18F-AV-1451-A16 SAP v1.0

A Clinico-Pathological Study of the Correspondence Between 18F-AV-1451 PET Imaging and Post-Mortem Assessment of Tau Pathology

NCT02516046

Approval date: 15 Aug 2018

STATISTICAL ANALYSIS PLAN

^{18} F-AV-1451-A16

A Clinico-Pathological Study of the Correspondence Between ^{18}F-AV-1451 PET Imaging and Post-Mortem Assessment of Tau Pathology

AUTHOR: [redacted]

VERSION NUMBER AND DATE: FINAL, V1.0, 15AUG2018

STATISTICAL ANALYSIS PLAN SIGNATURE PAGE

Final Statistical Analysis Plan v1.0 (dated 15Aug2018) for Protocol ^{18}F-AV-1451-A16.

	Name	Signature	Date
Author:	[redacted]		
Position:			
Company:			

Upon review of this document, the undersigned approves this version of the Statistical Analysis P an authorizing that the content is acceptable for the reporting of this study.

	ame		Date
Approved By:	[redacted]		
Position:			
Company:			
Approved By:	[redacted]		
Position:			
Company:			
Approved By:	[redacted]		
Position:			
Company:			
Approved By:	[redacted]		
Position:			
Company:			

MODIFICATION HISTORY

Unique Identifier for this Version	Date of the Document Version	Author	Significant Changes from Previous Authorized Version
Final, v1.0	15Aug2018	[redacted]	Not Applicable – First Version

TABLE OF CONTENTS

LIST OF TABLES

1. INTRODUCTION

This document describes the rules and conventions to be used in the presentation and analysis of safety and efficacy for Avid Radiopharmaceuticals protocol ^{18}F-AV-1451-A16 final analysis. It described the data to be summarized and analyzed, including specifics of the statistical anlayses to be performed.

This statistical analysis plan (SAP) is based on final protocol amendment 3, dated 03-Apr-2018.

2. STUDY OBJECTIVES

2.1. PRIMARY OBJECTIVE

This study is designed to test the relationship between ante-mortem ^{18}F-AV-1451 Positron Emission Tomography (PET) imaging and tau neurofibrillary pathology associated with Alzheimer's disease (AD), as measured at autopsy. To accomplish this goal, two primary analyses will be performed:

1. The diagnostic performance (sensitivity/specificity) of 5 independent readers' interpretations of ante-mortem ^{18}F-AV-1451 PET imaging for detection of a pattern of ^{18}F-AV-1451 neocortical uptake that corresponds to Neurofibrillary Tangle (NFT) scores of B3 (Hyman et al., 2012; Montine et al., 2012) as measured at autopsy will be evaluated; and if success criteria are met,
2. The diagnostic performance (sensitivity/ specificity) of 5 independent readers's interpretation of ante-mortem ^{18}F-AV-1451 PET imaging for detection of a pattern of ^{18}F-AV-1451 neocortical uptake that corresponds to high levels of AD neuropathologic change as defined by National Institute on Aging Alzheimer's Association (NIA-AA) criteria (Hyman et al., 2012) as measured at autopsy will be evaluated. For individuals with cognitive impairement, high levels of AD neuropathologic change are considered adequate to explain cognitive impairement or dementia symptoms.

The trial will be considered a success if both primary analyses are met.

2.2. SECONDARY OBJECTIVES

The secondary objectives are:

- To assess diagnostic performance of ante-mortem ^{18}F-AV-1451 PET imaging, based on majority of interpretation of 5 independent readers, for detection pattern of ^{18}F-AV-1451 neocortical uptake that corresponds to NFT scores of B3 at autopsy and for detection of a pattern of ^{18}F-AV-1451 neocortical uptake that corresponds to high levels of AD neuropathologic change as defined by NIA-AA crieria at autopsy.
- To assess agreement among readers of ^{18}F-AV-1451 PET scans.

3. STUDY DESIGN

3.1. GENERAL DESCRIPTION

This study is a clinico-pathological study of the correspondence between ante-mortem ^{18}F-AV-1451 PET imaging and post-mortem assessment of tau pathology and associated NIA-AA pathological diagnosis.

Approximately 200 subjects with terminal medical conditions and projected life expectancy of ≤ 6 months were to be enrolled and imaged with ^{18}F-AV-1451 in order to obtain post-mortem histological data on approximately 80 subjects (including up to 6 subjects in a Front-Runner Cohort and up to 74 subjects in the Primary Efficacy Analysis Cohort; see below for more details). Enrollment was not formally stratified, but an effort was made to enroll subjects with cognitive status ranging from normal cognition to mild cognitive impairment (MCI) and dementia to achieve a full range of tau pathology. It was expected that at least 65% of the subjects enrolled will have MCI or dementia.

The screening and imaging procedures were desgined to be tolerable for eligible end-of-life subjects. Individuals with terminal medical conditions with projected life expectancy of ≤ 6 months and/or their legally authorized representative (LAR) provided informed consent. The subject then received a PET scan with ^{18}F-AV-1451 and was followed to death.

In order to control for the potential progressive change in tau deposition over time, cognitively impaired subjects for who death did not occur within nine months of the ^{18}F-AV-1451 PET scan were given the opportunity to undergo a second ^{18}F-AV-1451 PET scan and continue in the protocol at the Sponsor's discretion. In order to obtain a second PET scan before the nine month expiry date, it was recommended that the second scan be performed any time between six and nine months after the first scan. If the subject agreed to a second ^{18}F-AV-1451 PET scan, screening asssessments were not repeated with exception of any updates to the medical history/concurrent medications and an updated neurological disease diagnosis (according to best available information) which were collected at the time of the second ^{18}F-AV-1451 PET imaging visit. If the subject subsequently came to autopsy, the data (neurological disease diagnosis and ^{18}F-AV-1451 PET scan) from the second ^{18}F-AV-1451 PET imaging visit (nearest to time of death) will be used in the analyses. Cognitively normal (unimpaired) subjects were not required to undergo repeat scans, even if death occurred more then nine months after the ^{18}F-AV-1451 PET scan, because cortical tau deposition was expected to occur very slowly, if at all, in such subjects. That is, cognitively normal subjects remained eligible for autopsy, regardless of time from first or second scan to autopsy, and will be included in the Primary Efficacy Analysis Cohort if the subject came to autopsy at any time before the trial stopping criteria were met (refer to Section 7.1).

Up to the first 6 subjects that come to autopsy were to be considered as 'front-runners'. Avid were not blinded to the front-runner imaging and pathology results. These results were analyzed on a patient-by-patient basis and were used to refine the PET or autopsy methods (e.g., pathology staining, quantitation methods, etc.). Upon completion of the final front-runner subjects analysis, the final imaging and autopsy methods were specified. No additional PET image reads or autopsy analyses were performed until the PET image read and autopsy methods were finalized. The front-runners will not be included in the Primary

Efficacy Analysis Cohort, but additional analyses of combined Primary Efficacy Analysis plus Front-Runner Cohorts will be performed. It is to be noted that the first three subjects who came to autopsy were considered as front runners. That is, all subjects who came to autopsy starting with the the fourth subject who came to autopsy were elligible to be part of the Primary Efficacy Analysis Cohort.

The Primary Efficacy Analysis Cohort includes all subjects with valid autopsy after the Front-Runner Cohort was completed, but before the sample size requirement/trial stopped criteria were met (refer to Section 7.1). Subjects for whom autopsies were ongoing at the time of the sample size requirement/trial stopped criteria were met are also included in the Primary Efficacy Analysis Cohort.

Each subject had a screening visit, up to two ^{18}F-AV-1451 PET scan imaging visits[1], and a follow-up phone call after each imagining visit. Prior to each ^{18}F-AV-1451 PET scan, all subjects received a single intravenous (IV) bolus administration target dose of 370 MBq (10 mCi) of ^{18}F-AV-1451 injection. Subjects who came to autopsy during study received a post-mortem evaluation. Specific details of assessments that were performed at each visit can be found in Table 1.

The first 30 PET scans collected (including front-runner subjects) were to be reviewed by an independent Data Monitoring Board (DMB) to assess whether the scans could be evaluated in a manner consistent with methods developed in Phase II (^{18}F-AV-1451 Exploratory Phase Cohort) as determined by criteria outlined in the DMB Charter. The DMB had the ability to recommend unblinding of the Sponsor to all 30 cases if a pre-specified number of scans in the autopsy population had anomalies not present in the early phase studies. Study enrollment continued during the DMB's review. In the event the Sponsor proceeded with a DMB recommendation for unblinding, the unblinded subjects were to be excluded from the Primary Efficacy Analysis Cohort (neither scans nor any autopsies were to be used in the primary analyses) or the study might have been terminated. If unblinding was to occur, it should have been completed before the final image read and autopsy methods were specified. It is to be noted that, as planned, the DMB reviewed scans from the first 30 cases in a pre-specified manner. The DMB did not recommend unblinding. Therefore, the first 30 cases (with the exception of the front-runners) are eligible to be included in the Primary Efficacy Analysis Cohort.

