A Visual Interpretation Algorithm for Assessing Brain Tauopathy with ^{18}F-MK-6240 PET

John P. Seibyl[1,2], Jonathan M. DuBois[3], Annie Racine[3], Jessica Collins[3], Qi Guo[4], Dustin Wooten[3], Eddie Stage[4], David Cheng[2], Roger N. Gunn[2], Lilly Porat[2], Alex Whittington[2], Phillip H. Kuo[5], Masanori Ichise[6], Robert Comley[4], Laurent Martarello[3], and Cristian Salinas[3]

[1]Institute for Neurodegenerative Disorders, New Haven, Connecticut; [2]Invicro, New Haven, Connecticut; [3]Biogen, Cambridge, Massachusetts; [4]AbbVie, North Chicago, Illinois; [5]University of Arizona, Tucson, Arizona; and [6]National Institute of Radiological Sciences, National Institutes for Quantum and Radiological Science and Technology, Chiba, Japan

In vivo characterization of pathologic deposition of tau protein in the human brain by PET imaging is a promising tool in drug development trials of Alzheimer disease (AD). 6-(fluoro-^{18}F)-3-(1H-pyrrolo[2,3-c] pyridin-1-yl)isoquinolin-5-amine (^{18}F-MK-6240) is a radiotracer with high selectivity and subnanomolar affinity for neurofibrillary tangles that shows favorable nonspecific brain penetration and excellent kinetic properties. The purpose of the present investigation was to develop a visual assessment method that provides both an overall assessment of brain tauopathy and regional characterization of abnormal tau deposition. **Methods:** ^{18}F-MK-6240 scans from 102 participants (including cognitively normal volunteers and patients with AD or other neurodegenerative disorders) were reviewed by an expert nuclear medicine physician masked to each participant's diagnosis to identify common patterns of brain uptake. This initial visual read method was field-tested in a separate, nonoverlapping cohort of 102 participants, with 2 additional naïve readers trained on the method. Visual read outcomes were compared with semiquantitative assessments using volume-of-interest SUV ratio. **Results:** For the visual read, the readers assessed 8 gray-matter regions per hemisphere as negative (no abnormal uptake) or positive (1%–25% of the region involved, 25%–75% involvement, or >75% involvement) and then characterized the tau binding pattern as positive or negative for evidence of tau and, if positive, whether brain uptake was in an AD pattern. The readers demonstrated agreement 94% of the time for overall positivity or negativity. Concordance on the determination of regional binary outcomes (negative or positive) showed agreement of 74.3% and a Fleiss κ of 0.912. Using clinical diagnosis as the ground truth, the readers demonstrated a sensitivity of 73%–79% and specificity of 91%–93%, with a combined reader-concordance sensitivity of 80% and specificity of 93%. The average SUV ratio in cortical regions showed a robust correlation with visually derived ratings of regional involvement ($r = 0.73$, $P < 0.0001$). **Conclusion:** We developed a visual read algorithm for ^{18}F-MK-6240 PET offering determination of both scan positivity and the regional degree of cortical involvement. These cross-sectional results show strong interreader concordance on both binary and regional assessments of tau deposition, as well as good sensitivity and excellent specificity supporting use as a tool for clinical trials.

Key Words: neurology; PET; Alzheimer disease; MK-6240; PET; tau proteinopathies

J Nucl Med 2023; 64:444–451
DOI: 10.2967/jnumed.122.264371

Received May 11, 2022; revision accepted Sep. 21, 2022.
For correspondence or reprints, contact John P. Seibyl (jpseibyl@seibyl.com).
Published online Sep. 29, 2022.

The recent introduction of tau PET imaging biomarkers for clinical and research applications provides a powerful tool for corroborating the patterns of pathologic progression in Alzheimer disease (AD) suggested by postmortem studies (*1*), as well as potentially offering a way to monitor response to treatments designed to interrupt AD brain pathology (*2,3*). Tau PET imaging shows a good correlation between regional brain uptake and clinical and psychometric measures in cross-sectional studies (*4–7*). Indeed, increased density and spread of abnormal uptake in tau PET images is consistent with progression of the disease in patients with early or mild AD and is associated with the degree of neuropsychologic impairment.

Tau PET studies have corroborated findings from pathologic postmortem examination of AD brains, which demonstrated initial cortical uptake in the entorhinal cortex and medial temporal structures, extending to the inferolateral temporal and superolateral temporal structures and to the neocortical occipital, posterior cingulate, parietal, and frontal cortices (*8*). Temporal lobe structures, especially the mesial temporal gyri and hippocampus, are the earliest neocortical regions to manifest neurofibrillary tangles (*9*), suggesting that visual read methods might focus particular attention here. The first-generation tau PET agent ^{18}F-flortaucipir (^{18}F-AV-1451), which was used in the largest clinicopathologic study of AD and mild cognitive impairment (MCI) (*10*), has off-target uptake in areas adjacent to the mesial temporal lobe, limiting the ability to assess the tau pathology of this important region.

6-(fluoro-^{18}F)-3-(1H-pyrrolo[2,3-c]pyridin-1-yl)isoquinolin-5-amine (^{18}F-MK-6240) is a second-generation tau PET imaging agent with high specificity and low off-target binding in gray and white matter, representing improvements over first-generation tracers (*11,12*). In cognitively normal individuals, ^{18}F-MK-6240 demonstrates homogeneous uptake such that some structural features such as the ventricles are visualized but without evidence of focal uptake in the neocortex. AD patients show a pattern of cortical uptake that is more intense, asymmetric, and focal (*13*), consistent with the distribution of the tau pathology reported in postmortem studies (*9*).

