of blinded visual interpretations of [18F]flutemetamol PET images alone for detecting brain Aβ. Secondary objectives included specificity, and sensitivity and specificity when computed tomography (CT) images were available during PET image review.

### Procedures

Consenting patients were enrolled at centers in the United States and England. For statistical analyses, patients were classified as having an entry diagnosis of AD, another dementing disorder, mild cognitive impairment, memory loss, or no cognitive impairment at screening based on a review of reported medical history data. The Mini-Mental State Examination was performed at screening.[21] Before PET imaging, participants underwent head CT or magnetic resonance imaging unless prior images (obtained within 12 months) were available. Flutemetamol injection labeled with fluorine 18 was administered intravenously at a dose of 185 to 370 MBq of radioactivity at physician discretion based on how long the patient could lie in the scanner. Positron emission tomography images were acquired in 2-minute frames on PET/CT cameras, beginning approximately 90 minutes postinjection. Five frames were summed to give a 10-minute scan, which was attenuation-corrected using CT data. Most images were reconstructed iteratively to form 128 × 128 axial slices and a gaussian postreconstruction smoothing filter was applied to some. Imaging parameters had been developed in a prior phase 2 study.[11] Safety was monitored from injection to 4 hours after scanning and 24 hours after injection (by telephone).

For the blinded read, PET images were randomized and approximately 10% of these images were duplicated and randomly combined with the other images to measure within-reader reproducibility. Five readers (4 nuclear medicine physicians and 1 radiologist) independently interpreted PET scans at a central review center. Each reader had at least 3 years' neuroimaging experience, was trained in person by an expert nuclear medicine physician in evaluating [18F]flutemetamol PET images, and was blinded to patient information. The 5 brain regions assessed were the combined lateral frontal cortex and anterior cingulate, combined posterior cingulate and precuneus, insula, lateral temporal lobes, and striatal regions. Images were displayed in color (Sokoloff color scale)[22] scaled from 0 to maximum intensity. All images showed nonspecific flutemetamol uptake by white matter; the absence of a similar level of uptake in any gray matter area indicated a negative image. Positive images showed at least 1 gray matter area with uptake similar to or greater than that seen in the white matter and cerebellar cortex.

After readers interpreted the PET images and locked their decisions, images were rerandomized and reread with CT im-

ages for anatomic guidance. Between-reader agreement and within-reader reproducibility were determined for the PET interpretations. The interpretation of each PET image made independently by at least 3 of 5 readers was considered the majority image interpretation.

Standard of truth results were obtained from participants who died during the study and underwent brain autopsy. Postmortem histopathology was performed at a central pathology laboratory. Brain regions assessed were the precuneus, midfrontal lobe, superior temporal, middle temporal, anterior cingulate, posterior cingulate, primary visual cortex, and inferior parietal. Two blocks per region and 3 slides per block were prepared. Sections were stained with Bielschowsky silver stain.[23] Two neuropathologists (blinded to all PET data and participant clinical information) assessed 5 randomly chosen fields of view (approximately 2.5 mm$^2$) per slide simultaneously, using a 2-headed microscope, and reached consensus on neuritic plaque counts following Vemuri modification of the Consortium to Establish a Registry for Alzheimer Disease criteria: 0 (none), 1 (sparse, 1-5 plaques), 2 (moderate, 6-19 plaques), or 3 (frequent, >20 plaques).[1,24] Field of view scores were averaged to give slide scores, which were averaged to give regional scores. Each brain was classified as either negative (all regional scores ≤1.5 [the midpoint of the Consortium to Establish a Registry for Alzheimer Disease scale]) or positive (any regional score >1.5).

The blinded visual interpretations of PET images were compared with the postmortem brain Aβ results and sensitivity and specificity were determined for each reader. To obtain a neuropathological diagnosis including cerebral amyloid angiopathy, tissue was also stained by immunohistochemistry for Aβ, tau, α-synuclein, and ubiquitin. Quantitative analysis of images (standard uptake value ratio) was performed and will be reported separately.