In addition, the Sponsor was aware of a number of cases who received ^{18}F-AV-1451 and came to autopsy in studies conducted by independent investigators outside of this protocol. The Sponsor identified all investigators/institutions known to have collected at least one case that was reasonably likely to be truth standard (TS) negative at autopsy (refer to Section 3.3 for TS negative definition). The Sponsor then worked with the investigators to obtain ^{18}F-AV-1451 images and autopsy tissue from all cases that have come to autopsy at that institution (both TS positive and negative cases). Although not recruited and imaged under protocol A16, these cases would have been used to expand the main efficacy analysis set i.e., the Efficacy Analysis Set (EFF; refer to Section 5.5), should the minimum number of cognitively normal

[1] Should a ^{18}F-AV-1451 PET image obtained during a subject first ^{18}F-AV-1451 PET scan imaging visit be deemed invalid (refer to Section 3.2), then this first imaging visit could have been repeated in order to obtain a valid ^{18}F-AV-1451 PET image. Similarly, should a subject for who death did not occur within nine months of the first valid ^{18}F-AV-1451 PET image agreed to come to another ^{18}F-AV-1451 PET scan imaging visit and should the ^{18}F-AV-1451 PET scan image obtained at that imaging visit be deemed as invalid, this second ^{18}F-AV-1451 PET scan imaging could have been repeated in order to obtain a second valid ^{18}F-AV-1451 PET image.

subjects required by the sample size would not have been reached among the subjects enrolled under this protocol. It is to be noted that, as mentioned in the protocol (Section 8.6), an independent statistician who is not an employee of the Sponsor and who is not involved in the conduct and analysis of the study examined the pathology consensus data after the third consensus panel meeting and notified the Sponsor that the stopping criteria had been met (i.e., at least 14 cases in each of the TS positive and negative categories came to autopsy; refer to Section 7.1). Hence, the cases recruited by independent investigators outside of this protocol will not be included in the EFF, but supportive analyses with these cases will be performed. These supportive analyses will be detailed in a separate SAP.

Table 1 Schedule of Events

Evaluations	Screening Visit	Pre-Dose	^{18}F-AV-1451 Dose	PET Imaging (20 minutes)	End of Imaging	Follow-Up Phone Call	Autopsy
Signed Consent[A]	X						
Demographics and Baseline Characteristics	X						
Medical History/Neurologic Disease History	X						X[B]
Concomitant Medication	X						X[B]
Physical Exam/ Neurological Exam	X						
Vital Signs[C]	X	X			X		
IQCODE	X						
Urine Pregnancy Test	X	X[D]					
MMSE	X						
^{18}F-AV-1451 Injection			X				
PET Imaging				X			
Evaluation by a physician	X[E]	X			X		
Adverse Events	X	X	X	X	X	X	
Serious Adverse Events	X	X	X	X	X	X	
Death							X

[A] Subjects might have consented to a second ^{18}F-AV-1451 PET scan 6 to 9 months after the initial scan. Cognitively normal subjects were not required to undergo repeat scans and remained eligible for autopsy, regardless of time from first or second scan to autopsy.

[B] Documentation of the events leading up to death was collected including concomitant medications, significant medical events, and cause(s) of death.

[C] Vital signs were performed in the supine position. Height and weight were collected from subject's medical records.

[D] A urine or serum pregnancy test was acceptable on imaging day.

[E] Evaluation by physician or designee at screening.

3.2. EVALUATION OF ^{18}F-AV-1451 PET IMAGES

Criteria for declaring an image invalid or not evaluable were specified in advance in the Image Review Charter. In this SAP, valid images will be considered as unevaluable only if at least 3 out of 5 independent readers declare the image unevaluable for the same reason(s).

^{18}F-AV-1451 retention of PET images were interpreted by visual examination as either

- AD-: neocortical uptake not consistent with an AD pattern i.e., no increased neocortical activity in any region or increased activity is isolated to the mesial temporal, anterolateral temporal or frontal region(s);
- AD+: neocortical uptake consistent with an AD pattern i.e., increased neocortical tracer activity in the posterolateral temporal (PLT), parietal (including precuneus) or occipital region(s); or
- AD++: neocortical uptake consistent with an AD pattern and with likelihood of disease progression i.e., increased neocortical tracer activity is found in the parietal (including percuneus) region(s) regardeless of uptake elsewhere (subset of AD+) or in the frontal region in addition to PLT, occipital or parietal region(s).

For the purpose of testing the primary analyses, a scan interpretation will be considered positive (hypothesized to correspond to a B3 NFT score for the first primary efficacy analysis and an NIA-AA score of high AD pathology for the second primary efficacy analysis) if the scan was interperted as at least consistent with an AD pattern (AD+ or AD++).

Images were also assessed quantitatively (computerized measurement of Standard Uptake Value Ratios [SUVRs]) for exploratory analysis purposes.

The ^{18}F-AV1451 PET imaging results will not be shared with the investigator sites or subjects and their families prior to the end of the study.

3.3. NEUROPATHOLOGICAL EVALUATION

All neuropathological measurements on brain tissue in this trial were evaluated in a standardized fashion in a qualified laboratory. In this SAP, subject autospy specimens for which technical errors resulted in invalid or uninterpretable autospy specimens (as determined by the pathology core laboratory in accordance with criteria set in the Technical Autopsy Manual and Neuropathology Analysis Plan) will be considered as being invalid and/or uninterpretable if the specimens were assessed as being 'not evaluable' for both hemispheres in the Braak Stage or NIA-AA autopsy diagnosis by the consensus panel, as appropriate for the primary efficacy endpoint #1 (refer to Section 15.1.1) and primary efficacy endpoint #2 (refer to Section 15.1.2).

Neuropathological assessment was derived from the NIA-AA guidelines (Hyman et al., 2012). Sections from brain regions recommended for pathological assessment of AD, as well as additional, pre-specified

neocortical regions were sampled. Staining followed the procedures recommended in the NIA-AA practical approach guidelines. NFT extent was assessed using standard procedures and NFT score (B0-B3) was recorded (Hyman et al., 2012; Montine et al., 2012). The TS for the first primary efficacy analysis was constructed from NFT scores at autopsy as shown in Table 2.

Table 2 Autopsy NFT Score Truth Standard

Braak Stage	NFT Score[A]	Truth Standard
0 (no NFTs)	B0	
I-II	B1	Negative
III-IV	B2	
V-VI	B3	Positive

[A] Adapted from Hyman et al., 2012.

Amyloid pathology was evaluated using Thal and CERAD rating systems. Distribution of amyloid (Thal plaque score A0-A3; Montine et al., 2012) was recorded. The frequency of neuritic amyloid plaques was evaluated semi-quantitatively (CERAD score: none, sparse, moderate or frequent). Potential alternative neuropathologies contributing to cognitive impairment were also considered, as recommended by Montine et al., 2012, and level of AD neuropathologic change was recorded as none, low, intermediate or high as per NIA-AA guidelines (Hyman et al., 2012; Montine et al., 2012). The TS for the second primary efficacy analysis was constructed from levels of AD neuropathological change as shown in Table 3.

Table 3 NIA-AA Autopsy Diagnosis Level of AD Neuropathological Change Truth Standard

Level of AD Neuropathological Change Present[A]	Truth Standard
None Low Intermediate	Negative
High	Positive

[A] Adapted from Hyman et al., 2012.

3.4. BLINDING

A single Avid member performed periodic quality assurance (QA) by reviewing a listing of the PET image header data throughout the course fo the study. These QA reviews served to ensure image quality control processes at the imaging core laboratory was adequately performed. Avid staff performing PET image data QA assessment was not aware of the clinical status for the subjects/images being reviewed and pathology data. After the independent visual reads was completed, PET scan images were transferred to Avid for SUVr analysis which was done blinded to visual read results and pathology data. Only designated Avid personnel responsible for image QC and SUVr analysis had access to the image data. All other Avid personnel are blinded to the Primary Efficacy Analysis Cohort ^{18}F-AV-1451 PET images and visual read results as well as to the pathology data until database lock (DBL).

The independent visual readers and DMB were blinded to the subject clinical status and pathology data.

The pathological core laboratory was blinded to the subject clinical status, PET scan image(s), and visual read results.

IQVIA Biostatistical team (including statistical programmers) are blinded to the Primary Efficacy Analysis Cohort ^{18}F-AV-1451 PET scan images, visual read results, and pathology data until DBL.

