27 June 2024
EMA/CHMP/287487/2024
Committee for Medicinal Products for Human Use (CHMP)

Assessment report

Tauvid

International non-proprietary name: Flortaucipir (18F)

Procedure No. EMEA/H/C/006064/0000

Note

Assessment report as adopted by the CHMP with all information of a commercially confidential nature deleted.

Official address Domenico Scarlattilaan 6 ● 1083 HS Amsterdam ● The Netherlands
Address for visits and deliveries Refer to www.ema.europa.eu/how-to-find-us
Send us a question Go to www.ema.europa.eu/contact **Telephone** +31 (0)88 781 6000 An agency of the European Union

Table of contents

List of abbreviations

Aβ	amyloid-β
AD	Alzheimer's disease
ADAS	Alzheimer's Disease Assessment Scale
ADNC	Alzheimer's disease neuropathologic change
AE	adverse event
CDR-SB	Clinical Dementia Rating Scale – Sum of Boxes
CERAD	Consortium to Establish a Registry for Alzheimer's Disease
CHMP	Committee for Medicinal Products for Human Use
CI	confidence interval
CSF	cerebrospinal fluid
CSR	clinical study report
CTE	chronic traumatic encephalopathy
DVR	distribution volume ratio
ECG	electrocardiogram
EMA	European Medicines Agency
FAQ	Functional Activities Questionnaire
FDG	fluorodeoxyglucose
hERG	human ether-a-go-go
IWG	International Working Group
Kd	dissociation constant
MAO	monoamine oxidase
MCI	mild cognitive impairment
MMSE	Mini-Mental State Examination
MRI	magnetic resonance imaging
NACC	National Alzheimer's Coordinating Center
NIA-AA	National Institute on Aging and the Alzheimer's Association
NFT	neurofibrillary tangle
PET	positron emission tomography
PHF	paired helical filament
PK	pharmacokinetic(s)
SAC	supplemental analysis cohort
SACFAS	Supplemental Autopsy Cohort Full Analysis Set

SAE	serious adverse event
SmPC	summary of product characteristics
SUVr	standardised uptake value ratio
TEAE	treatment-emergent adverse event

1. Background information on the procedure

1.1. Submission of the dossier

The applicant Eli Lilly Nederland B.V. submitted on 3 March 2023 an application for marketing authorisation to the European Medicines Agency (EMA) for Tauvid, through the centralised procedure under Article 3 (2)(a) of Regulation (EC) No 726/2004. The eligibility to the centralised procedure was agreed upon by the EMA/CHMP on 27 January 2022.

The applicant applied for the following indication:

This medicinal product is for diagnostic use only.

Flortaucipir (^{18}F) is a radiopharmaceutical indicated for positron emission tomography (PET) imaging of the brain to assess the neocortical distribution of aggregated tau neurofibrillary tangles (NFTs) in adult patients with cognitive impairment who are being evaluated for Alzheimer's disease (AD). Flortaucipir (^{18}F) is an adjunct to clinical and other diagnostic evaluations.

For limitations of use, see sections 4.4 and 5.1.

1.2. Legal basis, dossier content

The legal basis for this application refers to:

Article 8.3 of Directive 2001/83/EC - complete and independent application

The application submitted is composed of administrative information, complete quality data, non-clinical and clinical data based on applicants' own tests and studies and/or bibliographic literature substituting/supporting certain test(s) or study(ies).

1.3. Information on paediatric requirements

Pursuant to Article 7 of Regulation (EC) No 1901/2006, the application included an EMA Decision(s) P/0199/2022 on the granting of a (product-specific) waiver.

1.4. Information relating to orphan market exclusivity

1.4.1. Similarity

Pursuant to Article 8 of Regulation (EC) No. 141/2000 and Article 3 of Commission Regulation (EC) No 847/2000, the applicant did not submit a critical report addressing the possible similarity with authorised orphan medicinal products because there is no authorised orphan medicinal product for a condition related to the proposed indication.

1.5. Applicant's request(s) for consideration

1.5.1. New active substance status

The applicant requested the active substance Flortaucipir (^{18}F) contained in the above medicinal

product to be considered as a new active substance, as the applicant claims that it is not a constituent of a medicinal product previously authorised within the European Union.

1.6. Scientific advice

The applicant did not seek scientific advice from the CHMP.

1.7. Steps taken for the assessment of the product

The Rapporteur and Co-Rapporteur appointed by the CHMP were:

Rapporteur: Martina Weise Co-Rapporteur: Jayne Crowe

The application was received by the EMA on	3 March 2023
The procedure started on	23 March 2023
The CHMP Rapporteur's first Assessment Report was circulated to all CHMP and PRAC members on	20 June 2023
The CHMP Co-Rapporteur's first Assessment Report was circulated to all CHMP and PRAC members on	12 June 2023
The PRAC Rapporteur's first Assessment Report was circulated to all PRAC and CHMP members on	26 June 2023
The CHMP agreed on the consolidated List of Questions to be sent to the applicant during the meeting on	20 July 2023
The applicant submitted the responses to the CHMP consolidated List of Questions on	13 October 2023
The CHMP Rapporteurs circulated the CHMP and PRAC Rapporteurs Joint Assessment Report on the responses to the List of Questions to all CHMP and PRAC members on	20 November 2023
The PRAC agreed on the PRAC Assessment Overview and Advice to CHMP during the meeting on	30 November 2023
The CHMP agreed on a list of outstanding issues in writing and/or in an oral explanation to be sent to the applicant on	14 December 2023
The applicant submitted the responses to the CHMP List of Outstanding Issues on	22 January 2024
The CHMP Rapporteurs circulated the CHMP and PRAC Rapporteurs Joint Assessment Report on the responses to the List of Outstanding Issues to all CHMP and PRAC members on	09 February 2024
The CHMP agreed on a list of outstanding issues in writing and/or in an oral explanation to be sent to the applicant on	22 February 2024
The applicant submitted the responses to the CHMP 2nd List of Outstanding Issues on	24 May 2024

The CHMP Rapporteurs circulated the CHMP and PRAC Rapporteurs Joint Assessment Report on the responses to the 2nd List of Outstanding Issues to all CHMP and PRAC members on	12 June 2024
The outstanding issues were addressed by the applicant during an oral explanation before the CHMP during the meeting on	N/A
The CHMP, in the light of the overall data submitted and the scientific discussion within the Committee, issued a positive opinion for granting a marketing authorisation to Tauvid on	27 June 2024
Furthermore, the CHMP adopted a report on New Active Substance (NAS) status of the active substance contained in the medicinal product	27 June 2024

2. Scientific discussion

2.1. Problem statement

This centralised application concerns Flortaucipir (^{18}F) (international non-proprietary name; research number: 18F-AV-1451, [F 18]T807, [18F]T807, LY3191748; proposed proprietary name: Tauvid). Throughout this document and unless specifically noted, reference to flortaucipir in the text refers to the radioactive form of flortaucipir (18F) and AV 1451 refers to the nonradioactive 19F-AV-1451.

2.1.1. Disease or condition

Background information on the disease

Alzheimer's disease (AD) is the most common cause of dementia in the elderly.

A recent study estimated 6.9 million persons living with AD dementia in Europe, with an additional 15.2 million persons living with milder cognitive impairment due to AD pathology (Gustavsson et al. 2022). AD and other dementias were reported as the third leading cause of death for people over 70 years of age in the EU, accounting for about 7% of total deaths (IHME 2020). AD is a significant growing socio-economic issue and early diagnosis of AD is estimated to bring socio-economic benefits (Barnett et al. 2014; Dubois et al. 2016).

Accurate diagnosis maximises opportunities to prevent unnecessary hospitalisation and other medical diagnostics and interventions. It also allows for correct decision-making involving health, legal, and financial decisions (Festari et al. 2022). Early and accurate diagnosis is critical to patient management.

AD is currently defined by underlying neuropathology hallmarks, namely the presence of NFTs and Aβ plaques. The recent research framework (Jack et al. 2018) and the latest IWG clinical diagnostic criteria (Dubois et al. 2021) underline the importance of biomarkers for amyloid and tau pathology to define and diagnose AD.

2.1.2. Clinical presentation, diagnosis

Neuropathologic diagnosis of AD

The accepted "gold standard" criteria for determining the presence of AD pathology are outlined by Hyman et al. 2012 (referred to in this text as the "2012 NIA-AA criteria" or the "2012 criteria"), that sought to disentangle the pathology of AD from the clinical syndrome and introduced the concept of AD neuropathologic change (abbreviated ADNC). ADNC refers to the presence and extent of neuropathologic changes of AD observed at autopsy, regardless of the level of cognitive impairment at the time of death.

Determination of the level of ADNC requires assessment of the 2 main hallmarks of AD: NFTs and amyloid plaques. NFTs are comprised of aggregated hyperphosphorylated intracellular tau protein and are assessed using Braak Staging (Braak and Braak 1991), or "B score" (see below Table 1).