### Statistical Analysis

Populations analyzed included the safety cohort (participants receiving any flutemetamol dose), full cohort (participants with usable PET results), and postmortem cohort (participants with postmortem Aβ results). Sensitivity and specificity were determined as point estimates with exact 2-sided 95% CIs. Between-reader agreement and within-reader reproducibility were determined as percentage agreement and κ coefficient. All data were reported and no imputation for missing data was done. Areas under receiver operating characteristic (ROC) curves were determined by calculating the area of trapezoids defined by the curves.

The null hypothesis related to the precision of sensitivity estimation, as indicated by the width of the 95% CI; the null hypothesis was that the lower bound of the 95% CI for sensitivity determined for a blinded reader was 70% or less, assuming a true sensitivity of 92%. Sample size calculations showed that 31 brains that were Aβ positive post mortem would provide approximately 90% power to reject the null hypothesis. The study would be considered a success if the null hypothesis was rejected for at least 3 of the 5 readers. Based on similar calculations for specificity (assuming a true specificity of 95%), it was also planned to accrue at least 22 brains that were

Aβ negative post mortem. An independent data monitoring committee reviewed the pathology data on an ongoing basis and reported the numbers of Aβ-positive and Aβ-negative brains at regular intervals.

## Results

Between June 22, 2010, and November 23, 2011, 203 participants consented and enrolled at 19 centers (dementia clinics, memory centers, and hospice centers; 15 in the United States, 4 in England); eFigure 1 in the Supplement shows participant disposition and **Table 1** shows baseline demographics and characteristics. Most participants had existing neurologic (83%), cardiovascular (73%), digestive (64%), psychiatric (57%), and musculoskeletal (52%) conditions at baseline. For the 180 participants who were dosed and imaged, the mean (SD) dose (megabecquerel radioactivity) was 349.2 (38.358); 86% of participants received 370 ± 10% MBq, 2% received 185 ± 10% MBq, and 12% received between 185 and 370 MBq. All images were of good technical quality; none were considered unevaluable. Images from 4 participants (2.2%) were excluded from analysis because the participants did not complete the study. For these end-of-life patients, the time to remain motionless in a loud magnetic resonance imaging scanner was deemed excessive for practical and ethical reasons; hence, few magnetic resonance images were obtained and none were used because all patients had received a CT scan within 12 months.

Among participants who died, [18F]flutemetamol PET was performed a mean of 3.5 months (range, 0-13 months) before death. Of the 68 evaluable brains, 25 brains (37%) were Aβ negative and 43 (63%) were Aβ positive. Histopathology showed varying levels of neuritic plaque density from none to frequent, with as many as 60% of samples considered moderate or frequent (**Figure**).

**Table 2** shows by-reader results of the primary end point, which was sensitivity. Sensitivity was 81% to 93% (median, 88%; majority, 86%). The lower bound of the 2-sided 95% CI was more than 70% for at least 3 of the 5 readers, meeting the primary objective of the study. Specificity was 44% to 92% (median, 88%; majority, 92%). When readers interpreted PET scans with CT images, sensitivity and specificity improved for most readers but the differences were not significant (data not shown). Possible reasons for the outlying specificity value (reader 3) are discussed below. The area under the ROC curve for flutemetamol without CT (eFigure 2 in the Supplement) was 0.90 (95% CI, 0.82-0.97), comparing favorably with a theoretical value of 1.0 for a perfect test.

Pairwise between-reader agreement (**Table 3**) was 90% or higher without or with CT except for comparisons involving reader 3 (80%-88%); κ scores were also lower for reader 3. κ scores were higher with CT. Within-reader reproducibility without and with CT images was 88% to 100% and 82% to 100%, respectively, and κ ranges were 0.60 to 1.00 and 0.30 to 1.00, respectively (**Table 4**).