For additional details about blinding and unblinding, refer to the PET imaging read manual, pathological analysis manual, and Biostatistics Unblinding Plan, as appropriate.

Finally, an independent statistician who is not an employee of the Sponsor and who is not involved in the conduct and analysis of the study reviewed the pathology consensus data in an unblinded fashion after the third consensus panel meeting in order to determined if the sample size requirement/trial stopped criteria were met (i.e., at least 14 cases in each of the TS positive and negative categories have came to autopsy; refer to Section 7.1). The independent statistician was blinded to the PET scan images and visual read results

3.5. CHANGES TO ANALYSIS FROM PROTOCOL

As per protocol Section 8.4, only ^{18}F-AV-1451 PET images from subjects with valid autopsy data were to be interpreted by visual examination. However, all collected ^{18}F-AV-1451 PET images of subjects who were still active in the study (i.e., who could potentially come to autopsy) at the time the visual reads were conducted, were interpreted by visual examination, regardless of the validity of the autopsy data.

During the study, it was identified that the imaging core laboratory (InVicro) included ~10% randomly selected scans for a second read by each imaging reader. This was a violation of the protocol. That is, rereads were not planned for this protocol because 1) the expected sample size for reads was small and 2) an intra-reader reliability has already been performed in a larger, target population-like sample population in study A05. Therefore, the second read data from this study will not be used for any analyses exception of an intra-reader consistency evaluation (refer to Section 15.4).

4. PLANNED STATISTICAL ANALYSES

No interim analysis (IA) or analysis for a Data Safety Monitoring Board (DSMB) committee were performed. That is, the only statistical analysis that will be performed for this study is the final analysis. This analysis will be performed by IQVIA Biostatistics following Sponsor authorization of this SAP, DBL, and study unblinding. Considering that the definition of each analysis set defined in this SAP (refer to Section 5) does not include any subjective criterion, Sponsor approval of this SAP will stand for the Sponsor approval of the inclusion/exclusion of each subject in/from each analysis set.

5. ANALYSIS SETS

5.1. ALL SUBJECTS ENROLLED SET

The all subjects enrolled analysis set (ENR) will contain all subjects who provided informed consent and received at least one ^{18}F-AV-1451 injection. Subjects who provided informed consent, completed screening evaluation but did not receive any ^{18}F AV-1451 injection will be treated as screen failures and will not be analyzed.

This analysis set will be used to summarize subjects' disposition, major protocol devations, and analysis sets data (refer to Section 9). Summaries will be presented overall and by most recent neurological disease diagnosis i.e., normal, mild cognitive impairment, and dementia.

5.2. SAFETY ANALYSIS SET

The safety analysis set (SAF) will be the same as ENR i.e., include all subjects who provided informed consent and received at least one ^{18}F-AV-1451 injection.

This analysis set will be used to summarize demographic and other baseline characteristics (refer to Section 10), medical/surgical history, concurrent medical conditions/diseases, family neurological disease history, and subjects neurological disease history (refer to Section 11), concomitant medications (refer to Section 12), cognition assessments at screening (refer to Section 13), exposure to ^{18}F-AV-1451 (refer to Section 14), and all safety data (refer to Section 16). Summaries will be presented overall and by most recent neurological disease diagnosis.

5.3. MODIFIED INTENT-TO-TREAT ANALYSIS SET

The modified Intent-to-Treat analysis set (mITT) will include all SAF subjects (refer to Section 5.2) recruited under this protocol with available PET Scan visual interpretation data (regardless of PET Scan image validity and evaluability)

This analysis set will be used to conduct the secondary efficacy analysis for the inter-reader agreement among readers (refer to Section 15.2.2). Summaries will be presented according to the individual reader's interpretation of the ^{18}F-AV-1451 PET image (see Section 3.2 and APPENDIX 1).

5.4. FULL ANALYSIS SET

Two full analysis set (FAS) will be defined for this study, one for each primary efficacy analysis (refer to Section 15.1).

5.4.1. FULL ANALYSIS SET #1

The FAS #1 (FAS1) will include all SAF subjects (refer to Section 5.2) recruited under this protocol that:

- came to autopsy, *including* the Front-Runner Cohort (i.e., including the first three subjects who came to autopsy),
- had valid and evaluable PET data (refer to Section 3.2), and
- had valid and interpretable autopsy specimens (refer to Section 3.3), where interpretability will be assessed for the Braak Stage.

This analysis set will be used to summarize the consensus panel's assessment of each TS (refer to Section 15.1), conduct supportive analyses to the first primary efficacy analysis (refer to Section 15.1.1) and secondary analysis of majority of readers versus the NFT Score TS (15.2.1) as well as to conduct the exploratory efficacy analysis exception of the logistic regression model for the NIA-AA autopsy diagnosis TS (refer to Section 15.3). Data will be presented according to the individual reader's interpretation of the ^{18}F-AV-1451 PET image (see Section 3.2 and APPENDIX 1), majority of readers' interpretation of the ^{18}F-AV-1451 PET image (see Section 3.2 and APPENDIX 1) or by most recent neurological disease diagnosis, as appropriate.

5.4.2. FULL ANALYSIS SET #2

The FAS #2 (FAS2) will include all SAF subjects (refer to Section 5.2) recruited under this protocol that:

- came to autopsy, *including* the Front-Runner Cohort (i.e., including the first three subjects who came to autopsy),
- had valid and evaluable PET data (refer to Section 3.2), and
- had valid and interpretable autopsy specimens (refer to Section 3.3), where interpretability will be assessed for the NIA-AA autopsy diagnosis.

This analysis set will be used to conduct supportive analyses to the second primary efficacy analysis (refer to Section 15.1.2) and majority of readers secondary analysis for the NIA-AA autopsy TS (refer to Section 15.2.1) as well as to conduct the logistic regression model exploratory efficacy analysis for the NIA-AA autopsy diagnosis TS (refer to Section 15.3). Data will be presented according to the individual reader's or majority of readers' interpretation of the ^{18}F-AV-1451 PET image (see Section 3.2 and APPENDIX 1), as appropriate.

5.5. EFFICACY ANALYSIS SET

Two EFF will be defined for this study, one for each primary efficacy analysis (refer to Section 15.1)

5.5.1. EFFICACY ANALYSIS SET #1

The EFF #1 (EFF1) will include all SAF subjects (refer to Section 5.2) recruited under this protocol that:

- came to autopsy *after* the Front-Runner Cohort (i.e., excluding the first three subjects who came

to autopsy),
- had valid and evaluable PET data (refer to Section 3.2), and
- had valid and interpretable autopsy specimens (refer to Section 3.3) , where interpretability will be assessed for the Braak Stage.

This analysis set will be used to summarize demographic and other baseline characteristics (refer to Section 10), medical/surgical history, concurrent medical conditions/diseases, family neurological disease history, and subjects neurological disease history (refer to Section 11), concomitant medications (refer to Section 12), cognition assessments at screening (refer to Section 13), exposure to ^{18}F-AV-1451 (refer to Section 14), and other efficacy endpoint (refer to Section 15.4) as well as to conduct the first primary efficacy analysis (refer to Section 15.1.1) and majority of readers secondary efficacy for the NFT Score TS (refer to Section 15.2.1). Summaries will be presented according to the individual reader's interpretation of the ^{18}F-AV-1451 PET image (see Section 3.2 and APPENDIX 1), majority of readers' interpretation of the ^{18}F-AV-1451 PET image (see Section 3.2 and APPENDIX 1) or by most recent neurological disease diagnosis, as appropriate.

5.5.2. EFFICACY ANALYSIS SET #2

The EFF #2 (EFF2) will include all SAF subjects (refer to Section 5.2) recruited under this protocol that:
- came to autopsy *after* the Front-Runner Cohort (i.e., excluding the first three subjects who came to autopsy),
- had valid and evaluable PET data (refer to Section 3.2), and
- had valid and interpretable autopsy specimens (refer to Section 3.3) , where interpretability will be assessed for the NIA-AA autopsy diagnosis.

This analysis set will be used to conduct the second primary efficacy analysis (refer to Section 15.1.2) and majority of readers secondary efficacy analysis for the NIA-AA autopsy diagnosis (refer to Section 15.2.1). Summaries will be presented according to the individual reader's or majority of readers' interpretation of the ^{18}F-AV-1451 PET image (see Section 3.2 and APPENDIX 1), as appropriate.

5.6. SENSITIVITY ANALYSIS SET

The sensitivity analysis set (SENS) will include all SAF subjects (refer to Section 5.2) recruited under this procotol that came to autopsy after the Front-Runner Cohort (i.e., excluding the first three subjects who came to autopsy) with an available PET Scan visual interpretation (regardless of PET Scan image validity and evaluability; refer to Section 3.2) and available pathological diagnosis (regardless of its interpretability; refer to Section 3.3).