Table 1: Various stages and B score to describe NFTs in the brain

Braak Stage	**Brain Area(s) Affected**	**B Score**
0	No NFTs	B0
I	Lesions in specific projection cells in transentorhinal region	B1
II	Lesions extending into the entorhinal region	
III	Lesions in hippocampus and the temporal proneocortex	B2
IV	Lesions in association areas of the adjoining neocortex	
V	Lesions spread superolaterally	B3
VI	Lesions extend into the primary areas of the neocortex	

Amyloid plaques are comprised of extracellular deposits of β-amyloid proteins assessed by 2 scoring systems:

- Thal phase, which accounts for the location of neuritic and diffuse amyloid plaques throughout the brain ("A score"), and
- The Consortium to Establish a Registry for Alzheimer's Disease (CERAD) scoring system, which accounts for the density of neuritic amyloid plaques in several regions of the cortex ("C score").

The intersection of these 2 pathologies, assessed by the 3 scores, results in a schema (Table 2) for determining the level of ADNC in autopsy (Hyman et al., 2012; note that not all footnotes from the original are included for brevity).

Table 2: Levels of AD neuropathologic change (Hyman et al., 2012)

A: Aβ/amyloid plaque score (Thal phases)*	C: Neuritic plaque score (CERAD)†	B: NFT score (Braak stage)‡		
		B0 or B1 (None or I/II)	B2 (III/IV)	B3 (V/VI)
A0 (0)	C0 (none)	Not§	Not§	Not§
A1 (1/2)	C0 or C1(none to sparse)	Low	Low	Low¶
	C2 or C3 (mod. to freq.)**	Low	Intermediate	Intermediate¶
A2 (3)	Any C	Low††	Intermediate	Intermediate¶
A3 (4/5)	C0 or C1 (none to sparse)	Low††	Intermediate	Intermediate¶
	C2 or C3 (mod. to freq.)	Low††	Intermediate	High

Abbreviations: CERAD, Consortium to Establish a Registry for Alzheimer's disease; mod., moderate; freq., frequent.

NOTE. AD neuropathologic change is evaluated using an "ABC" score that derives from three separate 4-point scales: Aβ/amyloid plaques (A) by the method of Thal phases, NFT stage by the method of Braak (B), and neuritic plaque score by the method of CERAD (C). The combination of A, B, and C scores receives a descriptor of "Not," "Low," Intermediate," or "High" AD neuropathologic change. "Intermediate" or "High" AD neuropathologic change is considered sufficient explanation for dementia.

In order to achieve a high level of ADNC, the highest B score must be present (B3: Braak Stage V-VI)

- with the highest A score (A3: Thal Phase 4/5), and
- 1 of the 2 highest C scores (C2 or C3: CERAD moderate or frequent).

AD as a clinical-biological entity

The International Working Group (IWG) already provided a first revision of the NINCDS-ADRDA criteria in 2007 to propose Alzheimer´s disease as a clinical-biological entity based on the combination of in-vitro biomarkers and specific clinical phenotypes and to extend the definition to the prodromal (pre-dementia) stages (Dubois et al. 2007).

In the updated International Working Group (IWG) recommendations, AD can be diagnosed only in the presence of both a specific clinical phenotype of AD and evidence of amyloid and tau pathology, supporting its definition as a clinical-biological entity (Dubois et al. 2021). This concept is challenged by NIA-AA which published revised criteria for diagnosis and staging of Alzheimer´s disease in October 2023, with a purely biological definition proposing a distinction between an asymptomatic "disease" phase and a symptomatic "illness" phase. However, the purely biological definition of AD proposed by the recent NIA update falls short given the complex nature of the disease (Bianchetti et al. 2023). Disease progression depends not only on biomarkers (here Tau PET imaging) but also on comorbidities, polypharmacy, frailty status and variables related to sociocultural history and lifestyle factors, cognitive reserve, psychic stress, age and gender.

Available diagnostic tools for AD

Three Aβ amyloid PET tracers Amyvid (Florbetapir (18F); by the applicant of this MAA; EMEA/H/C/002422; approved in 2013), Vizamyl MA (Flutemetamol 18F; EMEA/H/C/002557; approved in 2014) and Neuraceq (Florbetaben (18F) by Life Molecular Imaging GmbH; EMEA/H/C/002553; approved in 2014) have been approved in the EU for detecting underlying β-amyloid plaque in the brain in the assessment of patients undergoing evaluations for AD. A negative Aβ PET scan indicates sparse to no neuritic plaques and is inconsistent with a neuropathological diagnosis of AD at the time of image acquisition. However, since amyloid pathology can be present in the absence of AD clinical syndrome, or may not be the cause of the presenting cognitive symptoms, it cannot be used to diagnose AD by itself (Clark et al. 2011, 2012). So far, no imaging for detection of tau/NFTs has been approved in the EU and Tauvid would be the first PET product in this indication.

Other ways of diagnosing the presence of beta amyloid and tau aggregated indirectly include combined amyloid and tau cerebro-spinal fluid (CSF) biomarkers, which can reach high levels of statistically meaningful specificity and sensitivity compared to clinical diagnosis and PET-based findings (Olsson et al. 2016; Tariciotti et al. 2018; Bouwman et al. 2022) and post-mortem AD pathology (Tapiola et al. 2009; Seeburger et al. 2015; Mattsson-Carlgren et al. 2022). However, while CSF diagnostic would allow the simultaneous assessment of beta amyloid and tau, lumbar punctures might not be suitable for all subjects, or may be refused (Brazel et al. 2020) and in this case PET imaging might be preferred, or even necessary.

The latest revised criteria for diagnosis and staging of Alzheimer´s disease (AD) published by the NIA-AA (October 9, 2023) recommend use of fluid biomarkers of tau as an addition to the amyloid PET as first-line ("Core 1") diagnostic tools. Tau-imaging (e.g., flortaucipir PET) is regarded as Core 2 diagnostic tool for e.g., staging rather than for primary diagnosis of AD. The revised criteria have not yet been finalised.

Other supportive diagnostic tools not directed to the detection of beta-amyloid or tau in the brain are MRI and FDG PET. MRI is commonly applied to evaluate other potential causes of cognitive impairment, as well as to characterise structural changes typical in patients with AD (Pitner and Bachman 2004; Frisoni et al. 2010; Dubois et al. 2021). FDG PET has been used to detect brain glucose hypometabolism patterns associated with AD (Dubois et al. 2021). Both MRI and FDG PET are available in Europe, with some heterogeneity in the approved indications across the member states (Frisoni et al. 2017). While they can be useful in the diagnostic workup of patients with suspected AD, they are non-specific markers of downstream neurodegeneration, and do not provide direct evidence of the underlying pathophysiology (Jack et al. 2018; Dubois et al. 2021).

Thus, there is a certain degree of an unmet medical need in the non-invasive and accurate imaging marker for detection of tau/NFT in the brain.

2.1.3. Management

Currently, symptomatic treatments are available for patients with a diagnosis of AD dementia that can slow symptomatic decline (Yiannopoulou and Papageorgiou 2013). Lifestyle modifications have also shown benefit (Morris et al. 2015). Further, management of social support resources and planning are critically important to patients and caregivers (Alzheimer's Association 2019). In addition, disease-modifying therapies, including anti-amyloid treatments, are under development that may become available in the near future.

2.2. About the product

Tauvid is an ^{18}F-labelled diagnostic positron emission tomography (PET) radiopharmaceutical that was designed to image density and distribution of aggregated tau neurofibrillary tangles (NFTs) in the brains of adult patients being evaluated for AD.

Flortaucipir (^{18}F) binds to aggregated tau protein. In brains of patients with AD, paired helical filament (PHF) tau forms aggregates that subsequently combine to form neurofibrillary tangles (NFTs), a required component of the neuropathological diagnosis of AD. *In vitro*, flortaucipir (^{18}F) binds to PHF tau purified from brain homogenates of donors with AD. Weak binding and poor co-localisation was observed for tau aggregates from other non-AD tauopathies. *In vivo*, flortaucipir (^{18}F) is differentially retained in neocortical areas that contain aggregated tau.

Pharmacotherapeutic group: diagnostic radiopharmaceutical, central nervous system, ATC code: V09AX07

The target indication proposed initially for this MAA was:

> *This medicinal product is for diagnostic use only. Flortaucipir (^{18}F) is a radiopharmaceutical indicated for positron emission tomography (PET) imaging of the brain to estimate the density and distribution of the aggregated tau neurofibrillary tangles (NFTs) of Alzheimer's disease (AD) in adult patients with cognitive impairment who are being evaluated for AD. Flortaucipir (^{18}F) can be used to help establish a diagnosis of AD and to inform risk for progression of cognitive and functional impairment due to AD. For diagnostic performance to confirm the presence of AD neuropathology, see section 5.1. Flortaucipir (^{18}F) is an adjunct to clinical and other diagnostic evaluations*.

With response to the Day 120 LoQ the claim that Flortaucipir (^{18}F) can help to inform the risk for progression of cognitive und functional impairment due to AD is no longer claimed. This is appreciated.