Based on the majority interpretations of PET, 6 of the 43 participants who were Aβ positive (14%) had false-negative (FN) interpretations and 2 of the 25 participants who were Aβ nega-

Table 1. Patient Demographics and Characteristics

| Statistic | Analysis Set | | |
|---|---|---|---|
| | Safety (n = 180) | Full (n = 176) | Postmortem (n = 68) |
| **Characteristic** | | | |
| Age, y[a] | | | |
| Sample size | 180 | 176 | 68 |
| Mean (SD) [range] | 79.9 (9.61) [47-98][b] | 80.0 (9.59) [47-98][b] | 80.8 (8.19) [60-95] |
| Sex, No. (%) | | | |
| Male | 77 (43) | 75 (43) | 33 (49) |
| Female | 103 (57) | 101 (57) | 35 (51) |
| Race/ethnicity, No. (%) | | | |
| Asian | 1 (1) | 1 (1) | 0 |
| Black | 7 (4) | 6 (3) | 2 (3) |
| White | 168 (93) | 165 (94) | 64 (94) |
| Other | 4 (2) | 4 (2) | 2 (3) |
| Race/ethnicity, No. (%) | | | |
| Non-Hispanic or Latino | 139 (77) | 135 (77) | 46 (68) |
| Hispanic or Latino | 41 (23) | 41 (23) | 22 (32) |
| MMSE | | | |
| Sample size[c] | 123 | 119 | 42 |
| Mean (SD) [range] | 8.5 (10.39) [0-30] | 8.5 (10.50) [0-30] | 9.8 (10.47) [0-30] |
| Clinical diagnosis, No. (%)[d] | | | |
| Alzheimer disease | 87 (48) | 84 (48) | 30 (44) |
| Other dementing disorder | 48 (27) | 47 (25) | 17 (25) |
| Parkinson disease with dementia | 1 | 1 | 1 |
| Frontotemporal dementia | 2 | 2 | 1 |
| Dementia with Lewy bodies | 4 | 4 | 1 |
| Vascular dementia | 1 | 1 | 1 |
| Normal pressure hydrocephalus | 2 | 2 | 2 |
| Unspecified | 38 | 37 | 11 |
| Mild cognitive impairment | 0 | 0 | 0 |
| No history of cognitive impairment | 44 (24) | 44 (25) | 21 (31) |
| Memory loss (unspecified) | 1 (1) | 1 (1) | 0 |

Abbreviation: MMSE, Mini-Mental State Examination.

[a] Age was calculated as (date of informed consent minus date of birth)/365.25 rounded down to the nearest integer.

[b] One patient with end-stage renal disease and a life expectancy of less than 6 months was given a waiver to enroll in the study because the age difference was not expected to affect subject safety or efficacy evaluation.

[c] Number of patients with nonmissing data for MMSE.

[d] Clinical diagnosis was determined based on reported medical history.

tive (8%) had false-positive (FP) interpretations. eTable 1 in the Supplement shows characteristics for participants with FN, FP, and correct image interpretations. eFigure 3 in the Supplement shows the PET and PET/CT images for the FP and FN cases that were not borderline in pathology (Figure), as well as examples of unequivocal negative and positive cases. Summaries of the pathology of these nonborderline cases are included. Although caution should be exercised in interpreting the data based on small sample sizes, the FN cases were more likely to have no history of cognitive impairment and the FP cases were more likely to have a history of a non-Alzheimer form of dementia. All of the FN and FP cases had sparse or moderate neuritic plaque densities and were more likely to have a low likelihood of AD according to the National Institute of Aging–Reagan Institute criteria.[25]

eTable 2 in the Supplement lists reported adverse events. Flushing (n = 2), considered possibly related to flutemetamol, was mild and did not require treatment. Deaths were expected based on the main inclusion criterion (life expectancy, ≤1 year). Of the 180 dosed participants, 69 (38%) died

during the study; 2 deaths occurred within the 24-hour follow-up period and were reported as the outcome of serious adverse events (prostate cancer and senile dementia). No death was considered related to flutemetamol. Nonfatal serious adverse events (considered unrelated to study drug) were reported for 2 participants, a case of severe anemia and a case of change in mental status caused by a combination of AD progression and *Clostridium difficile* infection. All other adverse events were mild or moderate. No clinically important trends or issues were noted in any safety parameter, including clinical laboratory values, vital signs, and electrocardiograms.