This analysis set will be used to conduct supportive analysis to the primary and secondary efficacy analyses exception of the inter-reader agreement among readers (refer to Sections 15.1 and 15.2). Data will be presented according to the individual reader's or majority of readers' interpretation of the ^{18}F-AV-1451 PET image (see Section 3.2 and APPENDIX 1), as appropriate.

6. GENERAL CONSIDERATIONS

Categorical data will be summarized by the number and percentage of subjects in each category. Continuous data will be summarized by descriptive statistics including n (number of subjects with available data), mean, standard deviation (SD), median, minimum, and maximum values.

6.1. REFERENCE START DATE AND STUDY DAY

Study Day will be calculated from the reference start date and will be used to show start/stop day of assessments and events.

Reference start date is defined as the date of the first ^{18}F-AV-1451 injection (Study Day 1 is the day of the first injection).

- If the date of the event/assessment is on or after the reference start date, then:
  Study Day = (date of event/assessment – reference start date) + 1.
- If the date of the event/assessment is prior to the reference start date, then:
  Study Day = (date of event/assessment – reference start date).

In the situation where the event/assessment date is partial or missing, Study Day and any corresponding durations will be missing.

6.2. BASELINE

Unless otherwise specified, baseline is defined as the last non-missing measurement taken prior to reference start date and time, i.e. date and time of first ^{18}F-AV-1451 injection, including unscheduled assessments. In case where the last non-missing measurement and the reference start date and time coincide, that measurement will be considered as pre-baseline, with the exception of the adverse events (AEs) and medications. AEs/medications commencing on the reference start date and time will be considered as post-baseline (i.e. treatment-emergent/concomitant).

6.3. REPEATS, UNSCHEDULED VISITS AND EARLY TERMINATION DATA

In order to control for the potential progressive change in tau deposition over time, cognitively impaired subjects for whom death did not occur within nine months of the ^{18}F-AV-1451 PET scan were given the opportunity to undergo a second ^{18}F-AV-1451 PET scan and continue in the protocol at the Sponsor's discretion. If the subject agreed to a second ^{18}F-AV-1451 PET scan, screening asssessments were not repeated with exception of any updates to the medical history/concurrent medications and an updated neurological disease diagnosis (according to best available information) which were collected at the time of the second ^{18}F-AV-1451 PET imaging visit. If the subject subsequently came to autopsy, the data (neurological disease diagnosis and ^{18}F-AV-1451 PET scan) from the second ^{18}F-AV-1451 PET imaging visit

(nearest to time of death) will be used in the analyses. Cognitively normal (unimpaired) subjects were not required to undergo repeat scans, even if death occurred more then nine months after the ^{18}F-AV-1451 PET scan, because cortical tau deposition is expected to occur very slowly, if at all, in such subjects. That is, cognitively normal subjects remained eligible for autopsy, regardless of time from first or second scan to autopsy. The safety analyses will include data from all scans.

This study does not accommodate unscheduled visits and no data will be collected at the time of an early termination beside the date and reason for discontinuation.

Listings will include scheduled, repeat, and early discontinuation data.

6.4. WINDOWING CONVENTIONS

No visit windowing will be performed for this study.

6.5. SOFTWARE VERSION

All analyses will be conducted using SAS® version 9.4 or higher.

7. STATISTICAL CONSIDERATIONS

7.1. SAMPLE SIZE JUSTIFICATION AND TRIAL STOPPING CRITERIA

Assuming sensitivity/specificity of approximately 80%, a minimum of 14 autopsy cases in each of the TS groups (pathological positive or negative as defined in Table 2 and Table 3) is required for the lower bound of the two-sided 95% confidence interval (CI) to be > 50% for each individual reader.

The Sponsor is blinded to the pathological results (TS) until after the study is completed and the database is locked. Therefore, in order to ensure that an adequate number of subjects were recruited and came to autopsy for each TS group, the Sponsor assumed that subjects diagnosed on clinical grounds as having dementia due to AD are TS positive and subjects diagnosed as having no cognitive impairment (cognitively normal) are TS negative. Based on these assumptions, the study sample size requirement and trial stopping criteria were to be deemed as having been met when:

- the number of subjects diagnosed as having impairment due to AD that came to autopsy was greater than or equal to 14, AND
- the number of subjects diagnosed as having no cognitive impairment that came to autopsy was greater than or equal to 14

The Sponsor recognizes that there might have been some error in clinical diagnosis such that some subjects diagnosed clinically as having impairment due to AD may have non-AD pathology, or limited amounts of AD tau pathology, and thus be TS negative, and less often, subjects considered clinically

normal may have AD pathology and thus be TS positive. However, given the excess frequency of cases with AD diagnoses relative to clinically normal subjects recruited in this trial population, any diagnosis errors will likely help ensure a minimum of 14 TS negative subjects without jeopardizing the criterion for TS positive subjects.

In addition to the above subjects diagnosed as clinically normal or AD patients, it was expected that approximately 10-15 subjects with non-AD impairment would be enrolled and came to autopsy in this trial. However, the proportion of these subjects that will have TS positive/negative AD tau pathology is very difficult to predict and will depend on the precise cases enrolled. Thus, it was not expected that the clinical profile of these cases would be useful in estimating when a sufficient number of cases had occurred in each TS group.

In accordance with protocol amendment 2, the actual number of cases that met the TS positive and TS negative criteria at autopsy were determined in an unblinded fashion using the neuropathology consensus panel results by an independent statistician who is not an employee of the Sponsor and who is not involved in the conduct and analysis of the study. Because the number of cases in this study were too large to review in a single session, this panel met periodically during the course of the trial and reviewed the cases in batches. After the third consensus panel meeting (held in May 2018), the independent statistician reviewed the consensus panel data and notified the Sponsor that the stopping criteria were met (i.e., at least 14 cases in each of the TS positive and negative categories had come to autopsy).

By the time the independent statistician notified the sponsor that the stopping criteria were met, more than 14 subjects diagnosed as having no cognitive impairment and more than 14 subjects diagnosed as having impairment due to AD had come to autopsy. Hence, both set of stopping criteria were met after the third consensus panel meeting that was held in May 2018.

7.2. MULTICENTER STUDIES

This study was conducted by multiple investigators at multiple centers. However, there are no plans to perform a formal analysis of homogeneity across centers.

7.3. MISSING DATA

Missing efficacy and safety data will not be imputed with the exception of missing AE severity and missing AE relationship to study medication/procedure (refer to Sections 16.1.1.1, 16.1.1.2, and 16.1.1.3, respectively). Missing or partial AEs and medications dates and/or time will be handle as described in APPENDIX 2.

It is possible that for some cases, only half of the brain tissues was avaible for pathological analysis while the whole brain ^{18}F-AV1451 PET scan images were available for visual reads. For these cases, only the visual reads from corresponding side of brain will be used for efficacy analyses.

7.4. HYPOTHESES TESTING AND MULTIPLE COMPARISONS/ MULTIPLICITY

For the primary analysis of the first primary efficacy endpoint (refer to Section 15.1.1), the hypothesis to be tested is that for at least the same 3 out of 5 independent readers, the lower bound of the two-sided 95% CI for both sensitivity and specificity of ^{18}F-AV-1451 PET reading interpretations will be ≥ 50%.

For the primary analysis of the second primary efficacy endpoint (refer to Section 15.1.2), the hypothesis to be tested is also that for at least the same 3 out of 5 independent readers, the lower bound of the two-sided 95% CI for both sensitivity and specificity of ^{18}F-AV-1451 PET reading interpretations will be ≥ 50%.

This trial will be considered as successful if both primary efficacy analyses reach statistical significance at a two-sided significance level of 0.05.

All other efficacy endpoints (refer Sections 15.2.1, 15.2.2, and 15.3) will be tested without multiplicity adjustments and nominal p-values will be provided. Although nominal p-values will be provided for the secondary efficacy endpoints, the hypothesis to be tested for these endpoints are:

- The majority read results will have the lower bound of the two-sided 95% CI ≥ 55% for both sensitivity and specificity for detection of tau neurofibrillary pathology that corresponds to NFT score of B3,
- The majority read results will have the lower bound of the two-sided 95% CI ≥ 55% for both sensitivity and specificity for detection of a pattern of ^{18}F-AV-1451 neocortical uptake that corresponds to a high levels of AD neuropathologic change as defined by NIA-AA criteria, and
- The observed kappa values are ≥ 0.64 and the lower bound of the two-sided 95% CIs are ≥ 0.55 for the inter-reader agreement among readers.