As the utility of tau PET radiotracers such as ^{18}F-MK-6240 expands, potential future clinical applications in AD may include aiding in the differential diagnosis of patients with cognitive impairment, screening for eligibility for long-term treatments to slow disease progression, monitoring the effectiveness of such treatments, assessing the course of disease, and serving as a prognostic biomarker potentially identifying at-risk cohorts. For any of these indications, a robust method for visual assessment will be an important way to evaluate ^{18}F-MK-6240 PET images, especially in the clinical setting, where an interpretation of the

tau PET scan as negative or positive is a primary goal. However, unlike amyloid PET, tau uptake patterns demonstrate dynamic heterogeneity in individuals with AD, both in the spatial extent within the brain and in the intensity of uptake within regions. These patterns offer an opportunity to obtain additional information from the visual read relevant to clinical research trials. Specifically, it may be possible to visually assess changes in the extent of uptake within regions over time, as well as between regions. Thus, it may be possible to elicit tau progression information from the visual interpretation.

The goals of the present study were to develop, field-test, and refine a visual read method for ^{18}F-MK-6240 PET as a potential tool for assessing in vivo brain tau accumulation, providing a readout of tau positivity and negativity and of the spatial extent of uptake within individual regions. The latter may be relevant to assessing within-patient changes in the context of clinical therapeutic drug trials in which the tau PET signal might be expected to be unchanged or even decrease on serial imaging.

MATERIALS AND METHODS

Imaging Data

Pooled imaging data provided under informed consent from 204 participants with various diagnoses (cognitively normal controls [CNs], patients with MCI due to AD, patients with AD dementia, and patients with non-AD brain disorders) were obtained from Cerveau Technologies, which gathered, curated, organized, and archived the data under contractual agreement from 9 separate studies at 9 clinical sites. All studies had Investigational Review Board approval, and participants provided written informed consent for the procedures, including handling of imaging data. Scans were received as reconstructed anonymized DICOM image volumes corrected for scatter, randoms, and attenuation. Data from each center were acquired according to its own imaging protocol, and scan time windows were therefore not guaranteed to overlap at later postinjection times. Therefore, to generate an average static image for visual assessments, we used the most common overlapping late time-frame window (6 × 5-min frames), which was between 60 and 90 min after injection.

Definition of MK-6240 Uptake Patterns

For developing the visual read method, we randomly selected a subset of the original database containing 102 participants (52 CN, 17 MCI, 29 AD, and 4 non-AD neurologic disease). ^{18}F-MK-6240 averaged images, obtained by creating a mean image from the serial static 5-min frames within the 60- to 90-min time window, were reviewed by an experienced nuclear medicine physician–researcher (reader 0) without a prespecified examination protocol and masked to each participant's diagnosis and imaging site. This process of unguided examination led to the identification and categorization of common uptake patterns that were generally consistent with current understanding of AD tau pathophysiology and patterns that were better described as non-AD or off-target binding. On the basis of this review, an initial visual read procedure was developed for field testing and refinement using the second half of the image dataset (102 different, nonoverlapping participants).

Visual Assessment Algorithm for ^{18}F-MK-6240

The visual assessment of ^{18}F-MK-6240 is a 3-step process involving, first, assessment of technical adequacy; second, systematic review of neocortical areas for the presence and spatial extent of increased radiotracer uptake; and third, application of a rule set to the findings of step 2 for determination of positivity and classification as an AD pattern or not. Details are described in section S-1 of the supplemental materials (supplemental materials are available at http://jnm.snmjournals.org).

Readers focus on 8 prespecified brain regions in each hemisphere of the cerebral cortex (16 regions total): hippocampus, mesial temporal, inferior temporal, lateral temporal, parietal, posterior cingulate, occipital, and frontal lobes (Fig. 1). Temporal regions are grouped under the designation cluster 1, whereas the extratemporal cortical regions are grouped under the designation cluster 2. Cluster 3 comprises subcortical regions (striatum/globus pallidus, thalamus, pons, and dentate nucleus) suggestive of binding related to non-AD tauopathy.

For each of the 16 cortical regions, the reader—informed by training examples—ascertains whether there is abnormally increased radiotracer in the region relative to the cerebellum. The reader also assigns a regional-extent score expressed as the percentage of the region showing abnormal increased uptake: none (0%), 1%–25%, 26%–75%, or more than 75% involvement of the region.

Regional positivity is defined by either focal or confluent uptake involving at least 1%–25% of the region (Table 1). Readers are asked to judge the extent (voxels with increased uptake within each region) rather than the intensity of uptake.

The initial algorithm for assessing positivity was according to 3 rules (Fig. 2). This initial set of rules was refined further on the basis of additional information and experience. Rule 1 is that a scan showing normal findings has no more than one region of focally increased uptake in clusters 1 or 2 (combined) and no regions of focal uptake in cluster 3. Rule 2 is that a scan positive for the AD pattern shows two or more cluster 1–positive or 2–positive regions, with at least one positive region in cluster 1 and no positive regions in cluster 3. Rule 3 is that a scan positive for the non-AD pattern shows two or more cluster 2–positive or cluster 3–positive regions, with no positive regions in cluster 1.