## Discussion

In this study, blinded interpretation of [18F]flutemetamol PET images allowed in vivo detection of Aβ in the brain with high sensitivity, specificity, between-reader agreement, and within-reader reproducibility. Flutemetamol injection labeled with ra-

Figure. Standard of Truth (SoT) Results for Each Participant Sorted by Mean Bielschowsky Score
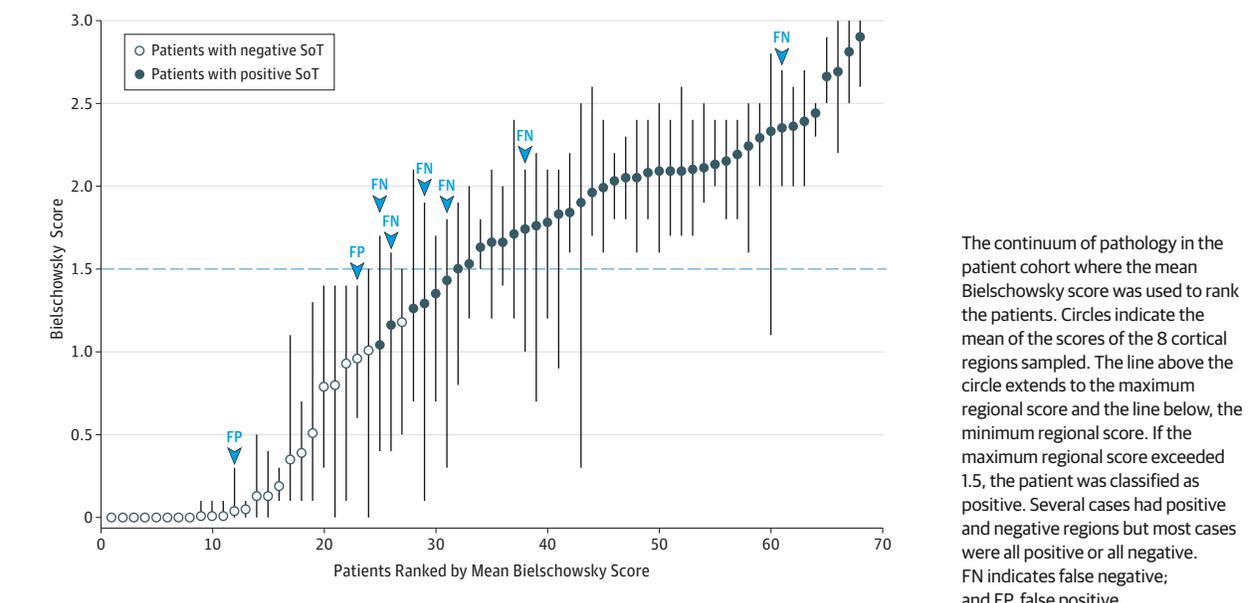


The continuum of pathology in the patient cohort where the mean Bielschowsky score was used to rank the patients. Circles indicate the mean of the scores of the 8 cortical regions sampled. The line above the circle extends to the maximum regional score and the line below, the minimum regional score. If the maximum regional score exceeded 1.5, the patient was classified as positive. Several cases had positive and negative regions but most cases were all positive or all negative. FN indicates false negative; and FP, false positive.

dioactive fluorine 18 was well tolerated in this population of terminally ill patients.

Despite a training program conducted in person by an experienced nuclear medicine physician, the specificity of one of the readers (reader 3) was lower compared with the other 4 readers. The lower specificity was owing to a high FP rate, indicating over-calling for amyloid. Subsequent to this study, an electronic reader training program was designed using more equivocal cases and was made available to image readers.[26] This program requires trainees to correctly interpret 14 of 15 cases to pass the training and we believe it will result in greater reader accuracy.