No statistical testing will be carried out for disposition and withdrawal (refer to Section 9), demographic or other baseline characteristics (refer to Section 10), medical/surgical history (refer to Section 11), concurrent diseases/conditions (refer to Section 11), prior and concomitant medications (refer to Section 12), exposure to ^{18}F-AV-1451 (refer to Section 14), other efficacy endpoint (refer to Section 15.4), and all safety outcomes (refer to Section 16).

7.5. STATISTICAL MODELS

The following general statistical models will be used in the analysis of the primary, secondary, and exploratory efficacy endpoints.

7.5.1. SENSITIVITY, SPECIFICITY, POSITIVE AND NEGATIVE PREDICTIVE VALUES, AND LIKELIHOOD RATIOS

The following table contains terms used to define sensitivity, specificity, accuracy, positive predictive value (PPV), negative predictive value (NPV), likelihood ratio positive (LR+) and likelihood ratio negative (LR-):

Diagnostic Test: ^{18}F-AV-1451 PET Scan Read Result	Truth Standard	
	Positive	**Negative**
Positive	True Positive (TP)	False Positive (FP)
Negative	False Negative (FN)	True Negative (TN)

The sensitivity, which evaluates how good a test is at detecting a positive disease, will be calculated as:

$$100\times\frac{TP}{TP\ +\ FN}$$

The specificity, which estimates how likely subjects without the disease can be correctly ruled out, will be calculated as:

$$100\times\frac{TN}{FP\ +TN}$$

The accuracy, which evaluates the overall proportion of observed agreement, will be calculated as:

$$100\times\frac{TP\ +\ TN}{TP\ +\ FN\ +\ FP\ +\ TN}$$

The PPV, which is the proportion of positive test results that are true positives, will be calculated as:

$$100\times\frac{TP}{TP\ +\ FP}$$

The NPV, which is the proportion of subjects with a negative test result who are correctly diagnosed, will be calculated as:

$$100\times\frac{TN}{FN\ +\ TN}$$

The 95% CIs provided for specificity, sensitivity, accuracy, PPV and NPV will be based on the Wilson score method (Wilson 1927; Newcombe 1998).

The LR+, which is true positive rate divided by the false positive rate, will be calculated as:

$$\frac{sensitivity}{1-specificity}$$

The LR-, which is the false negative rate divided by the true negative rate, will be calculated as:

$$\frac{1-\ sensitivity}{specificity}$$

7.5.2. SIMPLE KAPPA COEFFICIENT

The agreement between two readers will be calculated using the simple kappa coefficient, i.e. Cohen's kappa. The simple kappa coefficient along with its 95% CI can be obtained by using the FREQ procedure from SAS®.

7.5.3. FLEISS' KAPPA

The agreement between the five readers will be calculated using Fleiss' kappa along with its 95% CI (Fleiss 1981).

7.6. EXAMINATION OF SUBGROUPS

The following subgroups will be investigated for the primary efficacy endpoints:

- Age group i.e., <65, ≥65 to <75, and ≥ 75 years
- Birth gender i.e., male and female
- Most recent neurological disease diagnosis i.e., normal, mild cognitive impairment, and dementia
- Atypical tau pathology status i.e., atypical AD tau pathology present in both sides of the brain , atypical AD tau pathology present from either side of the brain, and absent from both side of brain. "Not evaluable" cases will be excluded from this analysis

If the data are too sparse for at least one age subgroup (i.e. less than 5 subjects), the concerned age subgroup will be pooled with another age subgroup, as appropriate. For example, should the <65 years age subgroup has less than 5 subjects, it will be pooled with the ≥65 to <75 years age subgroup yielding to the following two age subgroups: <75 years and ≥75 years.

If the data are too sparse for one of the birth gender, one of the most recent neurological disease diagnosis or one of the atypical tau pathology status subgroup, only descriptive statistics will be provided for that subgroup. That is no statistical inference will be performed.

8. OUTPUT PRESENTATIONS

APPENDIX 1 shows conventions for presentation of data in outputs.

The shells that will be provided with this SAP will describe the presentations for this study and therefore, the format and content of the summary tables, figures, and listings to be provided by IQVIA Biostatistics.

9. DISPOSITION AND WITHDRAWALS

All subjects who provide informed consent, meet all inclusion/exclusion criteria and have received at least one ^{18}F-AV-1451 injection will be accounted for in this study.

The numbers and percentages of subjects enrolled, who received at least one ^{18}F-AV-1451 injection, completed the study, and discontinued early from the study as well as the reason for early discontinuation will be summarized based on the ENR overall and by most recent neurological disease diagnosis. For the analysis purpose, study completion will be defined as subjects who died within nine months of their most recent ^{18}F-AV-1451 PET scan and had an autopsy performed.

The numbers and percentages of subjects with any critical and/or major protocol deviation (as defined by the clinical team during the course of the study and finalized before database lock) will be summarized based on the ENR overall and by most recent neurological disease diagnosis. The number and percentages of subjects with at least one protocol deviation in each type of protocol deviation will be summarized similarly by severity (critical/major), where type of protocol deviations are:

- Administrative criteria
- Compliance with investigational product
- Concomitant medication criteria
- Eligibility and entry criteria
- Informed consent
- Laboratory assessment criteria
- Other criteria
- Serious adverse event criteria
- Source document criteria
- Study procedures criteria
- Visit schedule criteria

Summaries will be presented in decreasing order of type of protocol deviation within each severity (critical/major). Minor protocol deviations which will be presented in subjects data listing only.

The numbers and percentages of subjects included in/excluded from each analysis set and reasons for exclusion from each analysis set will be summarized based on the ENR overall and by most recent neurological disease diagnosis.

The primary and secondary cause of death as well as clincial events leading up to death with potential to affect autopsy findings (i.e., stroke, brain infection, and other) will be summarized based on the SAF overall and by most recent neurological disease diagnosis.

10. DEMOGRAPHIC AND OTHER BASELINE CHARACTERISTICS

Demographic and other baseline characteristics will be summarized for the SAF overall and by most recent neurological disease diagnosis. A similar summary will be also presented for the EFF1.

The following demographic and other baseline characteristics will be reported for this study:

- Age (years) – calculated relative to the date of the signed informed consent
- Age group (<65, ≥65 to <75, and ≥ 75 years)

- Birth gender
- Race
- Ethnicity
- Highest level of education completed
- Subject handedness
- Weight (kg)
- Height (cm)
- Body mass index (BMI) (kg/m^2)
- Childbearing potential
- Alcohol, tobacco, and recreational drug use

Asian sub-race, country of birth and number of years patients lived in the country of birth will be presented in the subjects data listing only.

10.1. DERIVATIONS

- Age (years)= (year of signed informed consent – year of birth) + 1
- BMI (kg/m^2)= weight (kg)/ height (m)^2

11. MEDICAL HISTORY

Medical history (including surgical history) and concurrent condition/disease information collected at screening, repeat ^{18}F-AV-1451 PET scan visit(s) (if any), and autopsy will be classified according to the following categories and coded to system organ class (SOC) and preferred term (PT) using the Medical Dictionary for Regulatory Activities (MedDRA), version 18.1:

- Medical history finding are defined as any findings that stopped prior to the first ^{18}F-AV-1451 injection;
- Concurrent medical conditions/diseases are defined as any findings that were still present at or started after the first ^{18}F-AV-1451 injection.

Incidence of medical history findings, surgeries, and concurrent medical conditions/diseases will be presented separately by SOC and PT for the SAF overall and by most recent neurological disease diagnosis. Summaries will be presented by alphabetical order of SOC and decreasing order of total frequency of PT. Subjects with multiple occurrences of the same SOC/ PT will only be counted once per SOC/ PT in the tables.

Family history of Parkinson's disease, Alzheimer's disease, dementia with Lewy bodies (DLB), stroke, and

other dementing or movement disorder will be summarized by family relationship for the SAF overall and by most recent neurological disease diagnosis.

Number of months since neurological disease symptom onset and number of months since neurological disease diagnosis will be summarized for the SAF overall and by most recent neurological disease diagnosis. Number and percentage of subjects in each etiology category for the 'mild cognitive impairment' and 'dementia' most recent neurological disease diagnosis will be summarized for the SAF overall. Number and percentage of subjects who had a previous CSF test (including the Aβ and Tau levels, when available) as well as the number and percentage of subjects who had a previous ApoE test, are ApoE 4 carrier (including the genotype, when available) will be provided for the SAF overall and by most recent neurological disease diagnosis.

All these summaries will be also similarly presented for the EFF1.