Readers

Two board-certified nuclear medicine physicians with research backgrounds in brain molecular imaging and some experience with flortaucipir PET, but naïve to ^{18}F-MK-6240, served as testers of the read method. First, readers reviewed the scientific background, study rationale, and description of the read method with focus on brain region identification. After a group case review and individual testing for competency, field testing with the 2 independent readers and

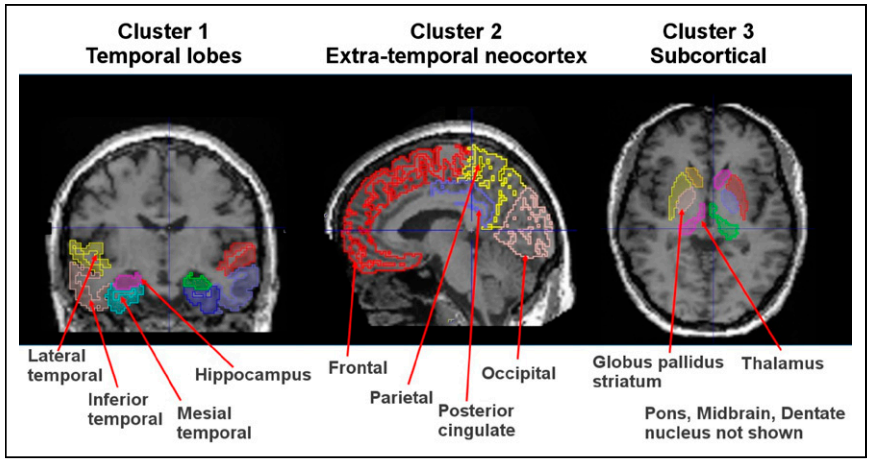

FIGURE 1. Regions for visual read are outlined and overlaid on T1-weighted MR images for anatomic reference.

TABLE 1
Regions for Visual Assessment

Brain area	Included regions	Visual rating	No. regions	Rationale
Cluster 1, temporal lobes	Hippocampus; mesial temporal; inferior temporal; lateral temporal	No uptake (0%); uptake 1%–25%; extension 26%–75%; >75% extension	8 regions: 4 each in left and right hemispheres	Earliest cortical regions involved in AD per Braak staging
Cluster 2, extratemporal neocortex	Occipital; posterior cingulate; parietal; frontal	No uptake (0%); uptake 1%–25%; extension 26%–75%; >75% extension	8 regions: 4 each in left and right hemispheres	Next regions involved in AD
Cluster 3, subcortical area	Striatum-globus; thalamus; dentate nucleus; pons; midbrain	Presence or absence	5 regions	May be positive in non-AD tauopathies

1 in-house expert reader was performed. In total, 112 test scans (102 cases + 10 repeat scans) were randomly presented without clinical or diagnostic information.

Scans were read in either linear gray scale or inverse gray scale. Readers were not provided MRI or CT results or other structural information. Images were maintained on a DICOM server accessed by remote desktop software to run PMOD, version 3.8 (PMOD Technologies), for visual display and adjustment of the PET scans. All interactions with the read platform were logged by PMOD. Readers recorded their findings in an electronic report form, which captured data, time, and user information for each case. Reads were conducted by 2 readers over 2 d and 1 reader over 7 d, all initiated within 1 d of the training.

Evaluation of Read Method

Binary readouts (positive or negative) for tau deposition by the 2 field test readers (readers 1 and 2) were compared with the gold standard read by the internal nuclear medicine reader (reader 0). Concordance for visual assessment among all 3 readers was tallied for positive and negative cases and expressed as a percentage. Cohen κ and Fleiss κ were used to assess reader-by-reader and group agreement, correcting for chance agreement for the overall scan assessment. Agreement was also evaluated on a region-by-region basis among all 3 readers for both the binary determination and regional extent using the Fleiss κ-statistic (*14*). In addition, an exploratory overall tau visual regional-extent score (VRES) was calculated for each region by assigning a value to the categoric region score of 1 for visual scores of more than 75%, 0.5 for scores of 26%–75%, 0.25 for scores of less than 25%, and 0 for scores of 0. Hence, a scan with complete bilateral uptake involving greater than 75% extension throughout each region has a total VRES of 16.

Other assessments included intrarater test–retest reliability for scan positivity or negativity, determined for those scans ($n = 10$) that were randomly presented twice to the readers; sensitivity, specificity, and area under the receiver-operating-characteristic curve, determined using the site clinical diagnosis for participants for whom this information was available ($n = 91$); and self-reported reader confidence in their assessments (supplemental materials, section S-2).

Comparison with SUV Ratio (SUVr)

SUVr was used to compare the brain uptake in regions involved in Braak stages 1–6 and the volume-of-interest (VOI) sampling of Jack et al. (*15*) to binary and VRES reads. T1-weighted MR images and ^{18}F-MK-6240 PET scans were obtained from the Cerveau database. ^{18}F-MK-6240 PET scans were processed and analyzed similarly to visual read images. Details of the image processing are provided in the supplemental materials, section S-3.

Statistical Analysis

Comparison between demographics in the development and testing scan groups used descriptive statistics and unpaired *t* tests. For visual reads, percentage concordance against the read standard (reader 0) was determined, and pairwise and group reader agreement was assessed with Cohen and Fleiss κ-statistics, respectively. The readers' total VRES (0−16) was compared using ANOVA and post hoc Welch *t* tests. VRES totals were also correlated with SUVr using Pearson correlation. Further, individual visual reads, VRES analyses, and SUVrs using different VOI strategies underwent receiver-operating-characteristic curve analysis for determination of sensitivity and specificity against the clinical site diagnosis as the truth standard.

Refinement of Algorithm for Assessing Positivity

After the masked read, a study close-out and image review with all readers assessed the adequacy of the initial rule set for describing scans encountered in the presented cases.