By the majority image interpretation, 8 incorrect results (6 FN cases and 2 FP cases) occurred. No FN or FP case was attributable to poor image quality. Detailed assessment of these cases (eTable 1 in the Supplement) showed the most striking common denominator to be a Consortium to Establish a Registry for Alzheimer Disease neuritic plaque density rating of sparse or moderate. In 4 FN cases, the neuritic plaque score was close to the threshold, suggesting these were borderline cases. Two FN cases had neuropathological evidence consistent with dementia with Lewy bodies. Two FPs occurred in patients who had a neuropathological diagnosis of dementia with Lewy bodies and had diffuse plaques as well as sparse or moderate neuritic plaques that were below the threshold for standard of truth. Patients with this presentation would have intermediate likelihood of AD using the more recent criteria.[27] These findings are consistent with results obtained with PiB in patients with dementia with Lewy bodies.[28] Each FN case also had some degree of cortical atrophy and 2 FN cases became true-positive cases when anatomic images were used, possibly through better delineation of the gray and white matter. eFigure 4 in the Supplement shows the PET/CT images of the FN and FP cases that did not have borderline pathology; patient summaries are included.

Two other amyloid PET imaging agents (florbetapir and florbetaben) are approved in the United States and Europe. None

of the 3 PET agents have been compared head to head in a clinical trial and the results from independent studies should be compared cautiously owing to differences in study design and population. For example, the distribution of neuritic plaque densities for the florbetapir participants was markedly different from the distributions for flutemetamol and florbetaben (eTable 3 in the Supplement). With these caveats in mind, by-reader sensitivity and specificity were determined from data reported in the US package inserts for florbetapir[9,29] and florbetaben[30] (eTable 4 in the Supplement). Direct comparison of these data is difficult because florbetaben and flutemetamol readers tended to have higher sensitivity than florbetapir readers, who tended to have higher specificity. To facilitate comparison, we constructed ROC curves for the readers (data not shown). The area under the ROC curve provides a single measure of diagnostic test performance; the area for a perfect test is 1 and the area for a useful test is greater than 0.5.[31] The areas under the ROC curves for flutemetamol (0.90), florbetaben (0.90), and florbetapir (0.85) were not significantly different (2-tailed, $P = .34$-.92).

Our study had some limitations. The number of readers (5), although typical for a clinical trial, may not represent the population of nuclear medicine physicians who will read PET amyloid scans. Majority interpretations provided a single number but did not simulate clinical practice where images are read by a single reader. Quantitative measurement of flutemetamol uptake, although not included in the primary outcome for this trial, may add important information to the interpretation of amyloid PET scans. Patients' cognitive capacity was recorded based on medical history and Mini-Mental State Examination; no further neurobehavioral testing was performed. However, it is unlikely that the results would have affected the correlation between imaging and pathology, which was the focus of the article. The variable, sometimes long, time between imaging and autopsy may have contributed to lack of agree-

ment between PET imaging and histopathology results in certain cases. The study's end-of-life population (required to allow autopsy) may not completely reflect the patient population expected to undergo amyloid PET imaging in clinical practice. For example, no participants with mild cognitive impairment, a population who may benefit from amyloid PET, were

included in this sample. Results of a recent study showed that patients with amnestic mild cognitive impairment who had positive [¹⁸F]flutemetamol scans were 2.6 times more likely to progress to probable AD than those with negative scans.[32] Further research is needed to elucidate the utility of amyloid PET in mild cognitive impairment. Finally, reader training in this study was conducted in person, whereas refined electronic versions have been developed.

Detecting amyloid is not synonymous with diagnosing AD. Amyloid PET should only be ordered in the appropriate clinical context, eg, after completion of a detailed evaluation by a dementia expert results in diagnostic uncertainty, AD is a serious consideration, and the result of the amyloid PET scan is likely to alter patient management.[6,7]

## Conclusions

In vivo detection of brain Aβ may help increase diagnostic accuracy in cognitively impaired patients compared with clinical diagnosis alone. This study showed that [¹⁸F]flutemetamol injection was safe and had high sensitivity and specificity in an end-of-life population with a broad and continuous range of Aβ levels in the brain.

**Table 2. Sensitivity and Specificity of Blinded Visual Interpretation of Flutemetamol Images[a]**

| | % (95% CI) | |
|---|---|---|
| Reader | Sensitivity[b] | Specificity[c] |
| 1 | 81 (67-92) | 88 (69-98) |
| 2 | 88 (74-96)[d] | 92 (74-99) |
| 3 | 93 (81-99)[d] | 44 (24-65) |
| 4 | 93 (81-99)[d] | 80 (59-93) |
| 5 | 88 (75-96)[d] | 92 (74-99) |
| Median | 88 (74-96) | 88 (69-98) |
| Majority | 86 (72-95) | 92 (74-99) |

[a] Data are rounded to the nearest integer.