11.1. DERIVATIONS

For subjects who will provide the date of neurological symptom onset and/or the date of neurological disease diagnosis instead of the number of months since neurological symptom onset/neurological disease diagnosis will be computed as follows:

- Number of months since neurological disease symptom onset= [(date of signed informed consent – onset date of neurological disease symptom) + 1]/ (30.4375 days/month)
- Number of months since neurological disease diagnosis= [(date of signed informed consent – date of neurological disease diagnosis) + 1]/ (30.4375 days/month)

12. MEDICATIONS

Medications will be classified according to the following categories:

- 'Prior' medications are defined as any medications which started and stopped prior to the first ^{18}F-AV-1451 injection.
- 'Concomitant' medications are defined as any medications taken on or after the first ^{18}F-AV-1451 injection.

APPENDIX 2 describes the handling of partial dates for medications. In the case where it is not possible to determine if a medication is prior or concomitant, the medication will be classified by the worst case; i.e. concomitant.

Medications will be coded to the anatomical therapeutic chemical (ATC) level 3 and preferred drug name using the World Health Organization Drug Dictionary (WHO-DD), version September 2015 C.

Concomitant medications will be summarized by ATC level 3 and preferred drug name for the SAF overall by most recent neurological disease diagnosis. Summaries will be presented by alphabetical order of ATC

level 3 and decreasing order of total frequency of preferred drug name . Subjects with multiple occurrences of the same ATC/preferred drug name will only be counted once per ATC/ preferred drug name in the tables. A similar summary will be also presented for the EFF1.

Prior medications will be provided in subjects data listing only.

13. IQCODE AND MMSE

Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE; Jorm 1994) is a short questionnaire designed to assess cognitive decline and dementia in elderly people. The questionnaire is filled out by a relative or friend who has known the elderly person for 10 years or more. To score the IQCODE, the score of each question will be added up and divided by the number of questions i.e., 16. The result is a score that ranges from 0 to 5. A score of 3 means that the subject is rated on average as 'no change'. A score of 4 means an average of 'a bit worse' and a score of 5, an average of 'much worse'. Should the score of at least one question be missing, the IQCODE score will be set to missing. It will be administered to the subject's caregiver/informant at screening. The IQCODE score will be summarized using descriptive statistics based on the SAF overall and by most recent neurological disease diagnosis

The Mini-Mental State Examination (MMSE; Folstein *et al.* 1975) is a brief instrument used to assess cognitive function in elderly patients. The instrument is divided into two sections. The first section measures orientation, memory, and attention. The maximum score for the first section is 21. The second section tests the ability of the patient to name objects, follow verbal and written commands, write a sentence, and copy figures. The maximum score for the second section is 9. The MMSE total score is the sum of the score of each section and hence, ranges from 0 to 30. It wil be administered at screening to subjects that are able to cooperate with testing. Only the MMSE total score will be captured into the eCRF. Descriptive statistics will be provided for the MMSE total score for the SAF overall and by most recent neurological disease diagnosis.

Similar summaries will be presented for the EFF1.

14. EXPOSURE TO ^{18}F-AV-1451

There is only one scheduled IV bolus injection of ^{18}F-AV-1541 in this study, unless the subject meets the requirements to undergo a second ^{18}F-AV-1451 PET (see Section 3.1). For each injection, the net activity administered (mCi) as determined by the PET imaging core laboratory(ies) will be summarized using descriptive statistics for the SAF overall and by most recent neurological disease diagnosis. A similar summary will be presented for the EFF1.

The injection site will be presented in subjects data listing only.

15. EFFICACY OUTCOMES

All efficacy analyses will be peformed using the EFF1 or EFF2, as appropriate, unless otherwise indicated.

15.1. PRIMARY EFFICACY

15.1.1. PRIMARY EFFICACY ANALYSIS #1

For the primary efficacy analysis #1, the diagnosis performance (sensitivity/specificity) of 5 independent readers' interpretation of ante-mortem ^{18}F-AV-1451 PET scan images (AD+/AD++ or AD-; refer to Section 3.2) for detection of a pattern of ^{18}F-AV-1451 PET neocortical uptake that corresponds to NFT score B3 (Hyman et al., 2012; Montine et al., 2012; refer to Section 3.3) at autopsy will be evaluated.

	Autopsy NFT Score TS	
^{18}F-AV-1451 PET Scan Read Result	**NFT Score B3 (True Positive)**	**NFT Score B0, B1 or B2 (True Negative)**
AD+/AD++: ^{18}F-AV-1451 Neocortical Uptake Consistent with AD Pattern	TP	FP
AD-: ^{18}F-AV-1451 Neocortical Uptake Not Consistent with AD Pattern	FN	TN

The sensitivity and specificity along with their two-sided 95% CI (based on the Wilson score method) will be calculated for each of the five readers. The first primary efficacy analysis will be considered to be met if for at least the same 3 out of 5 readers, the lower bound of the 95% CI for both sensitivity and specificity are ≥ 50% (i.e., statistically significant at two-sided significance level of 0.05).

The accuracy, PPV, NPV, LR+ and LR- of the ^{18}F-AV-1451 PET imaging classification for each of the five readers relative to the autopsy NFT score TS will be also calculated. Two-sided 95% CIs (based on the Wilson score method)will be provided for the accuracy, the PPV and NPV.

Addtionnally, the number and percentage of subjects for each Braak stage and NFT stage will be provided for each hemisphere overall and by most recent neurological disease diagnosis. The number and percentage of subjects in each highest hemisphere/stain NFT score and autopsy NFT score TS will be summarized overall and by most recent neurological disease diagnosis. A listing of subjects who are autopsy NFT score TS positive and TS negative will be provided including the same information as well as the readers'interpretation of the ^{18}F-AV-1451 PET scan image(s).

15.1.2. PRIMARY EFFICACY ANALYSIS #2

For the primary efficacy analysis #2, the diagnosis performance (sensitivity/specificity) of 5 independent readers' interpretation of ante-mortem ^{18}F-AV-1451 PET scan images (AD+/AD++ or AD-; refer to Section

3.2) for detection of a pattern of ^{18}F-AV-1451 neocortical uptake that corresponds to high levels of AD neuropathologic change as defined by NIA-AA criteria (Hyman et al., 2012; refer to Section 3.3) at autopsy will be evaluated statistically, if and only if the primary efficacy analysis #1 is a success. For individuals with cognitive impairment, high levels of AD neuropathologic change are considered adequate to explain cognitive impairment or dementia symptoms.

	NIA-AA Autopsy Diagnosis TS	
^{18}F-AV-1451 PET Scan Read Result	**High AD Neuropathologic Change (True Positive)**	**No, Low or intermediate AD Neuropathologic Change (True Negative)**
AD+/AD++: ^{18}F-AV-1451 Neocortical Uptake Consistent with AD Pattern	TP	FP
AD-: ^{18}F-AV-1451 Neocortical Uptake Not Consistent with AD Pattern	FN	TN

The sensitivity and specificity along with their two-sided 95% CI (based on the Wilson score method) will be calculated for each of the five readers. The second primary efficacy analysis will be considered to be met if for at least the same 3 out of 5 readers, the lower bound of the 95% CI for both sensitivity and specificity are ≥ 50% (i.e., statistically significant at two-sided significance level of 0.05).

The accuracy, PPV, NPV, LR+ and LR- of the ^{18}F-AV-1451 PET imaging classification for each of the five readers relative to the NIA-AA autopsy diagnosis TS will be also calculated. Two-sided 95% CIs (based on the Wilson score method) will be provided for the accuracy, the PPV and NPV.

Addtionnally, the number and percentage of subjects for each Thal phase, Aβ plaque score, atypical tau pathology status, neuritic plaque score (modified from CERAD), and NIA-AA AD neuropathological change will be provided for each hemisphere overall and by most recent neurological disease diagnosis. A listing of subjects who are NIA-AA autopsy diagnosis TS positive and TS negative will be provided including the same information as well as the readers'interpretation of the ^{18}F-AV-1451 PET scan image(s).

15.1.3. Supportive Analysis to Primary Efficacy Analyses

Primary efficacy analyses #1 and #2 will be repeated based the FAS1/FAS2, as appropriate, and SENS set.

15.1.4. Subgroup Analyses to Primary Efficacy Analyses

Primary efficacy analyses #1 and #2 will be repeated for each subgroup specified in Section 7.6.

15.2. SECONDARY EFFICACY

15.2.1. DIAGNOSTIC PERFORMANCE OF ANTE-MORTEM ^{18}F-AV-1451 PET IMAGING BASED ON MAJORITY INTERPRETATION OF 5 INDEPENDENT READERS

For each image, the majority interpretation of 5 independent readers will be derived using individual readers' interpretation results. Diagnostic performance of this majority interpretation wil be assessed relative to the autopsy NFT score TS as per primary efficacy analysis #1 (refer to Section 15.1.1) and again relative to the NIA-AA autopsy diagnosis TS as per primary efficacy analysis #2 (refer to Section 15.1.2).