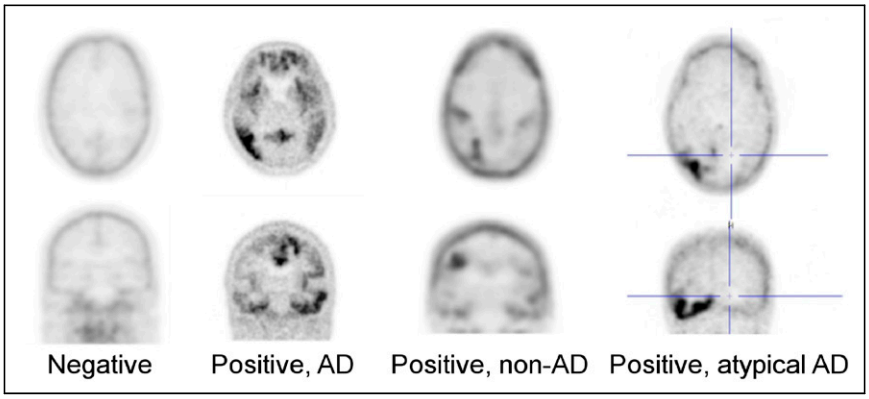

FIGURE 2. ^{18}F-MK-6240 PET in CN volunteer, AD patient, and patient with non-AD tauopathy. "Positive, atypical AD" was added in refined algorithm. Non-AD tauopathy patterns with only subcortical (cluster 3) and no cortical uptake (clusters 1 and 2), as would be expected in progressive supranuclear palsy (PSP), were not encountered in this primarily MCI/AD dataset.

This information was supplemented by other data sources, including published studies on the regional patterns of radiotracer uptake and review of another internal ^{18}F-MK-6240 PET dataset of MCI and mild-AD participants. On this basis, the algorithm was refined. These refinements were then applied to the original dataset for comparison to the preliminary algorithm.

RESULTS

Overall Binary Scan Assessment

The development (n = 102) and testing (n = 102) datasets were not significantly different with regard to age, self-reported sex, or diagnostic cohort (Table 2). No scans were excluded from review by any reader for technical deficiencies.

Comparing overall binary tau positivity (original rules 2 and 3) against the gold standard by reader 0, both reader 1 and reader 2 had a high level of concordance, with complete agreement on 107 (95%) and 108 (95%) of 112 scans (102 original + 10 repeated), respectively. All 3 readers were in complete agreement on 105 of 112 cases (94%). The Cohen κ for pairwise comparisons was 0.964 and 0.955, respectively, for reader 0 versus reader 1 and reader 0 versus reader 2, indicating excellent agreement. The Fleiss κ for interreader agreement among all 3 readers was 0.912, again indicating excellent agreement. Review of the 7 discordant cases showed 2 primary causes: the first was technical issues such as improper attenuation correction or reconstruction errors due to motion artifacts on scans judged to be still interpretable (n = 5), and the second was uptake in the inferior and mesial temporal lobes that was incorrectly attributed to off-target meningeal uptake at the base of the skull (n = 2) (Fig. 3). Discrimination of this nonspecific uptake from adjacent cortical regions may be improved by leveraging structural imaging (MRI or CT), which was not allowed in the present study (Supplemental Fig. 1).

Regional Assessments

Assessments were made in 1,792 regions (112 cases × 16 regions). Evaluation of the 16 regions for reader agreement on the binary determination of positivity or negativity for each region showed that 3 of 3 readers agreed in 1,329 regional reads (74.3% of regions), 2 of 3 readers agreed in 367 regional reads (20.4%), and 96 regions were read (5.3%) with complete discordance. Over the 16 regions, complete agreement among the 3 readers ranged from 69% to 86%. There was excellent agreement for some regions, with the Fleiss κ ranging from 0.726 (left mesial temporal) to 0.945 (right parietal) (Table 3).

TABLE 2
Demographic Data

	Development dataset				Test dataset			
			Sex (n)				Sex (n)	
Group	n	Age (y)	M	F	n	Age (y)	M	F
AD	29	72.4 (9.9)	20	9	24	70.4 (10.7)	12	12
MCI	17	71.2 (7.1)	11	6	21	69.9 (8.0)	12	9
CN	52	66.4 (12.1)	25	27	45	68.6 (7.5)	18	27
Other	4	63.3 (4.6)	0	4	12	65.9 (9.4)	5	7

Age is mean followed by SD in parentheses.

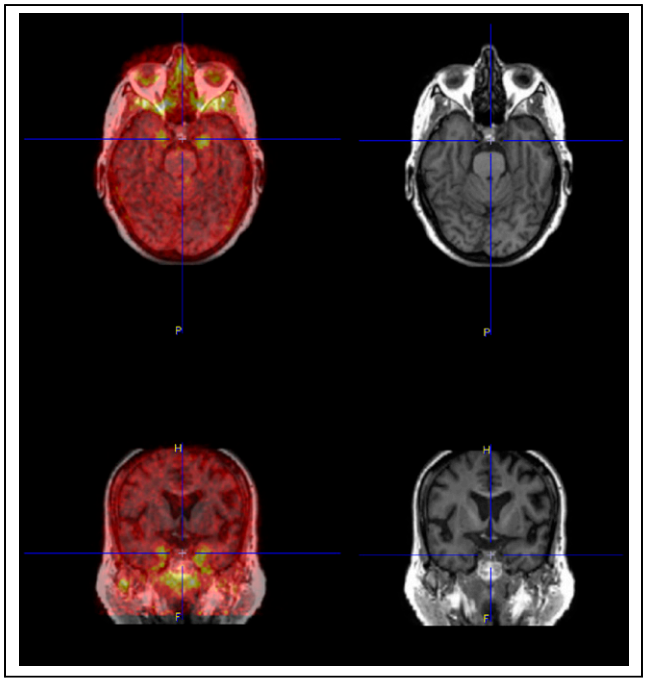

FIGURE 3. Difficult case showing bilateral anterior mesial temporal uptake, which can be confused with off-target uptake in meninges and floor of calvarium. Fused axial image of MK-6240 with the participant's T1-weighted MR image (top left), fused coronal image of MK-6240 with the participant's T1-weighted MR image (bottom left); axial (top right) and coronal (bottom right) views of MRI only.