[b] Sensitivity = true positives/(true positives + false negatives).

[c] Specificity = true negatives/(true negatives + false positives).

[d] Two-sided 95% CI with a lower bound greater than 70%.

**Table 3. Summary of Between-Reader Agreement for Blinded Visual Interpretations Without and With Anatomic Images**

| | Anatomic Images | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Without | | | | With | | | |
| Reader Pair | Sample Size[a] | Agreement, No. (%) | κ[b] | 95% CI | Sample Size[a] | Agreement, No. (%) | κ[b] | 95% CI |
| 1 vs 2 | 175 | 159 (91) | 0.79 | 0.69-0.89 | 175 | 166 (95) | 0.87 | 0.79-0.95 |
| 1 vs 3 | 176 | 140 (80) | 0.47 | 0.34-0.61 | 176 | 148 (84) | 0.56 | 0.43-0.70 |
| 1 vs 4 | 175 | 164 (94) | 0.85 | 0.77-0.94 | 176 | 171 (97) | 0.93 | 0.87-0.99 |
| 1 vs 5 | 176 | 169 (96) | 0.91 | 0.84-0.98 | 175 | 173 (99) | 0.97 | 0.94-1.00 |
| 2 vs 3 | 175 | 148 (85) | 0.55 | 0.41-0.69 | 175 | 154 (88) | 0.64 | 0.50-0.77 |
| 2 vs 4 | 175 | 158 (90) | 0.76 | 0.65-0.87 | 175 | 165 (94) | 0.86 | 0.77-0.94 |
| 2 vs 5 | 175 | 163 (93) | 0.84 | 0.75-0.93 | 174 | 167 (96) | 0.90 | 0.83-0.97 |
| 3 vs 4 | 175 | 141 (81) | 0.44 | 0.29-0.59 | 176 | 149 (85) | 0.56 | 0.43-0.70 |
| 3 vs 5 | 176 | 145 (82) | 0.52 | 0.39-0.66 | 175 | 149 (85) | 0.58 | 0.45-0.72 |
| 4 vs 5 | 175 | 168 (96) | 0.90 | 0.84-0.97 | 175 | 172 (98) | 0.96 | 0.91-1.00 |
| Readers 1-5 | 175 | 131 (75) | 0.72 | 0.67-0.76 | 174 | 143 (82) | 0.79 | 0.75-0.84 |

[a] Number of images with nonmissing values for the respective reader pair.

[b] Cohen κ coefficient reported for between-reader agreement. Fleiss κ coefficient reported for multiple reader agreement.

**Table 4. Summary of Within-Reader Reproducibility for Blinded Visual Interpretations Without and With Anatomic Images**

| | Anatomic Images | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Without | | | | With | | | |
| Reader | Sample Size[a] | Agreement, No. (%) | κ[b] | 95% CI | Sample Size[a] | Agreement to No. (%) | κ[b] | 95% CI |
| 1 | 17 | 16 (94) | 0.85 | (0.57 to 1.00) | 17 | 16 (94) | 0.85 | (0.57 to 1.00) |
| 2 | 17 | 17 (100) | 1.00 | NE | 17 | 17 (100) | 1.00 | NE |
| 3 | 17 | 15 (88) | 0.60 | (0.09 to 1.00) | 17 | 14 (82) | 0.30 | (−0.29 to 0.90) |
| 4 | 17 | 17 (100) | 1.00 | NE | 17 | 16 (94) | 0.85 | (0.57 to 1.00) |
| 5 | 17 | 17 (100) | 1.00 | NE | 17 | 17 (100) | 1.00 | NE |

Abbreviation: NE, not estimable.

[a] Number of images selected for reread.

[b] Cohen κ coefficient reported for within-reader agreement.