The updated indication is now:

> *This medicinal product is for diagnostic use only. Flortaucipir (^{18}F) is a radiopharmaceutical indicated for positron emission tomography (PET) imaging of the brain to assess the neocortical distribution of aggregated tau neurofibrillary tangles (NFTs) in adult patients with cognitive impairment who are being evaluated for Alzheimer's disease (AD). Flortaucipir (^{18}F) is an adjunct to clinical and other diagnostic evaluations.*
>
> *For limitations of use, see sections 4.4 and 5.1.*

This indication wording was proposed by the CHMP and has been accepted by the applicant.

It should be noted that in 2020 flortaucipir (^{18}F) was approved by the FDA for the following indication based on the same studies submitted for this application:

TAUVID is indicated for use with positron emission tomography (PET) imaging of the brain to estimate the density and distribution of aggregated tau neurofibrillary tangles (NFTs) in adult patients with cognitive impairment who are being evaluated for Alzheimer's disease (AD).

2.3. Type of application and aspects on development

Eli Lilly and Company (Lilly) submitted this marketing authorisation application for flortaucipir (^{18}F).

There have been no interactions with any regulatory organisations outside the US regarding flortaucipir (^{18}F) development. The development programme for flortaucipir (^{18}F) took into account advice from FDA. Flortaucipir (^{18}F) was approved in the US in 2020. Flortaucipir (^{18}F) is not yet approved, however it has been submitted for approval, in several other countries. Registration outside the US was postponed due to the lack of reimbursement and access to radiopharmaceuticals. Consequently, Lilly focused on supplying flortaucipir (^{18}F) as a biomarker for AD therapeutic development rather than a diagnostic radiopharmaceutical.

The development programme is generally aligned with the EMA guidance on the clinical development of diagnostic agents, including radiopharmaceutical imaging agents (CPMP/EWP/1119/98/Rev. 1). While Lilly did not conduct a study on diagnostic thinking and patient management, flortaucipir (^{18}F) has been widely studied globally and the literature contains data on the use and performance of flortaucipir (^{18}F) PET in different research and clinical settings, including a large, appropriately-designed, independently-conducted, counter-balanced head-to-head crossover study comparing impact of flortaucipir PET and amyloid PET on diagnostic thinking (Altomare et al. 2021).

2.4. Quality aspects

2.4.1. Introduction

The finished product is presented as solution for injection containing 800 MBq/mL and 1900 MBq/mL of flortaucipir (^{18}F) at the date and time of calibration (ToC) as active substance.

The activity per vial ranges from 800 MBq to 12 000 MBq and from 1900 MBq to 28 500 MBq, respectively at the ToC in approximately 1 mL to 15 mL.

Fluorine (^{18}F) decays to stable oxygen (^{18}O) with a half-life of approximately 110 minutes by emitting a positron radiation of 634 keV, followed by photonic annihilation radiation of 511 keV.

Other ingredients are:

800 MBq/mL strength: anhydrous ethanol, sodium chloride, and water for injections.

1900 MBq/mL strength: disodium phosphate (for pH adjustment), diluted hydrochloric acid, anhydrous ethanol, sodium chloride, and water for injections.

The product is available in clear Type I borosilicate glass vial with chlorobutyl or FluroTec-coated elastomeric stoppers and aluminium seals as described in section 6.5 of the SmPC.

2.4.2. Active Substance

Active Substance (flortaucipir (^{18}F))

General information

The chemical name of the active substance is 7-[6-(18F)fluoropyridin-3-yl]-5H-pyrido[4,3-b]indole corresponding to the molecular formula $C_{16}H_{10}[^{18}F]N_3$. It has a relative molecular mass of 263.0859 amu and the following structure (Figure 1):

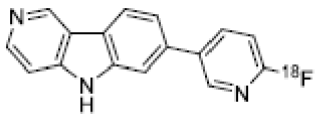


Figure 1: Active substance structure ((flortaucipir (^{18}F))

Structural characterisation was performed on the non-radioactive compound. The chemical structure of the active substance was elucidated by a combination of optical rotation, elemental analysis, Fourier Transform Infrared (FTIR) analysis, nuclear magnetic resonance (NMR) spectroscopy analysis, and X-Ray Structure Determination.

A radioactive active substance cannot be isolated as a pure substance in the form of a crystal or a powder but solved in solution. Its chemical substance amount is very low. In the case of flortaucipir (^{18}F) the chemical substance amount per dose is below nanomoles down to picomoles. Therefore, it is obvious that typical properties of solid substances as polymorphism and hygroscopicity are not relevant for a radioactive active substance.

Manufacture, characterisation and process controls (flortaucipir (^{18}F))

The active substance is not isolated during the manufacturing process. The active substance and the finished product are manufactured in one continuous process.

The active substance is synthesised in 2 main stages:

1. Manufacture of precursor (AV-1622)
2. Manufacture of final active substance (flortaucipir (^{18}F))

Precursor: AV-1622

According to the Guideline on Radiopharmaceuticals, information on chemical precursors is presented in a separate section 3.2.S in Module 3.

General information

The chemical name of the precursor is (5-[5-(tert-butoxycarbonyl)-5H-pyrido[4,3-b]indol-7-yl]-N,N,N-trimethylpyridin-2-aminium 4-methylbenzenesulfonate corresponding to the molecular formula $C_{24}H_{27}N_4O_2 \bullet C_7H_7SO_3$. It has a relative molecular weight of 403.2134 (as trimethylaminium cation) and the following structure (Figure 2):

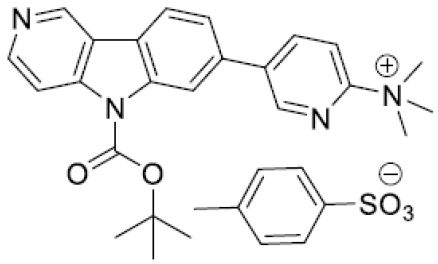


Figure 2: Precursor structure

The chemical structure of the precursor was elucidated by a combination of optical rotation, elemental analysis, Fourier Transform Infrared (FTIR) analysis, ultraviolet (UV) analysis, nuclear magnetic resonance (NMR) spectroscopy analysis, mass spectrum (MS), and X-Ray Structure Determination.

Solubility of the precursor was found to be approximately 77 mg/mL in dimethyl sulfoxide and approximately 8 mg/mL in acetonitrile.

Manufacture, characterisation and process controls (AV-1622)

The precursor is manufactured by one manufacturing site.

The precursor is synthesised in 7 main steps using well defined starting materials with acceptable specifications.

Adequate in-process controls are applied during the synthesis. The specifications and control methods for intermediate products, starting materials and reagents have been presented.
The characterisation of the active substance and its impurities are in accordance with the EU guideline on chemistry of new active substances.

Potential and actual impurities were well discussed with regards to their origin and characterised.

AV-1622 is stored in 4 or 40 mL Type I borosilicate amber glass vials fitted with polypropylene caps with PTFE/silicone septa which complies with Commission Regulation (EU) 10/2011, as amended.

Specification (AV-1622)

The precursor specification includes tests for appearance (visual), identity (FTIR, ^{1}H NMR, mass spectrometry), chromatographic purity (RP-HPLC), chromatographic impurities (RP-HPLC), assay (RP-HPLC), tosylate content (RP-HPLC), water content (KF), palladium (ICP-MS), inorganic impurities (residue on ignition), residual solvent content (GC), bacterial endotoxins (Ph. Eur.), and bioburden (Ph. Eur.).

Impurities are controlled as follows: 1) during the manufacturing process by in process controls; 2) during QC release testing by specifications; and 3) during stability studies by specifications. The manufacturing process impurities and degradants were synthesised and characterised using LC/MS and ^{1}HNMR spectroscopy except for two impurities which were assigned structures based on LC/MS data and expected by products of the manufacturing process.

Other impurities (e.g. residual solvents, inorganic impurities, palladium content, and water content) are controlled according to the specifications.

The control strategy for residual solvents and elemental impurities is deemed acceptable. The analytical methods used have been adequately described and (non-compendial methods) appropriately validated in accordance with the ICH guidelines. Satisfactory information regarding the reference standards used for testing has been presented.

Batch analysis data (n= 3 commercial scale batches) of the active substance are provided. The results are within the specifications and consistent from batch to batch.

Stability (AV-1622)

Stability data from 7 commercial scale batches of the precursor from the proposed manufacturer stored in the intended commercial package for up to 36 months under -20°C, for up 48 months and for up to 48 months under 5°C , under long term conditions (25°C/ 60% RH) for up 48 months, under intermediate conditions (30°C/ 65% RH) for up 12 months and for up 6 months under accelerated conditions (40ºC / 75% RH) according to the ICH guidelines were provided.

The following parameters were tested: appearance, water content, chromatographic purity, chromatographic impurities, and assay.

All out of specifications (OOS) results were observed at the accelerated stress condition (40°C/ 75% RH). The precursor has been shown in stress stability studies to degrade in high temperature / humidity conditions, and as such, all OOS results observed were anticipated.

Stress stability studies were conducted using the validated stability-indicating HPLC method. The effects of acid, base, and oxidizing media and light in solution were evaluated. Additionally, the effects of elevated heat, elevated heat/humidity, and light on solid samples were evaluated. Analysis of the stressed samples showed no interference between the precursor and degradation product peaks. Mass balance was achieved for all stressed conditions.