Evaluation of reader agreement across the 4 possible regional responses (0, <25%, 25%–75%, and >75%) showed substantial agreement for 8 of 16 regions and moderate agreement for the other 8 regions. The regions demonstrating the highest reader agreement were the lateral temporal lobes and posterior cingulate cortex, whereas those regions demonstrating the least reader agreement were the occipital lobes, right hippocampus, and left mesial temporal lobe.

Reproducibility

The reproducibility of the 10 randomly selected cases for binary determination of positive or negative was excellent, with readers 0 and 2 achieving 100% reproducibility and reader 1 achieving 90%. The reproducibility of the binary positive-or-negative determination for the 16 cortical regions for the 10 repeated case pairs was also very good, with readers 0, 1, and 2 showing 93.8%, 95.0%, and 98.8% self-agreement, respectively, for the 10 case pairs of 16 regions. Overall, of the 480 region pairs (10 case pairs of 16 regions each for 3 readers) assessed with the VRES score, 460 (95.8%) of the reads were scored identically. Regions that demonstrated the most intrareader disagreement were the left and right inferior temporal lobes, which together had 8 intrareader disagreements.

Comparison with Clinical Diagnosis

In 91 of the 102 (89%) cases for which a clear, well-supported clinical diagnosis of AD, MCI, or CN was available from the site, we assessed binary visual reads, total VRES, and diagnostic sensitivity and specificity. Among the CN volunteers, 42 of 45 had negative visual reads; positive visual reads were noted in 22 of 28 AD cases and 12 of 18 MCI cases. The consensus read for the overall binary determination of scan positivity (2/3 or 3/3 readers agree) relative to a site diagnosis of cognitive impairment due to AD or

TABLE 3
Reader Agreement on Visual Assessment as Positive or Negative for Regional Tau

Regional binary agreement on positive or negative			Complete agreement on spatial extent scoring		
Agreement	Region	Fleiss κ	Agreement	Region	Fleiss κ
Almost perfect	Right parietal	0.945	Substantial	Right lateral temporal	0.748
	Right frontal	0.929		Left lateral temporal	0.716
	Right lateral temporal	0.907		Right posterior cingulate	0.702
	Left parietal	0.906		Left posterior cingulate	0.678
	Left posterior cingulate	0.904		Right parietal	0.655
	Left lateral temporal	0.888		Left parietal	0.621
	Right posterior cingulate	0.87		Right frontal	0.618
Substantial	Left frontal	0.801		Left frontal	0.602
	Left occipital	0.801	Moderate	Right inferior temporal	0.597
	Left hippocampus	0.783		Left inferior temporal	0.576
	Right inferior temporal	0.782		Right hippocampus	0.552
	Right occipital	0.759		Right mesial temporal	0.552
	Right hippocampus	0.756		Left hippocampus	0.551
	Left inferior temporal	0.753		Left mesial temporal	0.545
	Right mesial temporal	0.742		Right occipital	0.537
	Left mesial temporal	0.726		Left occipital	0.503

Overall binary assessment, κ = 0.912.

AD dementia had a sensitivity of 81% and specificity of 93%. Comparing the visual findings of each reader with the site clinical diagnosis as the ground truth demonstrated excellent specificity and moderate to good sensitivity for the visual read method (Table 4).

Comparison with SUVr

SUVr was calculated for 87 subjects who had T1-weighted MR images for coregistration and regional segmentation. When parsed by visual read status, negative and positive read groups demonstrated a significantly different mean SUVr ($P < 0.001$) for all individual Braak regions, Braak regional combinations (Braak 1–2, 3–4, and 5–6), and VOI regions of Jack et al. (*15*) (Supplemental Table 1). Individual read data (Fig. 4) demonstrated clustering of negative scans around an SUVr of 1 (no specific binding), whereas visually positive scans showed higher SUVrs spread over a wider range. Although most negative scans clustered around an SUVr of 1, we showed a strong association ($r = 0.73$, $P < 0.0001$) between quantitative SUVrs from Braak 3–4 and semiquantitative VRES summed across all readers (Supplemental Fig. 2). In addition, we examined

TABLE 4
Sensitivity and Specificity of MK-6240 Visual Reads and SUVr Analyses Using Clinical Diagnosis as Standard of Truth

Method	Sensitivity	Specificity	Youden index
Visual read			
Reader 0	0.79	0.93	0.72
Reader 1	0.75	0.93	0.68
Reader 2	0.79	0.91	0.70
Consensus	0.81	0.93	0.74
Visual VRES (cutoff, 1.5)	0.742	0.890	0.63
SUVr			
Braak 1–2 (cutoff, 1.4)	0.649	0.888	0.54
Braak 3–4 (cutoff, 1.3)	0.645	0.963	0.61
Braak 5–6 (cutoff, 1.2)	0.611	0.960	0.57
Jack VOI (cutoff, 1.4)	0.650	0.880	0.53

Jack VOI = VOI regions of Jack et al. (*15*).

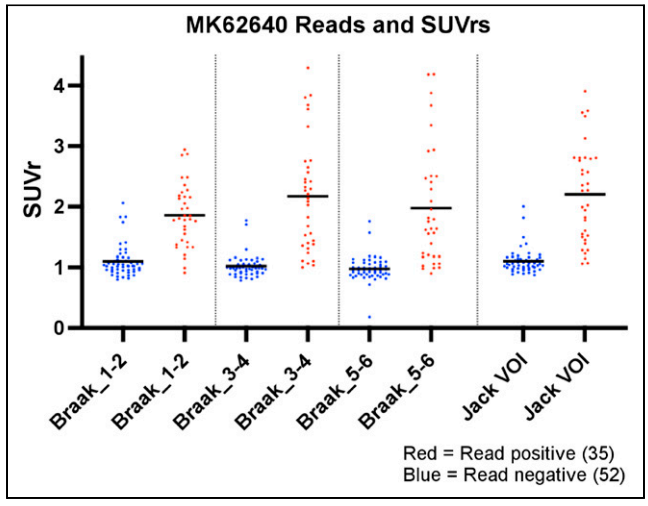

FIGURE 4. SUVr for different regional VOIs parsed by visual read of negative or positive. Jack VOI = VOI regions of Jack et al. (*15*).

the regional distribution of total VRES and found a Braaklike staging, with a greater number or cases showing higher VRESs in the temporal regions than in the extratemporal cortical regions (Supplemental Fig. 3).