## REFERENCES

1. Mirra SS, Heyman A, McKeel D, et al. The Consortium to Establish a Registry for Alzheimer's Disease (CERAD), part II: standardization of the neuropathologic assessment of Alzheimer's disease. *Neurology*. 1991;41(4):479-486.

2. Beach TG, Monsell SE, Phillips LE, Kukull W. Accuracy of the clinical diagnosis of Alzheimer disease at National Institute on Aging Alzheimer Disease Centers: 2005-2010. *J Neuropathol Exp Neurol*. 2012;71(4):266-273.

3. Salloway S, Sperling R, Fox NC, et al; Bapineuzumab 301 and 302 Clinical Trial

Investigators. Two phase 3 trials of bapineuzumab in mild-to-moderate Alzheimer's disease. *N Engl J Med*. 2014;370(4):322-333.

**4**. Schipke CG, Peters O, Heuser I, et al. Impact of beta-amyloid-specific florbetaben PET imaging on confidence in early diagnosis of Alzheimer's disease. *Dement Geriatr Cogn Disord*. 2012;33(6):416-422.

**5**. Frederiksen KS, Hasselbalch SG, Hejl AM, Law I, Højgaard L, Waldemar G. Added diagnostic value of (11)C-PiB-PET in memory clinic patients with uncertain diagnosis. *Dement Geriatr Cogn Dis Extra*. 2012;2(1):610-621.

**6**. Johnson KA, Minoshima S, Bohnen NI, et al; Alzheimer's Association; Society of Nuclear Medicine and Molecular Imaging; Amyloid Imaging Taskforce. Appropriate use criteria for amyloid PET: a report of the Amyloid Imaging Task Force, the Society of Nuclear Medicine and Molecular Imaging, and the Alzheimer's Association. *Alzheimers Dement*. 2013;9(1):e-1-e-16.

**7**. Johnson KA, Minoshima S, Bohnen NI, et al; Amyloid Imaging Task Force of the Alzheimer's Association and Society for Nuclear Medicine and Molecular Imaging. Update on appropriate use criteria for amyloid PET imaging: dementia experts, mild cognitive impairment, and education. *Alzheimers Dement*. 2013;9(4):e106-e109.

**8**. Clark CM, Schneider JA, Bedell BJ, et al; AV45-A07 Study Group. Use of florbetapir-PET for imaging beta-amyloid pathology. *JAMA*. 2011;305 (3):275-283.

**9**. Clark CM, Pontecorvo MJ, Beach TG, et al; AV-45-A16 Study Group. Cerebral PET with florbetapir compared with neuropathology at autopsy for detection of neuritic amyloid-β plaques: a prospective cohort study. *Lancet Neurol*. 2012;11 (8):669-678.

**10**. Barthel H, Gertz HJ, Dresel S, et al; Florbetaben Study Group. Cerebral amyloid-β PET with florbetaben (18F) in patients with Alzheimer's disease and healthy controls: a multicentre phase 2 diagnostic study. *Lancet Neurol*. 2011;10(5):424-435.

**11**. Vandenberghe R, Van Laere K, Ivanoiu A, et al. 18F-flutemetamol amyloid imaging in Alzheimer disease and mild cognitive impairment: a phase 2 trial. *Ann Neurol*. 2010;68(3):319-329.

**12**. Rowe CC, Pejoska S, Mulligan RS, et al. Head-to-head comparison of 11C-PiB and

18F-AZD4694 (NAV4694) for β-amyloid imaging in aging and dementia. *J Nucl Med*. 2013;54(6):880-886.

**13**. Ikonomovic MD, Klunk WE, Abrahamson EE, et al. Post-mortem correlates of in vivo PiB-PET amyloid imaging in a typical case of Alzheimer's disease. *Brain*. 2008;131(pt 6):1630-1645.

**14**. Yang L, Rieves D, Ganley C. Brain amyloid imaging: FDA approval of florbetapir F18 injection. *N Engl J Med*. 2012;367(10):885-887.

**15**. Ye L, Morgenstern JL, Gee AD, Hong G, Brown J, Lockhart A. Delineation of positron emission tomography imaging agent binding sites on beta-amyloid peptide fibrils. *J Biol Chem*. 2005; 280(25):23599-23604.