The stability studies demonstrate that the precursor is stable when packaged at the intended commercial packaging. Therefore, the stability results justify the proposed retest of 42 months when stored at controlled room temperature in the proposed container closure.

Manufacture, characterisation and process controls (flortaucipir (^{18}F))

The active substance ((flortaucipir (^{18}F)) is not isolated during the manufacturing process. Active substance and finished product are manufactured in one continuous process in which active substance is produced using an automated radiosynthesiser, formulated, and sterile filtered to form the finished product. Active substance is manufactured according to one of two manufacturing processes (Process 1 and Process 2).

Process 2 consists of 5 main steps using well defined starting materials with acceptable specifications.

Processes 1 and 2 are complimentary processes that will allow the greatest flexibility in growing the manufacturing footprint across the European Union (EU).

Adequate in-process controls are applied during the synthesis. The specifications and control methods for intermediate products, starting materials and reagents have been presented.
The characterisation of the active substance and its impurities are in accordance with the EU guideline on chemistry of new active substances.

Potential and actual impurities were well discussed with regards to their origin and characterised.

The active substance is not isolated during the manufacturing process but directly lead over to finish product manufacturing.

Specification (flortaucipir (^{18}F))

The active substance and the finished product are manufactured in one continuous process. The active substance is not isolated during the manufacturing process. Specifications for the active substance are incorporated into the specifications for the finished product.

Stability (flortaucipir (^{18}F))

The active substance and the finished product are manufactured in one continuous process. The active substance is not isolated during the manufacturing process. The stability of the active substance is incorporated into the stability studies of the finished product.

2.4.3. Finished Medicinal Product

Description of the product and pharmaceutical development

The finished product is presented as clear, colourless solution.

The 800 MBq/mL strength solution has a pH of 4.5 to 8.0 and an osmolality of approximately 2,356 mOsm/kg.

The 1900 MBq/mL strength solution has a pH of 6.0 to 8.0 and an osmolality of approximately 2,373 mOsm/kg.

The concentration of radioactive flortaucipir (^{18}F) and non-radioactive flortaucipir (^{19}F) is required to be NMT 2.0 µg/mL. The strength (radioactive concentration) of active substance in the finished product is required to be 800 MBq/mL or 1900 MBq/mL at End of Synthesis and NLT 37 MBq/mL at expiry.

The shelf-life (expiry) specification means that a 370 MBq dose of the finished product will be contained in not more than 10 mL of solution.

The active substance is poorly soluble in aqueous solution. Ethanol was considered to be the most physiologically compatible solubilising agent. The solubility of flortaucipir in the formulation requires 10% ethanol to ensure solubility of flortaucipir (at a maximum concentration of 2 µg/mL at pH 8.0) following dilution up to 1:5 by the end user.

The 800 MBq/mL formulation is not buffered but remains within the physiologically compatible range of pH 4.5 to 8.0. The 1900 MBq/mL formulation is buffered within the physiologically compatible range of pH 6.0 to 8.0. Since the active substance remains soluble in formulation in these pH ranges, this level of pH control was deemed adequate.

The 800 MBq/mL finished product is formulated in 10% v/v ethanol in 0.9% sodium chloride aqueous solution. The 1900 MBq/mL finished product is formulated in 10% v/v ethanol, up to 0.05% dibasic sodium phosphate, anhydrous and up to 1.5 mM hydrochloric acid in 0.9% sodium chloride aqueous solution.

All excipients are well known pharmaceutical ingredients and their quality is compliant with Ph. Eur standards. There are no novel excipients used in the finished product formulation. The list of excipients is included in section 6.1 of the SmPC.

Three formulations have been used during development. Initially, the finished product was formulated in 10% v/v ethanol in 0.9% sodium chloride aqueous solution which was used as 800 MBq/mL strength. The second formulation was developed while conducting early phase clinical trials. The ethanol concentration was increased to improve the solubility of the active substance. It is intended to be used for commercial manufacturing (800 MBq/mL).

The third formulation was developed after completion of the pivotal studies and is also intended to be used for commercial manufacturing (1900 MBq/mL). Based on development studies, the finished product formulation was modified to optimise and simplify the manufacturing process while also increasing the batch size (total radioactivity). Dibasic sodium phosphate was added to buffer the formulation due to the acidity of the purification mobile phase.

The formulations intended for commercial manufacturing were developed to be compatible with intravenous administration, to comprise common pharmacopeial excipients, to solubilise the active substance, flortaucipir (^{18}F), and carrier flortaucipir (^{19}F), and to allow sterilisation of the finished product by filtration.

These objectives were achieved as the finished product formulations intended for commercial manufacturing contain common pharmacopeial excipients, are at a physiologically compatible pH, are near isotonic (except for the alcohol which renders the solution hypertonic), contain ethanol, and the product is suitable for intravenous injection at a concentration sufficient to solubilise the active substance, flortaucipir (^{18}F) and flortaucipir (^{19}F).

Active substance and finished product are manufactured in one continuous process. The finished product steps include formulation, sterile filtration, dilution, and dispensing. A summary of the change in the formulation, sterile filtration, dilution manufacturing steps for the manufacturing processes utilised during clinical development and the intended commercial manufacturing processes have been provided and considered satisfactory.

The finished product manufactured using two processes: Process 1 for 800 MBq/mL formulation using the automated radiosynthesiser, and Process 2 for 1900 MBq/mL formulation which was modified and optimised to increase the yield and batch size, improve process robustness, and simplify the process. A comparison of Process 1 and Process 2 for the manufacture of the finished product using the automated radiosynthesiser was provided.

Automated and/or semi-automated dispensing systems utilizing a syringe driver, or a peristaltic pump may be used for the process of dilution and filling of the quality control (QC) and Flortaucipir (^{18}F) Solution for Injection dose vials. The dispensing systems utilise a chemically resistant, sterile, disposable fluid pathway.

The aseptic steps of the manufacture of the finished product are: the assembly of the intermediate and bulk product vials, the sterile filtration of the finished product and diluent, and the dispensing of the QC vials and multi-dose vials

All aseptic processing is carried out in a Class A distribution isolator at each of the manufacturing sites. The isolator is subjected to routine particulate, environmental, and microbiological monitoring per local methods established at each of the respective manufacturing sites. In addition, process simulations (media fills) are performed to qualify the aseptic steps of the manufacturing process. A minimum of three consecutive process simulations are performed at each manufacturing site to qualify the manufacturing process.

The container closure system is a sterile, apyrogenic 15 mL clear Type I borosilicate glass serum vial closed with a 20 mm elastomeric stopper sealed with a 20 mm crimp seal. These container closure systems were selected based on development studies that demonstrated compatibility of the closure with Flortaucipir (^{18}F) Solution for Injection. The material complies with Ph.Eur. and EC requirements. The choice of the container closure system has been validated by stability data and is adequate for the intended use of the product.

Manufacture of the product and process controls

The finished product is manufactured by two manufacturing sites.

The active substance (flortaucipir (^{18}F)) and finished product (Flortaucipir (^{18}F) Solution for Injection) are manufactured in one continuous process in which the active substance is produced using an automated radiosynthesiser, formulated, and sterile filtered to form the finished product. The active substance is not isolated during the manufacturing process. The steps in the finished product manufacturing process begin with formulation and result in finished product dispensed in multi-dose vials. These steps, which are performed in a Class A environment, comprise sterile filtration, final formulation (dilution with sterile filtered diluent), and dispensing.

The finished product is manufactured by two manufacturing processes (Process 1 and Process 2).

The formulation step of the finished product manufacturing process is carried out using the automated radiosynthesiser programme in a closed system in a Class C shielded enclosure.

Sterile filtration, dilution, and dispensing are carried out aseptically in a Class A environment.

Major steps of the manufacturing process have been validated by a number of studies in process validation batches at all manufacturing sites. It has been demonstrated that the manufacturing process is capable of producing the finished product of intended quality in a reproducible manner. The in-process controls are adequate for this type of manufacturing process.

Product specification

The finished product release specifications include appropriate tests for this kind of dosage form: appearance (visual), pH, radiochemical purity (radiometric HPLC), radiochemical impurities (radiometric HPLC) radiochemical impurities (radiometric HPLC), chemical impurities (UV HPLC), radiochemical identity (radiometric, UV HPLC), radionuclidic identity (radioactive half-life, MCA),

Flortaucipir (^{19}F) concentration (UV HPLC), strength (radioactivity), assay of ethanol (GC/FID), cryptand 222 (colorimetry), residual solvents (GC/FID), bacterial endotoxins (Ph. Eur.), filter integrity (bubble – point test), sterility (Ph. Eur.) radionuclidic purity (Gamma spectroscopy).

Chemical impurities are very low. The main radiochemical impurity to be expected is [^{18}F]fluoride.

The specification of 99.9 % radionuclide purity of radionuclides with a longer physical half-life than fluorine-18 is a general established limit. Because the strong radiation deriving from fluorine-18 will jam the gamma - detector a measurement of the low amounts of radionuclide impurities is only possible after the decay of fluorine-18 and by this after the release and patient use of the finished product. Reasonable amounts of fast decaying impurities, as for example the PET radionuclide nitrogen-13 with a physical half-life of 10 minutes would be detected by the radionuclide identity test on fluorine-18 by half–life.