Receiver-operating-characteristic analyses were performed to assess different VOI template sampling for SUVr relative to the clinical site diagnosis as the gold standard. Comparing VOI sampling strategies with the site clinical diagnosis as the truth standard using the highest Youden index for each analysis as the SUVr cutoff (Supplemental Fig. 4) again demonstrated excellent specificity and moderate sensitivity with Braak 3–4. The VOI regions of Jack et al. (*15*) showed slightly higher sensitivity and specificity than did Braak 1–2 or Braak 5–6 analyses, although the difference did not achieve statistical significance ($P = 0.59$).

Refinement of Algorithm for Assessing Positivity

The original algorithm for assessing positivity performed well in separating tau-positive from tau-negative scans consistently across readers. Nevertheless, a few changes were suggested after reader review, review of additional ^{18}F-MK-6240 PET datasets, and consideration of the published literature (*16*). These resulted in the adjustments described in Table 5.

The revised algorithm was applied to the regional data, with little effect on the determination of positivity by the readers. The final assessment was changed for only 1 reader in only 2 scans, both of which were reclassified from negative to positive, atypical AD pattern, improving the intrareader concordance from 90% to 100% for that reader. Overall concordance among the 3 readers showed 106 of 112 (95%) to be in complete agreement on the binary assessment of positivity.

DISCUSSION

This paper describes the development, initial evaluation, and refinement of a visual read method for ^{18}F-MK-6240 brain PET assessing the uptake patterns consistent with tau deposition in patients with AD. The method provides 2 sets of related information: the binary determination of positive or negative for the presence of pathologic tau, and the regional extent of tau deposition in brain areas thought to be involved in the pathologic progression of AD. The utility of this method lies in its potential use in clinical trials, particularly for eligibility assessments to confirm the presence of the targeted pathology, as well as to measure disease progression and treatment effect.

We demonstrated a high reader concordance for binary (positive or negative) and regional assessment of brain uptake consistent with expected tau pathology. Discordance among readers was relatively low and was due primarily to technical artifacts or to confusion of uptake at the base of the skull for uptake in the inferior and mesial temporal lobes. As expected, absolute agreement for VRES (4 possible categories) was lower but still substantial to moderate and varied by region. Agreement among readers was lower in the temporal lobe regions (excepting the lateral temporal lobes) than in other cortical regions (e.g., parietal, posterior cingulate), perhaps because of more difficult anatomic localization relative to the more accessibly definable and larger regions and high meningeal uptake adjacent to inferomedial temporal regions.

Our data show robust test–retest reproducibility and good accuracy relative to clinical site diagnoses. The region with the lowest reproducibility was the inferior temporal lobes; signal in this region was infrequently misattributed to extracerebral uptake in surrounding regions. Off-target uptake of ^{18}F-MK-6240 involves primarily the meninges, as well as uptake at the base of the calvarium, both of which had only a minor impact on scan interpretation in this cohort and were minimized with the aid of structural imaging. This result compares favorably with that of ^{18}F-AV-1451, for which it is typically difficult to evaluate medial temporal structures because of the proximity of off-target uptake within the choroid plexus.

We observed a strong concordance between SUVr and the binary read, with visually negative reads clustered around an SUVr of 1.0, consistent with negligible binding. By contrast, visually positive scans had higher SUVrs that were about twice the value of negative

TABLE 5
Initial and Revised Algorithms for Interpreting MK-6240 PET

Scan assessment	Original algorithm	Refined algorithm	Reason for adjustment
Negative	No more than one region positive in cortex	All clusters negative	Allowing one region to be positive was to prevent interpretation of scans as positive when meningeal uptake near inferolateral temporal lobes could be misread as positive region; this was dropped with improved methods or instructions to identify this confounder
Positive, AD pattern	Evidence of increased uptake in two cortical regions with at least one region in temporal lobes	At least one cluster 1 region positive and no cluster 3 regions positive	Revision was made after observation of multiple cases with isolated cluster 1 abnormality in just one region
Positive, atypical AD pattern	Not assessed	Increased uptake in one or more regions in cluster 2 but not in clusters 1 or 3	This rare pattern was noted from review of other MK-6240 datasets and was expected pattern based on published literature on other tau PET tracers
Positive, non-AD pattern	Any positive scan not fitting AD criteria	Increased uptake in one or more regions in clusters 1–3, with at least one region in cluster 3	Adjustment was made after formalized uniform assessment for regions involved in non-AD tauopathies

reads. Analysis of the semiquantitative VRES found that uptake occurred more frequently in temporal cortical structures than in extratemporal regions, as predicted from the models of tau spread from the postmortem data. However, these differences could also be due to the lack of normalization for region size, which ran from relatively small in the mesial temporal cortex to very large in the frontal lobe. Partial-volume error correction, which was not performed, would be expected to increase the differences between temporal and nontemporal regions.