**16**. Swahn BM, Sandell J, Pyring D, et al. Synthesis and evaluation of pyridylbenzofuran, pyridylbenzothiazole and pyridylbenzoxazole derivatives as $^{18}$F-PET imaging agents for β-amyloid plaques. *Bioorg Med Chem Lett*. 2012;22(13):4332-4337.

**17**. Nelissen N, Van Laere K, Thurfjell L, et al. Phase 1 study of the Pittsburgh compound B derivative 18F-flutemetamol in healthy volunteers and patients with probable Alzheimer disease. *J Nucl Med*. 2009;50(8):1251-1259.

**18**. Leinonen V, Rinne JO, Virtanen KA, et al. Positron emission tomography with [18F]flutemetamol and [11C]PiB for in vivo detection of cerebral cortical amyloid in normal pressure hydrocephalus patients. *Eur J Neurol*. 2013;20(7):1043-1052.

**19**. Rinne JO, Wong DF, Wolk DA, et al. [(18)F]Flutemetamol PET imaging and cortical biopsy histopathology for fibrillar amyloid β detection in living subjects with normal pressure hydrocephalus: pooled analysis of four studies. *Acta Neuropathol*. 2012;124(6):833-845.

**20**. Wong DF, Moghekar AR, Rigamonti D, et al. An in vivo evaluation of cerebral cortical amyloid with [18F]flutemetamol using positron emission tomography compared with parietal biopsy samples in living normal pressure hydrocephalus patients. *Mol Imaging Biol*. 2013;15(2):230-237.

**21**. 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*. 1975;12(3):189-198.

**22**. Goochee C, Rasband W, Sokoloff L. Computerized densitometry and color coding of [14C] deoxyglucose autoradiographs. *Ann Neurol*. 1980;7(4):359-370.

**23**. Litchfield S, Nagy Z. New temperature modification makes the Bielschowsky silver stain reproducible. *Acta Neuropathol*. 2001;101(1):17-21.

**24**. Vemuri P, Whitwell JL, Kantarci K, et al. Antemortem MRI based STructural Abnormality iNDex (STAND)-scores correlate with postmortem Braak neurofibrillary tangle stage. *Neuroimage*. 2008;42(2):559-567.

**25**. Hyman BT, Trojanowski JQ. Consensus recommendations for the postmortem diagnosis of Alzheimer disease from the National Institute on Aging and the Reagan Institute Working Group on diagnostic criteria for the neuropathological assessment of Alzheimer disease. *J Neuropathol Exp Neurol*. 1997;56(10):1095-1097.

**26**. Vizamyl (flutemetamol F 18 injection) Electronic Reader Training Program. http://www .readvizamyl.com. Accessed August 4, 2014.

**27**. Hyman BT, Phelps CH, Beach TG, et al. National Institute on Aging-Alzheimer's Association guidelines for the neuropathologic assessment of Alzheimer's disease. *Alzheimers Dement*. 2012;8(1): 1-13.

**28**. Kantarci K, Lowe VJ, Boeve BF, et al. Multimodality imaging characteristics of dementia with Lewy bodies. *Neurobiol Aging*. 2012;33(9): 2091-2105.

**29**. Prescribing information for Amyvid. http://www .accessdata.fda.gov/drugsatfda_docs/label/2013 /202008s020lbl.pdf. Accessed July 21, 2014.

**30**. Prescribing information for Neuraceq. http: //www.accessdata.fda.gov/drugsatfda_docs/label /2014/204677s001lbl.pdf. Accessed July 21, 2014.

**31**. Zhou X-H, Obuchowski N, McClish D. *Statistical Methods in Diagnostic Medicine*. 2nd ed. Hoboken, NJ: Wiley; 2011.

**32**. Wolk DA, Duara R, Sadowsky CH. [18F]flutemetamol amyloid PET imaging: outcome of phase III study in subjects with amnestic mild cognitive impairment after 3-year follow up [abstract]. Paper presented at: Alzheimer's Association International Conference; 2014; Copenhagen, Denmark.