The sensitivity and specificity along with their two-sided 95% CIs (based on the Wilson score method) will be calculated for the majority interpretation of 5 independent readers relative to the appropriate TS.

The accuracy, PPV, NPV, LR+ and LR- of the ^{18}F-AV-1451 PET imaging classification for the majority interpretation of 5 independent readers relative to the appropriate TS will be also calculated. Two-sided 95% CIs (based on the Wilson score method) will be provided for the accuracy, the PPV and NPV.

15.2.2. INTER-RATER AGREEMENT AMONG READERS OF ^{18}F-AV-1451 PET IMAGING INTERPRETATION

Fleiss' Kappa statistics will be used to assess inter-reader agreement for the diagnostic decision associated with the primary efficacy analysis #1 and #2. For evaluation of agreement in the context of primary efficacy analysis #1, Fleiss' Kappa statistic will be calculated according to the readers' interpretation classification as specified in Section 15.1.1 along with its two-sided 95% CI and associated p-value. A second Fleiss' Kappa statistic will be calculated using the readers' interpretation classfication as specifed in Section 15.1.2 along with its two-sided 95% CI and associated p-value. Percentage of agreement will also be provided i.e., number of images for which a pair of readers in agreement divided by the total number of images evaluated by any pair of readers multiplied by 100%.

In addition, the agreement between each pair of readers (reader 1 vs. readers 2, 3, 4, and 5 separately; reader 2 vs. readers 3, 4, and 5 separately; reader 3 vs. readers 4, and 5 separately; and reader 4 vs. reader 5) will be assessed for each diagnostic decision using the simple kappa coefficient, its two-sided 95% CI, and associated p-value. Percentage of agreement will also be provided for each pair of reader i.e., number images for which the pair of readers are in agreement divided by the total number of images evaluated by the pair of readers multiplied by 100%.

15.2.3. SUPPORTIVE ANALYSIS TO SECONDARY EFFICACY ANALYSES

The diagnosis performance of ante-mortem ^{18}F-AV-1451 PET imaging based on majority interpretation of 5 independent readers analysis will be repeated based on the FAS1/FAS2, as appropriate, and SENS.

The inter-reader agreement among 5 readers of ^{18}F-AV-1451 PET imaging interpretation analysis will be repeated based on the FAS1 and SENS.

15.2.4. SUBGROUP ANALYSES TO SECONDARY EFFICACY ANALYSES

No subgroup analyses will be performed for the secondary efficacy analyses.

15.3. EXPLORATORY EFFICACY

Descriptive statistics will be presented for the quantitative analysis result of ^{18}F-AV-1451 PET scan (i.e., SUVr) by most recent neurological disease diagnosis based on the FAS. The reference region for SUVr values will be a selected white matter region derived using a parametric estimation of signal intensity (PERSI) method. The global SUVr assessment will be computed using a data driven target region using multi-block baryercentric discriminant analysis (MUBADA) method. The target region for regional SUVr are based on automated anatomical labeling (AAL) atlas.

The objective of exploratory analysis is to find a cut-off value for MUBADA SUVr for each pathological truth standard as defined in Section 3.2 and 3.3 respectively, to define a positve (corresponding to patholodical positive) and negative (corresponding to patholoical negative) scan using MUBADA SUVr values. For each pathological truth standard, a logisitical regression model will be applied using pathological truth standard as the dependent variable and MUBADA SUVr value as independent variable. Receiver operating curves (ROC) will be generated and a Youden's Index method will be used to select the best cut-off values. The ROC figure and area under the Curve (AUC) of the ROC will be provided as well as its p-value and the SUVr cut-off value.

Additionnally, the number and percentage of accurate classification for the TS positive, negative, and both combined will be provided where an accurate classifcation is defined as follows for TS positive

$$\frac{\text{Number of TS positive subjects with global SUVr value on or above the cut-off value}}{\text{Total number of TS positive subjects}} \times 100\%$$

Similarly, accurate classfication for TS negative is defined as follows:

$$\frac{\text{Number of TS negative subjects with global SUVr value below the cut-off value}}{\text{Total number of TS negative subjects}} \times 100\%$$

Finally, accurate classification for TS postive/negative combined is defined as follows:

$$\frac{\text{[(Number of TS postive subjects with global SUVr value on or above the cut-off value) + (Number of TS negative subjects with global SUVr value below the cut-off value)]}}{\text{Total number of TS positive/negative subjects}} \times 100\%$$

15.3.1. SUPPORTIVE ANALYSIS TO EXPLORATORY EFFICACY ANALYSES

No supportive analyses will be performed for the exploratory efficacy endpoints.

15.3.2. SUBGROUP ANALYSES TO EXPLORATORY EFFICACY ANALYSES

No subgroup analyses will be performed for the exploratory efficacy endpoints.

15.4. OTHER EFFICACY

To evaluate the intra-rater reliability of the ^{18}F-AV-1451 PET scan interpretation, eleven images were randomly selected to be reviewed twice by each of the 5 readers. Given the limited number of images that were reviewed twice by each reader and that no stratification on the most recent neurological disease diagnosis was performed (to ensure balanced potentially TS negative cognitively normal and potentially TS positive dementia cases), no statistical inferences will be presented. That is, only frequencies will be provided for the intra-rater reliability.

15.4.1. SUPPORTIVE ANALYSIS TO OTHER EFFICACY ANALYSIS

No supportive analyses will be performed for the other efficacy endpoint.

15.4.2. SUBGROUP ANALYSES TO OTHER EFFICACY ANALYSIS

No subgroup analyses will be performed for the other efficacy endpoint.

16. SAFETY OUTCOMES

All outputs for safety outcomes will be summarized overall and by most recent neurological disease diagnosis based on the SAF.

16.1. ADVERSE EVENTS

AEs will be coded to the SOC and PT using the MedDRA central coding dictionary, version 18.1.

TEAEs are defined as AEs that started or worsened in intensity or frequency on or after an ^{18}F-AV-1451 injection and up to 48 hours after the ^{18}F-AV-1451 injection. APPENDIX 1 describes the handling of partial dates and/or times for AEs. In case it is not possible to determine if an AE is treatment emergent or not, the AE will be classified by the worst case; i.e. treatment emergent.

An overall summary of number and percentage of subjects within each of the categories described in the sub-sections below will be provided.

Listings will include all AEs.

16.1.1. ALL TEAEs

Incidence (frequencies and percentages) of TEAEs will be presented by SOC and PT, and also broken down further by maximum severity, relationship to study medication, and relationship to study procedures. In addition, incidence of TEAEs will be presented by PT only (i.e., not broken down by SOC). Subjects with multiple occurrences of a SOC/ PT will only be counted once for that SOC/PT in the tables. Summaries will be presented by alphabetical order of SOC and decreasing order of total frequency of PT.

16.1.1.1. Severity

Severity will be classified as mild, moderate or severe. TEAEs with a missing severity will be classified as severe. If a subject reports a TEAE more than once within a SOC/PT, the AE with the worst case severity will be used in the corresponding severity summaries.

16.1.1.2. Relationship to Study Medication

Relationship to study medication, as indicated by the Investigator on the eCRF, will be classified as related or not related to study medication. TEAEs with a missing relationship to study medication will be regarded as related to ^{18}F-AV-1451. If a subject reports the same AE more than once within a SOC/ PT, the TEAE with the worst case relationship to ^{18}F-AV-1451 will be used in the corresponding relationship summaries.

16.1.1.3. Relationship to Study Procedures

Relationship to study procedures, as indicated by the Investigator on the eCRF, will be classified as related or not related to study procedures. TEAEs with a missing relationship to study procedures will be regarded as related to study procedures. If a subject reports the same AE more than once within a SOC/ PT, the TEAE with the worst case relationship to study procedures will be used in the corresponding relationship summaries.

16.1.2. INJECTION SITE REACTIONS

Injection site reactions are those events that will be classified under the 'Injection site reactions' MedDRA High Level Term (HLT). Incidence of injection site reactions will be presented by PT. A listing containing all injection site reactions will also be prepared.

16.1.3. ADVERSE EVENTS WITH AN OUTCOME OF DEATH

AEs with an outcome of death are those events which have a resolution recorded as "Fatal" on the AEs page of the eCRF. Incidence of AEs with an outcome of death will be presented by SOC and PT. A listing containing all AEs with an outcome of death will also be prepared.

16.1.4. SERIOUS ADVERSE EVENTS

Serious adverse events (SAEs) are those events recorded as "Serious" on the AEs page of the eCRF. Incidence of SAEs will be presented by SOC and PT. A listing containing all SAEs will also be prepared.