Relative to clinical diagnosis, both the visual read and SUVR analysis showed excellent specificity and moderate to good sensitivity, with the best combination of sensitivity and specificity (81% and 93%, respectively) in the consensus read. Not surprisingly, the consensus read is the most commonly used PET interpretation method for eligibility assessments in clinical therapeutic trials, where amyloid PET negativity rates may be as high as 10%–20% among individuals thought to have AD on clinical examination (*17*). In addition, lower sensitivity relative to clinical diagnosis is common in cross-sectional datasets given the observed tendency for the baseline clinical diagnosis to identify AD with higher sensitivity and lower specificity than does imaging when the gold standard is final clinical diagnosis after longitudinal follow-up over a year or more (*18*). Both the semiquantitative SUVR analysis and the visual read performed similarly for sensitivity and specificity, with visual reads having slightly higher Youden indices. Although the SUVr cutoffs used here represent an optimized Youden index (Supplemental Fig. 2), they may not be ideal for addressing other research questions, which may require a different point on the receiver-operating-characteristic curve.

Limitations

Although the readers were naïve to ^{18}F-MK-6240 PET, they were brain imaging specialists with experience reading other tau and amyloid PET tracers. Reader selection was intentional to evaluate a method designed to visually assess brain tau burden in research and clinical trials, with regional information for tracking change and comparing patients, rather than the simple binary positive/negative determination normally required in clinical routine. Hence, compared with readers who have less experience, the study readers may have more readily handled the difficult task of accurately identifying subregions within the temporal lobes and other neocortical regions. Future studies will aim to evaluate the method and training paradigm on less experienced readers to determine the generalizability of the binary and regional-extent aspects of the method. Perhaps more germane to routine clinical use is that a simpler version of this method can be used for the easier task of determining overall scan positivity.

Limitations of this study are a lack of longitudinal data, no pathologic diagnostic confirmation, and limited data on amyloid status or clinical measures to support the accuracy of clinical diagnosis. Moreover, our development sample included a limited number of MCI patients ($n = 17$), which may limit the applicability of the algorithm to clinical diagnosis and clinical trial inclusion. Specifically, in cases with lower ^{18}F-MK-6240 binding, such as in MCI, off-target binding in meningeal tissue may theoretically lower the accuracy of visual reads in proximal brain regions. Future research will aim to include a better-characterized sample, more representative of patients screened for initial memory complaints. Furthermore, there are exceptions to the rules; specifically, logopenic primary progressive aphasia will score as typical AD because of unilateral left hemispheric uptake in clusters 1 and 2. The case used as an example of positivity for the non-AD pattern (Fig. 2) raises additional points. The case shows uptake bilaterally in the caudate and also in the right occipital and parietal regions. Is this a case of atypical AD or of non-AD tauopathy? In these rare cases, additional data may prove necessary to clarify and adjust the algorithm.

Finally, we included the midbrain in cluster 3 while instructing readers not to rate the substantia nigra as positive even though it is localized in the midbrain. This may be confusing for some, although the readers very rarely endorsed the midbrain regardless of the evident uptake in the substantia nigra.

The selection of a time window of 60–90 min was driven by the availability of the data and may not be ideal for separating negative from positive scans. In high-binding regions of positive scans, SUVrs may continue to rise past 90 min as washout proceeds (*13*). When one is visually comparing early versus late PET images, there is some minor residual background uptake in some of the scans. However, this is most apparent in the positive scans, which are typically easier to interpret than negative scans because of the relatively low off-target uptake. In summary, we believe that given the small impact on scan interpretability, it was preferable to open the time window rather than reduce the size of the cohort.

Future Work

Further validation of this read method is needed for greater confidence in the utility of the visual assessment toward the intended research application. The present data have resulted in a minor adjustment of the algorithm to be more reflective of the potential range of cases that a reader may encounter. Although we are confident that the proposed algorithm captures most AD-related cases, minor adjustments may be needed to better marry the scan phenomenology with the intended cohort. Future development should also focus on evaluation of patients who have serial ^{18}F-MK-6240 scans with months to years of longitudinal data, interpretation of scans over a wider range of tauopathies, evaluation in a larger group of readers with a greater range of experience with PET images, and comparison of ^{18}F-MK-6240 PET with clinical metrics and postmortem brain pathology.

CONCLUSION

This work provides a method for the visual interpretation of ^{18}F-MK-6240 brain PET images. This initial cross-sectional study demonstrated that experienced readers can use this algorithm for robust and reproducible visual interpretation of ^{18}F-MK-6240 brain scans. The overall determination of scan positivity may be useful in aiding diagnosis or enhancing the accuracy of clinical trial enrollment. Scan information derived from cortical regional assessment may be most valuable for within-patient evaluation of change over time, as well as for determining the efficacy of new treatments designed to alter the course of progression.

DISCLOSURE

The following investigators and their funding organizations contributed subject data used in this study: Davangere P. Devanand, MD, Department of Psychiatry, Columbia University Irving Medical Center, New York (grant support from R01 AG055422); Eric D. Hostetler, PhD, Merck & Co., Inc.; Keith Johnson, MD, Department of Radiology and Neurology, Massachusetts General Hospital, Harvard Medical School; Sterling C. Johnson, PhD, Alzheimer's Disease Research Center, University of Wisconsin School of Medicine and Public Health (grant support from AG021155, AG062285, AG027161, AG062715, AG062167, and S10OD025245); William C. Kreisl, MD, Taub Institute for Research on Alzheimer's Disease