The CHMP requested that a risk assessment in view of nitrosamine impurities (according the document "Questions and answers for marketing authorisation holders/applicants on the CHMP Opinion for the Article 5(3) of Regulation (EC) No 726/2004 referral on nitrosamine impurities in human medicinal products" (EMA/409815/2020 Rev.15)) should be provided for the cold chemical precursor, the active substance and the finished product. The applicant provided a risk assessment concerning the potential presence of nitrosamine impurities considering all suspected and actual root causes in line with the "Questions and answers for marketing authorisation holders/applicants on the CHMP Opinion for the Article 5(3) of Regulation (EC) No 726/2004 referral on nitrosamine impurities in human medicinal products" (EMA/409815/2020) and the "Assessment report- Procedure under Article 5(3) of Regulation EC (No) 726/2004- Nitrosamine impurities in human medicinal products" (EMA/369136/2020). Based on the information provided, it is accepted that there is no risk of nitrosamine impurities in the active substance or the related finished product. Therefore, no specific control measures are deemed necessary, and the major objection is considered solved.

The analytical methods used have been adequately described and appropriately validated in accordance with the ICH guidelines. Satisfactory information regarding the reference standards used for testing has been presented.

Batch analysis results are provided for 3 commercial scale batches confirming the consistency of the manufacturing process and its ability to manufacture to the intended product specification.

The finished product is released on the market based on the above release specifications, through traditional final product release testing.

Stability of the product

A stability study was performed to confirm the stability of the 800 MBq/mL finished product, manufactured according to Process 1. Stability data from three commercial scale batches stored in upright multi-dose vials at controlled room temperature (20 – 25°C) for up to 7.5 hours post end-of-synthesis (EOS) and elevated temperature (40°C) for up to 2 hours (post storage). There were no significant changes in the test results over the course of the study. Hence, this study demonstrated that the finished product manufactured at 800 MBq/mL according to Process 1 is stable for at least 7.5 hours when stored at room temperature.

Another stability study was performed to confirm the stability of the 1900 MBq/mL finished product, manufactured according to Process 2. Stability data from three commercial scale batches stored in upright multi-dose vials at room temperature for up to 10 hours post end-of-synthesis (EOS) and elevated temperature (40°C) for up to 2 hours (post storage). There were no significant changes in the test results over the course of the study. Hence, this study demonstrated that the finished product

manufactured at 1900 MBq/mL according to Process 2 is stable for at least 10 hours when stored at room.

800 MBq/mL and 1900 MBq/mL formulations. Finished product from three batches were manufactured and stored in 5 mL syringes with a needle. The finished product was then stored at -20°C until it was visually confirmed frozen (approximately 2-3 h). Once allowed to thaw, the QC tests to indicate stability and confirm composition were repeated. There were no significant changes in the test results over the course of the study. Hence, this study demonstrated that the finished product is stable following a freeze-thaw cycle.

An exaggerated stress stability study was performed to measure the effect of hydrogen peroxide, sodium hydroxide (NaOH), hydrochloric acid (HCl), high temperature (90°C) and full spectrum light on the radiochemical stability of the finished product. Considering the similarities between the two strength formulations, 800 MBq/mL formulation (Process 1) was used in the study to represent both800 MBq/mL and 1900 MBq/mL formulations. The finished product under the stress conditions was assessed in terms of radiochemical purity. There was no significant change in radiochemical purity of the finished product.

An in-use stability study after dilution with 0.9% sodium chloride aqueous solution was performed.

Independent of the strength at EOS, 800 MBq/mL or 1900 MBq/mL, the strength at the time of use by the end-user is dependent on the time of dose administration. Therefore, the study was designed to evaluate finished product at a maximum strength that may be obtained by an end-user to represent the worst case from a radiolysis standpoint. The finished product was diluted with 0.9% sodium chloride aqueous solution and tested over a 3-hour period. No precipitation was observed by visual examination of all diluted samples.

Stability of the finished product after first opening was evaluated in two different studies. The first study evaluated the physical and chemical stability of the product, and the second study evaluated the microbiological quality of the product.

Hold time studies were carried out to evaluate the effect of the simulated, non-radioactive finished product solution diluted 1:50 with 0.9% aqueous sodium chloride solution on mesophilic Ph. Eur. 2.6.12 indicator microorganisms with the addition of E. coli. The results of the studies demonstrate that simulated finished product samples diluted 1:50 with 0.9% sodium chloride aqueous solution can be held at 2-8°C for up to 8 days before sterility testing without loss of viability of microorganisms.

No significant changes were observed in any of the stability indicating parameters. The stability of finished product at both strengths was confirmed following successive punctures of the vial and throughout its respective shelf-life in all tested sample vials.

A stability study after first use and throughout its shelf-life of up to 10 hours after successive use was performed in the second study. After 10 hours, each sample vial was analysed for appearance, endotoxin, and sterility. All samples met the microbiological acceptance criteria. This study confirms the microbiological quality of the product following successive punctures and withdraws from each vial in an aseptic environment.

Based on available stability data, the proposed shelf-life of up to 7.5 hours from ToC for 800 MBq/mL strength and up to 10 hours from ToC for 1900 MBq/mL strength at 25°C without any special storage conditions as stated in the SmPC (sections 6.3 and 6.4) are acceptable.

Product diluted according to the preparation must be used within 3 hours of dilution and prior to the radiopharmaceutical expiry, whichever is soonest. From a microbiological point of view, unless the method of opening or dilution precludes the risk of microbial contamination, the product should be

used immediately. If not used immediately, in-use storage times and conditions are the responsibility of the user.

Adventitious agents

No excipients derived from animal or human origin have been used.

2.4.4. Discussion on chemical, pharmaceutical and biological aspects

Information on development, manufacture and control of the active substance and finished product has been presented in a satisfactory manner.

During evaluation, one major objection was raised by the CHMP in relation to risk assessment of nitrosamines. The responses from the applicant to this major objection were considered satisfactory and all the issues were considered to be resolved, as explained above.

The results of tests carried out indicate consistency and uniformity of important product quality characteristics, and these in turn lead to the conclusion that the product should have a satisfactory and uniform performance in clinical use.

2.4.5. Conclusions on the chemical, pharmaceutical and biological aspects

The quality of this product is considered to be acceptable when used in accordance with the conditions defined in the SmPC. Physicochemical and biological aspects relevant to the uniform clinical performance of the product have been investigated and are controlled in a satisfactory way.

2.4.6. Recommendation(s) for future quality development

Not applicable.

2.5. Non-clinical aspects

2.5.1. Introduction

Pivotal studies on safety pharmacology and toxicity were performed in accordance with Good Laboratory Practice (GLP) regulations and were consistent with the Organisation for Economic Cooperation and Development (OECD) standards in effect at the time. Other pharmacodynamic and pharmacokinetic studies do not fall under these regulations. Applicable ICH and EMA guidance documents were also consulted.

Studies to demonstrate the nonclinical pharmacology and safety of F18-flortaucipir (LY3191748) were designed to be appropriate for the expected clinical utilisation of the agent, i.e., a rapidly excreted diagnostic PET reagent, which is intended for microdose administration to elderly humans and are in accordance with the draft *Guideline on the non-clinical requirements for radiopharmaceuticals* (EMA/CHMP/SWP/686140/2018; November 2018) and *Microdose Radiopharmaceutical Diagnostic Drugs: Nonclinical Study Recommendations Guidance for Industry* (FDA; August 2018). This testing meets or exceeds the current non-clinical safety and toxicity recommendations to support human clinical trials and marketing applications of a microdose radiopharmaceutical diagnostic drug (Nonclinical Safety Studies for the Conduct of Human Clinical Trials and Marketing Authorization for Pharmaceuticals *ICH M3(R2)*).

2.5.2. Pharmacology

2.5.2.1. Primary pharmacodynamic studies

The applicant has conducted 3 in-vitro studies to show F18-flortaucipir's ability to bind to paired helical filaments of human origin. Studies TR-AV1451-013 and TR-AV-1451-156.00 used immunoaffinity purification isolated - tau aggregates from human brains affected by AD as described by Jicha et al. 1999. The presence of phosphorylated tau protein was confirmed in these studies by a MC-1 and AT8 ELISA assay, which was reported to recognise phosphorylated tau at both Ser 202 and Thr 205 (Goedert et al. 1995). In these studies, a dissociation constant K_d of 0.7 and 0.57 nM respectively, was reported. In study TR-AV-1451-173.00, human brain slices with confirmed AD were used. PHF tau rich regions were characterised by immunohistochemistry using Mouse anti-human PHF-tau AT100 and rabbit anti-human Aβ42 antibodies. Binding of various concentrations of F18-flortaucipir to serial slices with PHF tau rich and PHF tau poor regions were compared to each other. Preferred binding of F18-flortaucipir to PHF tau rich brain regions in comparison to tau poor regions could be shown. A Kd of 4.5 nM was determined by non-linear regression. The applicant provides further two published studies investigating F18-flortaucipir's pharmacodynamic activities (Chien et at al 2013, Xia et al. 2013). The focus of the paper of Chien et al 2013 is clearly on the clinical imaging and non-clinical results are only described in passing. However, it becomes clear that F18-flortaucipir binds especially to PHF-tau rich brain sections. Unfortunately, immunohistochemistry (IHC) staining methods were not further described. No such effects were shown for samples with low/negative tau but β-amyloid positive regions. The Xia et al (2013) study also assessed the correlation in AD brain sections between F18-flortaucipir autoradiography signal and immunohistochemistry signals for PHF-tau and amyloid beta, respectively. Here, the r2 for PHF-tau was determined at 0.90, while for β-amyloid no specific binding was evident, r2 being 0.08. This result further demonstrates the high selectivity of F18-flortaucipir for PHF-tau over amyloid beta.