16.1.5. TEAEs LEADING TO DISCONTINUATION OF STUDY

TEAEs leading to discontinuation from study are those events with an action taken recorded as "permanently discontinued" on the AEs page of the eCRF. Incidence of TEAEs leading to discontinuation from study will be presented by SOC and PT. A listing containing all TEAEs leading to discontinuation from study will also be prepared.

16.2. LABORATORY EVALUATIONS

Not applicable.

16.3. ECG EVALUATIONS

Not applicable.

16.4. VITAL SIGNS

Vital sign measurements (pulse rate, respiratory rate, and blood pressure) will be taken in the supine position at screening, immediately prior to the ^{18}F-AV-1451 injection, and after completion of the imaging prior to discharge.

The change in vital sign measurements from baseline to completion of imaging prior to discharge will be computed as the value collected after the completion of the imaging prior to discharge minus the baseline value. Observed and change from baseline values will be summarized through use of descriptive statistics for the each ^{18}F-AV-1451 injection.

Incidence of potentially clinically significant abnormal values will also be provided for each ^{18}F-AV-1451 injection and a listing of subjects with potentially clinically significant abnormal values will be prepared.

16.4.1. POTENTIALLY CLINICALLY SIGNIFICANT ABNORMAL VITAL SIGNS CRITERIA

Potentally clinically significant abnormal vital sign measurements will be identified in accordance with the following criteria:

Table 4 Potentially Clinically Significant Abnormal Vital Signs Criteria

Variable	Unit	Low	High
SBP	mmHg	≤ 90 mmHg AND change from baseline ≤ -20 mmHg	≥ 180 mmHg AND change from baseline ≥ 20 mmHg
DBP	mmHg	≤ 50 mmHg AND change from baseline ≤ -15 mmHg	≥ 105 mmHg AND change from baseline ≥ 15 mmHg
Pulse rate	bpm	≤ 50 bpm	≥ 110 bpm AND change from baseline ≥ 20 bpm
Respiratory rate	breaths/ min	≤ 10 breaths/min	NA

16.5. PHYSICAL AND NEUROLOGICAL EXAMINATIONS

A brief medical examination will be conducted by a medically qualified study team member at the screening visit. Clinically significant changes from screening will be recorded as AEs.

A brief neurological examination will also be performed by a medically qualified study team member at the screening visit to evaluate cranial nerves, gait, sensory, motor function, coordination, and tendon reflexes. Clinically significant changes from screening will be recorded as AEs.

Incidence of physical and neurological abnormalities at screening will be provided.

17. REFERENCES

- Fleiss JL. (1981). *Statistical Methods for Rates and Proportions*. New York, NY: John Wiley & Sons.
- Folstein MR, Folstein SE, and McHugh PR. (1975). *Mini Mental State: A Practical Method for Grading the Cognitive Status of Patients for the Clinician.* Journal of Psychiatric Research; 12: 189-198.
- Hyman BT, Phelps CH, Beach TG at al. (2012). *National Institute on Aging-Alzheimer's Association Guidelines for the Neuropathologic Assessment of Alzheimer's Disease*. Alzheimers Dement.; 8(1): 1-13.
- Jorm AF. (1994). *A Short Form of the Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE): Development and Cross-validation*. Psyschological Medicine; 24: 145-153
- Mackinnon AJ and Mulligan R. (1998). *Combining Cognitive Testing and Informant Report to Increase Accuracy in Screening for Dementia*. American Journal of Psychiatry; 155: 1529-1535.
- Montine TJ et al. (2012). *National Institute on Aging-Alzheimer's Association Guidelines for the Neuropathologic Assessment of Alzheimer's Disease: A Practical Approach.* Acta Neuropathologica; 123.1: 1-11
- Newcombe RG.. (1998). *Two-Sided Confidence Intervals for the Single Proportion: Comparison of Seven Methods.* Statist. Med.; 17: 857-872.
- Wilson EB. (1927). *Probable inference, the law of succession, and statistical inference. Journal of the American Statistical Association*; 22: 209-212

APPENDIX 1. PROGRAMMING CONVENTIONS FOR OUTPUTS

OUTPUT CONVENTIONS

PAPER SIZE, ORIENTATION AND MARGINS

The size of paper will be letter for the United States. The page orientation should be landscape. Margins should provide at least 1 inch (2.54 centimeters) of white space all around the page, regardless of the paper size.

FONTS

The font type 'Courier New' should be used as a default for tables and listings, with a font size of 8. The font color should be black. No bolding, underlining, italics or subscripting should be permitted. Super-scripts will be avoided, unless absolutely necessary. Single spacing should be used for all text.

SDTM TERMINOLOGY

When possible, SDTM controlled terminology (e.g., race) will be used in the tables, listings and figures outputs.

DATES & TIMES

Depending on data available, dates and times will take the form YYYY-MM-DD HH:MM.

SPELLING FORMAT

English US.

OUTPUTS PRESENTATION

For outputs presented by most recent neurological disease diagnosis category will be represented as follows and in that order:

Most recent neurological disease diagnosis	Order for Tables and Listings
Normal	1
Mild Cognitive Impairment	2
Dementia	3

For outputs presented by autopsy NFT score TS, NIA-AA autopsy diagnosis levels of AD neuropathologic change TS, individual reader's interpretation of the ^{18}F-AV-1451 PET image or majority of readers' interpretation of the ^{18}F-AV-1451 PET image, groups will be represented as follows and in that order:

Autopsy NFT score TS	**Order for Tables and Listings**
B3	1
B0, B1 or B3	2

NIA-AA Autopsy Diagnosis Levels of AD Neuropathologic Change TS	**Order for Tables and Listings**
High Levels of AD Neuropathologic Change	1
No, Low or Intermediate Levels of AD Neuropathologic Change	2

Individual or Majority of Reader's Interpretation	**Order for Tables and Listings**
AD+/AD++	1
AD-	2

Inter-reader Agreement	**Order for Tables and Listings**
AD+/AD++	1
AD-	2

LISTINGS

All data collected on the eCRF will be presented in data listings. The listings will be ordered by the following (unless otherwise indicated in the shells): most recent neurological disease diagnosis, subject number, date, and time (where applicable).

APPENDIX 2. PARTIAL DATE AND/OR TIME CONVENTIONS

ALGORITHM FOR PRIOR / CONCOMITANT MEDICATIONS

START DATE	STOP DATE	ACTION
Known/ Partial/ Missing	Known/ Ongoing	If stop date < ^{18}F-AV-1451 injection date, assign as prior If stop date ≥ ^{18}F-AV-1451 injection date or medication is ongoing, assign as concomitant.
	Partial	If known components of stop date show that: • Stop date < ^{18}F-AV-1451 injection date, assign as prior • Stop date ≥ ^{18}F-AV-1451 injection date, assign as concomitant.
	Missing	Assign as concomitant

ALGORITHM FOR TREATMENT EMERGENCE OF ADVERSE EVENTS

START DATE/TIME	STOP DATE/TIME	ACTION
Known	Known, Partial, Missing	If start date/time < ^{18}F-AV-1451 injection date/time, then not TEAE If ^{18}F-AV-1451 injection date/time ≤ start date/time ≤ ^{18}F-AV-1451 injection date/time + 48 hours, then TEAE If start date/time > ^{18}F-AV-1451 injection date/time + 48 hours, then not TEAE
Partial, but known components show that it cannot be on or after ^{18}F-AV-1451 injection date/time and within 48 hours post-injection	Known, Partial, Missing	Not TEAE
Partial and known components show that it could be on or after ^{18}F-AV-1451 injection date/time and within 48 hours post-injection	Known, Partial, Missing	Assumed TEAE
Missing	Known	If stop date/time < ^{18}F-AV-1451 injection date/time, then not TEAE If stop date/time ≥ ^{18}F-AV-1451 injection date/time, then TEAE

START DATE/TIME	STOP DATE/TIME	ACTION
	Partial	If known components of stop date/time show that: • Stop date/time < ^{18}F-AV-1451 injection date/time, then not TEAE • Stop date/time ≥ ^{18}F-AV-1451 injection date/time, then TEAE
	Missing	Assumed TEAE

Document Approval Signature(s)

This document has been signed electronically in compliance with Quintiles electronic signature policies and procedures. This page is a manifestation of the electronic signatures, which may be confirmed within the Quintiles source system in accordance with Quintiles records retention policies.

Note: the Title field in the signature block has been intentionally left blank.

UserName: [redacted]
Title:
Date: Thursday, 16 August 2018, 10:32 AM Eastern Time
Meaning: I Approve the Document
==========

UserName: [redacted]
Title:
Date: Thursday, 16 August 2018, 10:48 AM Eastern Time
Meaning: I Approve the Document
==========