and the Aging Brain, Columbia University (grant support from R01AG063888 and K23AG052633); Pedro Rosa-Neto, MD, PhD, Translational Neuroimaging Laboratory, McGill University Research Centre for Studies in Aging, Alzheimer's Disease Research Unit, Douglas Research Institute, Le Centre Intégré Universitaire de Santé et de Services Sociaux (CIUSSS) de l'Ouest-de-l'Île-de-Montréal, and Departments of Neurology and Neurosurgery, Psychiatry, and Pharmacology and Therapeutics, McGill University; Christopher C. Rowe, MD, Department of Molecular Imaging and Therapy Austin Health, Australian Dementia Network, University of Melbourne, Australian Imaging, Biomarker, and Lifestyle Study of Aging (AIBL), Florey Institute of Neuroscience and Mental Health (grant support from NHMRC APP1132604, APP1140853, and APP1152623); and Yaakov Stern, PhD, Taub Institute for Research on Alzheimer's Disease and the Aging Brain, Columbia University (grant support from 5R01 AG038465). This study was funded by Biogen. John Seibyl is a consultant to Biogen and a distinguished scientist at Invicro. David Cheng, Roger Gunn, Lilly Porat, and Alex Whittington are employed by Invicro. Jonathan DuBois, Annie Racine, Jessica Collins, Laurent Martarello, and Cristian Salinas are employed by Biogen. Qi Guo, Dustin Wooten, Eddie Stage, and Robert Comley are employed by AbbVie. Phillip Kuo is a consultant to and employed by Invicro. Masanori Ichise is a consultant. No other potential conflict of interest relevant to this article was reported.

ACKNOWLEDGMENT

We acknowledge the very helpful scientific input and logistic support provided by Megan Stark.

KEY POINTS

QUESTION: Can a visual read method for brain tau deposition using ^{18}F-MK-6240 PET provide robust overall information on positivity or negativity and information about the regional extent of abnormal tracer uptake?

PERTINENT FINDINGS: Readers can use this algorithm for robust and reproducible visual interpretation of ^{18}F-MK-6240 brain scans at the whole-brain and regional levels.

IMPLICATIONS FOR PATIENT CARE: This read method may be a useful tool for AD clinical drug development to enhance the accuracy of clinical trial enrollment, to evaluate within-patient changes over time, and to determine the efficacy of new treatments designed to alter the course of progression.

REFERENCES

1. Braak H, Del Tredici K. Top-down projections direct the gradual progression of Alzheimer-related tau pathology throughout the neocortex. *Adv Exp Med Biol.* 2019;1184:291–303.
2. Dronse J, Fliessbach K, Bischof GN, et al. In vivo patterns of tau pathology, amyloid-beta burden, and neuronal dysfunction in clinical variants of Alzheimer's disease. *J Alzheimers Dis.* 2017;55:465–471.
3. Knopman DS, Lundt ES, Therneau TM, et al. Joint associations of beta-amyloidosis and cortical thickness with cognition. *Neurobiol Aging.* 2018;65: 121–131.
4. Bao W, Jia H, Finnema S, Cai Z, Carson RE, Huang YH. PET imaging for early detection of Alzheimer's disease: from pathologic to physiologic biomarkers. *PET Clin.* 2017;12:329–350.
5. Bejanin A, Schonhaut DR, La Joie R, et al. Tau pathology and neurodegeneration contribute to cognitive impairment in Alzheimer's disease. *Brain.* 2017;140:3286–3300.
6. Buckley RF, Hanseeuw B, Schultz AP, et al. Region-specific association of subjective cognitive decline with tauopathy independent of global beta-amyloid burden. *JAMA Neurol.* 2017;74:1455–1463.
7. Hansson O, Grothe MJ, Strandberg TO, et al. Tau pathology distribution in Alzheimer's disease corresponds differentially to cognition-relevant functional brain networks. *Front Neurosci.* 2017;11:167.
8. Schwarz AJ, Yu P, Miller BB, et al. Regional profiles of the candidate tau PET ligand ^{18}F-AV-1451 recapitulate key features of Braak histopathological stages. *Brain.* 2016;139:1539–1550.
9. Braak H, Del Tredici K. Potential pathways of abnormal tau and alpha-synuclein dissemination in sporadic Alzheimer's and Parkinson's diseases. *Cold Spring Harb Perspect Biol.* 2016;8:a023630.
10. Meyer PF, Pichet Binette A, Gonneaud J, Breitner JCS, Villeneuve S. Characterization of Alzheimer disease biomarker discrepancies using cerebrospinal fluid phosphorylated tau and AV1451 positron emission tomography. *JAMA Neurol.* 2020;77:508–516.
11. Hostetler ED, Walji AM, Zeng Z, et al. Preclinical characterization of ^{18}F-MK-6240, a promising PET tracer for in vivo quantification of human neurofibrillary tangles. *J Nucl Med.* 2016;57:1599–1606.
12. Koole M, Lohith TG, Valentine JL, et al. Preclinical safety evaluation and human dosimetry of [^{18}F]MK-6240, a novel PET tracer for imaging neurofibrillary tangles. *Mol Imaging Biol.* 2020;22:173–180.
13. Betthauser TJ, Cody KA, Zammit MD, et al. In vivo characterization and quantification of neurofibrillary tau PET radioligand ^{18}F-MK-6240 in humans from Alzheimer disease dementia to young controls. *J Nucl Med.* 2019;60:93–99.
14. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* 1977;33:159–174.
15. Jack CR, Wiste HJ, Botha H, et al. The bivariate distribution of amyloid-beta and tau: relationship with established neurocognitive clinical syndromes. *Brain.* 2019; 142:3230–3242.
16. Phillips JS, Nitchie FJ VI, Da Re F, et al. Rates of longitudinal change in ^{18}F-flortaucipir PET vary by brain region, cognitive impairment, and age in atypical Alzheimer's disease. *Alzheimers Dement.* 2022;18:1235–1247.
17. Chapleau M, Iaccarino L, Soleimani-Meigooni D, Rabinovici GD. The role of amyloid PET in imaging neurodegenerative disorders: a review. *J Nucl Med.* 2022; 63(suppl 1):13S–19S.
18. Veitch DP, Weiner MW, Aisen PS, et al. Using the Alzheimer's Disease Neuroimaging Initiative to improve early detection, diagnosis, and treatment of Alzheimer's disease. *Alzheimers Dement.* 2022;18:824–857.