Because the compound did not detect tau protein in imaging studies on 'humanised' transgenic mouse models, no *in vivo* PD studies could be carried out. This is acceptable and in line with ICH M3 guideline. The applicant refers very briefly to the papers of Marquie et al. 2015; Xia et al. 2013 and Brendel et al. 2018. Marquie et al 2015 and Brendel et al. 2018 investigated transgenic mice expressing mutant tau protein, however not of the paired helical filament type. No binding in the straight filament tau type (rTg4510) mice could be found. However, some binding could be demonstrated in the P301 mice that carry tau filaments, which appear mostly as half-twisted ribbons in the brainstem. Larger paired helical tau filaments reminiscent of those seen in human AD are observed less frequently. In wildtype mice as used by Xia et al. (2013) no specific binding could be observed.

The proportion of radioactive F18-flortaucipir amounts to approximately 0.008% of the total F19-flortaucipir in a human dose. It is acknowledged that this is a very small proportion. Nevertheless, due to the short half-life of ~110 minutes, appreciable amounts of 18O- are generated during the β+ decay process. A likely follow on reaction would be for the $^{18}O^-$ to react with a proton (H^+) to become a hydroxyl group (OH).

2.5.2.2. Secondary pharmacodynamic studies

Several studies on secondary pharmacodynamics investigating effects on α-synuclein, MAO-A, MAO-B, and selectivity of F18-flortaucipir for AD tau relative to PSP, PiD and CTE, were provided.

α-synuclein:

F18-flortaucipir was tested (Study TR-AV-1451-074.01) on a panel of fixed brain tissue samples from α-synucleinopathies including Dementia with Lewy Bodies (DLB), Parkinson's disease (PD), Parkinson's disease dementia (PDD), Multiple System Atrophy (MSA), as well as brain tissues from age-matched non-demented patients (CTL). Signal strength was classified by visual inspection of slides under a microscope on a relative scale (-/+/++/+++) and compared to a panel of immunohistochemistry signals from disease protein specific antibodies and a beta amyloid PET tracer, 18F AV-45 (florbetapir). The study shows, that in αSYN rich slices of human brain tissues F18-flortaucipir binds preferentially to tissues with phosphorylated tau but not to α-synuclein rich tissues devoid of pathological tau proteins.

MAO-A:

Due to the structural similarity between ^{18}F-AV-1451 and known reversible monoamine oxidase-A (MAO-A) ligands, including harmine and ^{18}F-fluoroethyl harmol (^{18}F-FEH), the binding affinity of ^{18}F-AV-1451 to MAO-A was evaluated in-vitro (TR-AV-1451-162.00) and in-vivo (TR-AV-1451-161.00).

The in-vitro study investigating the binding of F18-flortaucipir to recombinant MAO-A F18-flortaucipir showed a rather high affinity to MAO-A with a K_d of 2.0 nM. Deeper analysis investigating association rate constant (k_{on}) and dissociation rate constant (k_{off}) showed that F18-flortaucipir has a rather high dissociation rate constant compared to other MAO-A substrates such as ^{18}F-FEH and to the binding target PHF. As a result, F18-flortaucipir may associate preferentially with PHF within the brain and not with MAO-A. Interestingly, the experimentally determined Kds in this kinetic binding assay are 5.2 and 6.7 times lower for Tauvid:PHF (0.57 nM/ 0.11 nM) and Harmol:MAOA (1.2 nM/ 0.18 nM) than in the saturation binding assay, but identical for Tauvid:MAO-A (2.0 nM/ 1.97 nM) (values from study report 162.00). In response to a request the applicant argued that the difference in the calculated Kd is due to the slow ligand off rates of Tauvid:PHF and Harmol:MAOA. The applicant's line of argumentation is in line with the overall conclusion from secondary Pharmacodynamics studies.

The in-vivo study in rats showed a slightly higher C_{max} of F18-flortaucipir in the brain than in the muscle (which may be explained by perfusion differences between brain and muscle) and nearly identical concentrations in both tissues after approximately 20 minutes for the rest of the observation period of 2 hours. No retention within the brain could be shown. Pre-treatment with pargyline did not change the Cmax or the retention within rat brain or muscle.

MAO-B:

In studies investigating specific binding to MAO-B (TR-AV-1451-189, TR-AV-1451-163.00, TR-AV-1451-179.00), no specific binding could be detected. This finding was considered rather conflicting by the applicant, since rather high affinities of F18-flortaucipir to MAO-B has been reported in the past (Vermeiren et al. 2018). However, the applicant´s investigations provided evidence, that a possible adherence of F18-flortaucipir to assay filter plates may have caused the reported high affinity of F18-flortaucipir to MAO-B in the study of Vermeiren et al. It is agreed to, that the in-vitro results show a rather weak binding of F18-flortaucipir to MAO-B. Based on this finding, no significant binding in-vivo is expected from the non-clinical point of view.

Selectivity of F18-flortaucipir for AD tau relative to PSP, PiD and CTE (TR-AV-1451-180):

The study *TR-AV-1451-180* shows that F18-flortaucipir has the ability to bind to tau deposits in the non-AD tauopathies PSP, PiD and CTE as shown by immunofluorescence methods against hyperphosphorylated-tau AT8 in postmortem brain sections.

Although AD slices were clearly identified by F18-flortaucipir, no discrimination between PID, PSP and CTE in comparison to AD slices was possible, since some weak binding was detected. In AD brain sections, the autoradiography (ARG) signal was very strong in the frontal and temporal lobe areas, the

cingulate and hippocampus, and vastly overlapped with the immunofluorescence (IF) signal using AT8 monoclonal antibody, recognising phosphorylated tau. In PiD and PSP tissue samples, in contrast, the ARG signal was of weaker intensity and overlapped significantly less with the IF signal, which was also slightly reduced overall, in comparison to AD tissues analysed.

The applicant provided further references, discussing the differences of tau-deposits in the different non-AD tauopathies. From the non-clinical point of view, the relevance appears to be rather low, since the applicant´s studies show that F18-flortaucipir may detect AD (as defined as neurofibrillary tangles which can be isolated by MC1 antibodies). However, a reliable discrimination between PID, PSP and CTE in comparison to AD appears to be unlikely. The applicant refers in the following to clinical studies investigating the use of F18-flortaucipir in chronic traumatic encephalopathy, which is out of the scope of this application. It is important to note that the study report does not specifically mention if all IF pictures were taken using identical microscope settings. Quantitation of both ARG and IF signals occurred by visual scoring. Although not a perfect system, this strategy is commonly used by other experts in the field.

2.5.2.3. Safety pharmacology programme

CNS:

The applicant provided a non-GLP compliant study (TR-AV-1451-007), investigating a panel of 72 common CNS targets. Only monoamine transporters and MAO-A showed binding and inhibition in a high nanomolar range. All other targets investigated showed a lower binding and inhibition. The applicant estimates a maximum brain concentration of 4 nM, which would equal at least 100 times less than the IC_{50} of the monoamine transporter and MAO-A. The applicant´s view that the potential for adverse effects due to inhibition of the investigated targets is low is shared.

F18-flortaucipir´s effects on CNS parameters were further investigated in a GLP compliant dedicated conventional safety pharmacology study (8286450) in male rats. No effects on the investigated parameters were observed up to the highest dose tested, which equals approximately 100 times the maximum human dose based on a human equivalent dose. The applicant´s conclusions are therefore agreed to. No effects on the CNS in humans are anticipated.

Cardiovascular/ hERG:

The applicant provided initially a non-GLP compliant study (130812.FMD) investigating F18-flortaucipirs potential ability to block hERG potassium channels. In this assay, F18-flortaucipir had the ability to block potassium currents in a high nanomolar range. A second GLP-compliant study, conducted as response to a request, confirmed that finding. In this study F18-flortaucipir blocks hERG potassium channels with an IC_{50} of 0.61 µM (161 ng/ml), which results (based on free plasma concentrations) in a safety margin of 338-fold for clinical use.

The evaluation of in-vivo cardiovascular safety was part of this repeated dose toxicity study in dogs (8286448). The NOEL for Day 1 of this study was determined to be at least 60 µg/kg (100x MHD, allometrically scaled), the highest dose level tested and on Day 29 the NOEL in male dogs and No Observed Adverse Effect Level (NOAEL) in female dogs, respectively, was determined to be 30 µg/kg (50x MHD, allometrically scaled).

Respiratory

F18-flortaucipir´s effects on respiratory parameters were investigated in a GLP compliant dedicated plethysmography study in male rats. No effects on respiratory parameters were observed up to the

highest dose tested, which equals approximately 100 times the maximum human dose based on a human equivalent dose. The applicant´s conclusions are therefore agreed to.

The applicant presents furthermore a non-GLP compliant whole-body PET/CTE dosimetry study in rhesus monkeys (TR-AV-1451-009), which appears to be somewhat unrelated to safety pharmacology. However, it becomes clear which organs may be mainly exposed to radiation during clinical use. Prominently, lungs during the first minutes after administration and later the urogenital system with the bladder wall as most exposed tissue followed by the small intestine, which showed however a great variability. The pattern seems to reflect the elimination of the ^{18}F-label out of the organism and no retention could be observed, which would require further safety pharmacology studies.

2.5.2.4. Pharmacodynamic drug interactions

No pharmacodynamics drug interactions were investigated, which is considered acceptable.

2.5.3. Pharmacokinetics

F18-flortaucipir had been administered intravenously to mice, rats, and monkeys to characterise its distribution, metabolism, and excretion properties. In all species PET scans were performed, and in mice, in addition to PET scans, radioactivity was determined in organ and tissue samples using a gamma counter.

Following IV administration, F18-flortaucipir rapidly distributed to the brain in all species evaluated. In mice and monkeys, species in which presence of F18-flortaucipir in several organs and tissues was investigated, F18-flortaucipir distributed to multiple tissues, including, but not limited to, brain, liver, and kidney.

In vitro, the plasma protein binding of the stable isotope, F19-flortaucipir, was generally low (about 81 to 91%) in all nonclinical species evaluated. In humans, plasma protein binding was slightly higher (94,7%).

The *in vitro* metabolism of F19-flortaucipir suggests that the compound is cleared primarily by oxidation and direct glucuronidation. Structural information on the identified metabolites is limited, but based on this limited information, the metabolites observed in human hepatocytes were also observed in at least one non-clinical species, although the relative amounts of individual metabolites identified vary to partially large extents between hepatocytes of non-clinical species and humans, respectively.

In mice, F18-flortaucipir was the predominant circulating entity in plasma, along with 3 additional metabolites that were smaller and more hydrophilic compared to parent, based on peak intensity with radio-detection. In brain homogenates, essentially all of the radioactivity was attributed to F18-flortaucipir, indicating that only parent compound distributes to the brain and that metabolites do not contribute to the PET signal.

There is limited excretion data for F18-flortaucipir; however, data from *in vivo* mouse and monkey studies suggests that renal elimination is a primary route of elimination, however excretion in the faeces cannot be excluded.

The bioanalytical methods employed in nonclinical studies were not validated to GLP standards. In response to a request, the applicant has provided details of the assays run in the PK studies. These assays were performed in accordance with standard protocols. The methods employed are considered appropriate taking into account the small dose and physical characteristics of the medicinal product.

2.5.4. Toxicology

2.5.4.1. Single dose toxicity

Single dose toxicity of F19-flortaucipir was tested in rats with doses up to 150x MHD (allometrically scaled). The NOAEL was considered to be the highest dose tested.

2.5.4.2. Repeat dose toxicity

The potential toxicity of F19-flortaucipir was tested in rats with a 1 month of daily repeated dosing (up to 50x MHD, allometrically scaled).

In beagle dogs, toxicity was assessed for 1 month of daily repeat dosing (up to 50x MHD, allometrically scaled). In all of these studies the NOEL was determined to be the highest dose level tested, with the exception of the female dog in the repeat dose toxicity study.

In this study, the NOEL was 15µg/kg (25x MHD, allometrically scaled) and the NOAEL was determined to be at 50x MHD, allometrically scaled, based on the transient moderate increase in heart rate observed at Day 29 at 30µg/kg.

2.5.4.3. Genotoxicity

A standard battery of genotoxicity tests was performed with AV-1451 in line with ICH S2(R1). AV-1451 was positive in the Ames test and in an *in vitro* Chromosome aberration assay with and without metabolic activation and can thus be considered as directly mutagenic and clastogenic *in vitro*. A subsequent *in vivo* Micronucleus Test AV-1451 was negative up to the highest test concentration at exposures 750-fold the MRHD based on allometric scaling. As the *in vivo* Micronucleus test mainly detects chromosome aberrations but not necessarily point mutations, it cannot be excluded that AV-1451 is mutagenic *in vivo*. However, flortaucipir will be administered at a low single dose of 0.33 µg/kg (16.5 µg/50 kg patient). This is still below the less-than-lifetime (LTL) TTC of 120 µg/day for a treatment duration ≤ 1month for mutagenic impurities according to ICH M7(R1). Thus, the lifetime cancer risk of a single dose of flortaucipir is far below the acceptable 1:100.000 excess cancer risk accepted for mutagenic impurities.

2.5.4.4. Carcinogenicity

Carcinogenicity studies with AV-1451 were not performed in line with Guidance for Industry Developing Medical Imaging Drug and Biological Products, Part 1: Conducting Safety Assessments, FDA, June, 2004 and the draft Guideline on the non-clinical requirements for radiopharmaceuticals (EMA/CHMP/SWP/686140/2018). The sponsor has also received a waiver from FDA for conducting carcinogenicity studies.

Based on the nature of AV-1451 being a (single) microdose radiodiagnostic agent for elderly patients and in line with current guidelines the lack of carcinogenicity studies is acceptable.

2.5.4.5. Reproductive and developmental toxicity

No studies on reproductive and developmental toxicity have been conducted. Given the indication, this is considered acceptable.

2.5.4.6. Toxicokinetic data

N/A

2.5.4.7. Local tolerance

No dedicated local tolerance studies have been performed. Instead, the injection site was part of the clinical examination in the repeat-dose toxicity studies and it was concluded that F19-flortaucipir does not contribute to local effects at the injection site.

2.5.4.8. Other toxicity studies

In-vitro cytotoxicity results demonstrate that F19-flortaucipir has no effect on cell viabilities for the four cell lines tested.

2.5.5. Ecotoxicity/environmental risk assessment

An ERA Phase I for the active ingredient flortaucipir (^{18}F) and flortaucipir was provided.

Two independent, experimental logK_{ow} studies were provided. Both are not very detailed and not conducted under GLP (Table 3). Nevertheless, in this particular case, even though each study alone would not be acceptable to waive the PBT screening, since both studies independently resulted in de facto the same log D_{ow} at pH 7.4 of 3.28 and 3.3±0.1, in combination the results are convincing enough to do so.

For the $PEC_{surfacewater}$ calculation the focus is on flortaucipir, since flortaucipir (^{18}F) will quickly be degraded by radioactive decay. The maximal dose of flortaucipir is 20 µg/inh/day. For the F_{pen} refinement one treatment per year and a prevalence of dementia of 2 % is assumed. The resulting $PEC_{surfacewater}$ is with 5.5×10^{-7} µg/L below the action limit of 0.01 µg/l for a phase II assessment. According to this the ERA can stop in Phase I.

Table 3. Summary of main study results

Substance (INN/Invented Name): Flortaucipir (^{18}F)			
CAS-number (if available): 1522051-90-6			
PBT screening		Result	**Conclusion**
Bioaccumulation potential- log K_{ow}	OECD107 (not GLP) Additionally available: potentiometric (pH-metric) Study (no GL; not GLP)	log D_{ow} (pH 7.4) = 3.3±0.1 log K_{ow} = 3.79 log D_{ow} (pH 5) = 1.10 log D_{ow} (pH 7) = 2.98 log D_{ow} (pH 7.4) = 3.28 log D_{ow} (pH 9) = 3.76 pK_s = 7.74 (Pyridin-N)	Potential PBT (N)
Phase I			
Calculation	**Value**	**Unit**	**Conclusion**
$PEC_{surfacewater}$, refined (prevalence and treatment regime)	5.5×10^{-7}	μg/L	> 0.01 threshold (N)
Other concerns (e.g. chemical class)	radioactive		(Y)

2.5.6. Discussion on non-clinical aspects

The applicant has conducted in-vitro studies showing that F18-flortaucipir is able to bind to isolated paired helical filaments of human origin and to the respective sections of brain slices of humans with confirmed AD. The applicant has not conducted own in-vivo studies due to the lack of an appropriate animal model. Published investigations in different animal models showed conflicting results. However, the models investigated in the public domain do not always develop the molecular target of F18-flortaucipir. Overall, the applicant´s primary pharmacodynamics studies are considered sufficient for the indication applied for.

Several studies on secondary pharmacodynamics investigating effects on α-synuclein, MAO-A, MAO-B, and selectivity of F18-flortaucipir for AD tau relative to PSP, PiD and CTE, were provided. F18-flortaucipir did not bind to α-synuclein, MAO-A, and MAO-B to a relevant magnitude. It becomes clear that F18-flortaucipirs is not able to detect other tauopathies (PID, PSP and CTE) than AD.

The core battery of safety pharmacology has been investigated in GLP compliant studies according to Guideline ICH S7A. The only noteworthy finding is that F18-flortaucipir blocks hERG currents in the low nanomolecular range. However, the safety margin of 338-fold between *in vitro* IC_{50} and clinical exposure to unbound drug substance is considered high enough to conclude that a torsadogenic potential is unlikely.

No pharmacodynamics drug interactions were investigated, which is considered acceptable.

The applicant performed PET scans in mice and monkeys in order to draw conclusions on the distribution of flortaucipir. The applicant also performed an in-vivo metabolite screening in mice.

Flortaucipir $PEC_{surfacewater}$ value is below the action limit of 0.01 μg/L. and is not a PBT substance as log K_{ow} does not exceed 4.5.

Therefore, flortaucipir is not expected to pose a risk to the environment.

Assessment of paediatric data on non-clinical aspects

N/A

2.5.7. Conclusion on the non-clinical aspects

The non-clinical programme is adequate to support marketing authorisation.

2.6. Clinical aspects

2.6.1. Introduction

GCP aspects

The Clinical trials were performed in accordance with GCP as claimed by the applicant.

The applicant has provided a statement to the effect that clinical trials conducted outside the Community were carried out in accordance with the ethical standards of Directive 2001/20/EC (clinical trials – 18F-AV-1451-A16, 18F-AV-1451-A05, 18F-AV-1451-FR01 and 18F-AV-1451-PX01 i.e., key trials supporting this MAA - conducted outside the European Union; USA and Australia).

A total of 2 clinical studies in this application were inspected by regulatory authorities (Table 4):

Table 4: Clinical studies inspected by Regulatory Authorities

Study Identifier	Site of Inspection	Date of Inspection	Regulatory Authority That Conducted Inspection
^{18}F-AV-1451-A05	Covance (formerly known as Chiltern)	02 Mar 2020	US FDA
^{18}F-AV-1451-A16	Banner Sun Health Research Institute	16 Mar 2020	US FDA

The provided information on the outcome of these inspections does not raise a concern.

2.6.2. Clinical pharmacology

2.6.2.1. Pharmacokinetics

Two formulations of flortaucipir were used in clinical development and the second formulation has also been used in Alzheimer's disease (AD) therapeutic development studies and other biomarker studies of natural disease progression. The first early clinical formulation was used in the clinical development through Phase 2 and discontinued before flortaucipir phase 3 studies were initiated. The second or commercial formulation 1 was used in phase 1 to 3 clinical development and is currently planned for commercial use for the 800-MBq/mL strength; it is formulated in a solution containing 10% (v/v) ethanol in 0.9% sodium chloride injection. Furthermore, a third formulation or commercial formulation 2 was developed following completion of the Phase 3 studies; it differs from commercial formulation 1 by the addition of dibasic sodium phosphate as buffer due to the acidity of the purification mobile phase. This second commercial formulation is currently planned for commercial use of the 1900-MBq/mL strength. According to the applicant, both formulations in use during flortaucipir clinical development were manufactured to the same specifications and used a common reference standard; thus no comparison and analyses of results across studies were performed.

The biodistribution, plasma levels, and metabolism of flortaucipir were studied *in vitro* in human biomaterials and *in vivo* in animal models and in humans. The *in vitro* pharmacology studies using human biomaterials and studies in non-human species conducted to assess the pharmacology, PK, and toxicology of flortaucipir are summarised in the non-clinical modules.

Study 18F-AV-1451-A01 (A01), titled 'An Exploratory Evaluation of the Tau Protein Binding Properties, Whole-Body Biodistribution and Safety of 18F-AV-1451 Injection in Healthy Volunteers and Cognitively Impaired Subjects' included a whole-body dosimetry cohort of 9 healthy volunteers (HV) to establish the radiation-effective dose equivalent of flortaucipir. This phase I study was conducted between 16 August 2013 and 09 December 2013 at two centres in the USA.

Study 18F-AV-1451-A15 (A15), titled 'A Multicenter Study to Obtain Data to Assist in Radiation Dosimetry Calculations from Subjects Enrolled in Human Studies Employing 18F-AV-1451' evaluated urine excretion data from 6 HVs enrolled in the Phase 2/3 Study 18F-AV-1451-A05 (A05) to augment radiation dosimetry calculations from Study A01. This phase I study was conducted between 20 January 2015 and 02 June 2015 at one centre in the USA.

Study 18F-AV-1451-A10 (A10), titled 'Evaluation of 18F-AV-1451 Kinetic Modelling in Patients in Alzheimer's Disease and Healthy Controls evaluated tracer kinetic models for quantifying specific binding of flortaucipir. Subjects with AD and healthy control subjects were equally distributed across 2 cohorts; healthy controls were included to study the normal, nonspecific binding patterns of flortaucipir. Subjects in Cohort 1 received a flortaucipir PET scan at baseline and again approximately 1 year later, to investigate the effect of passage of time on measures of tau pathology, and to determine whether the kinetics of the tracer are reproducible between the 2 visits, 1 year apart. Subjects enrolled in Cohort 2 received a single flortaucipir PET scan. All subjects were scanned after a 240-MBq dose and had a dynamic 60-minute PET scan immediately following injection; at approximately 80 minutes post injection, an additional dynamic PET scan was obtained for 50 minutes. For quantitative analyses, SUVrs for 80- to 100-minute scan window were expressed as a ratio of an AD-targeted global composite region referred to as AD-signature weighted neocortical composite (Devous et al. 2018) over a subject-specific, white matter reference region referred to as PERSI (Southekal et al. 2018). Study A10 also included evaluations of flortaucipir metabolism and blood clearance from plasma. This phase II study was conducted between 13 August 2015 and 28 April 2017 at a single centre in the Netherlands.

For a tabular summary of studies evaluating PK, biodistribution, and metabolism of flortaucipir see the Table 5 below.

Table 5: Summary of Clinical Pharmacokinetic Studies

Study ID	Study Drug Dose, Route, Frequency, Image Acquisition	Study Objectives	Subjects Entered and Cohort Populations	Inclusion Criteria	Clinical Pharmacology Results
18F-AV-1451-A01	370 MBq, IV, single dose Cohort 1: Dynamic images collected from 0 to 60, 80 to 100, and (optional) 110 to 130 minutes post-dose Cohort 2: 10 emissions scans over 6 hours	• To address feasibility of further development of flortaucipir ^{18}F as a tau protein-targeted radio-pharmaceutical by: - evaluating brain uptake and retention - evaluating safety and dosimetry - comparing brain uptake and retention of flortaucipir ^{18}F to beta-amyloid status, cognitive function, and MRI scan	Cohort 1: N=15 n=4 Group 1 HV n=4 Group 2 HV n=3 Group 3 MCI n=4 Group 4 AD Cohort 2: N=9 WBD HV Cohort 3: N=6 MRI/amyloid extension from Study T807000	HV: Age ≥20 to ≤40 years (Group 1) or ≥65 years (Group 2), MMSE >29 AD: Age ≥50 years, MCI due to AD (Group 3) or possible/probable AD dementia (Group 4) per Alzheimer's Association guidelines MRI/amyloid extension cohort: subjects imaged in Study T807000	• Quantitative assessment of global composite SUVr relative to entire cerebellum showed higher flortaucipir ^{18}F uptake in AD/MCI compared with CN HV • Body region receiving highest mean flortaucipir ^{18}F dose was upper large intestine wall • Significant correlation between flortaucipir ^{18}F and florbetapir ^{18}F SUVr for both composite and individual regions
18F-AV-1451-A15	370 MBq, IV, single dose No PET scan collected	• To obtain urinary excretion data from subjects who were administered flortaucipir ^{18}F in Study A05, to augment the calculation of radiation dosimetry estimates determined in Study A01	N=6 HV	HV: Passed screening for Study A05	• Mean urine cumulated activity of 0.141 (0.072) MBq-hr/MBq • Mean (SD) body effective dose per modified organ source data of 0.0235 (0.0016) mSv/MBq • Estimated effective dose of 8.70 mSv for anticipated 370 MBq (10 mCi) injection • Body region receiving highest mean dose was upper large intestinal wall
18F-AV-1451-A10	240 MBq, IV, initial single dose Cohort 1 had a follow-up dose (240 MBq IV) and PET scan 1 year after initial dose/scan Cohort 2 received the single dose and PET scan Dynamic images collected from 0 to 60 and 80 to 130 minutes post-dose	• To evaluate tracer kinetic models for quantifying specific binding of flortaucipir ^{18}F in cross-sectional and longitudinal applications • To evaluate simplified methods for quantitation of flortaucipir ^{18}F uptake	Cohort 1: N=10 n=5 AD n=5 HV Cohort 2: N=10 n=5 AD n=5 HV	HV: Age ≥50 years, no evidence of cognitive impairment as per cognitive neurologist AD: Age ≥50 years, probable AD dementia, Amsterdam Dementia Cohort, MMSE ≥18, amyloid-positive biomarker data	• 2TC kinetic model provided a good fit to the data, adequately separated subjects with AD vs. CN HV • Flortaucipir was gradually metabolised, with parent flortaucipir accounting for ~86% of plasma radioactivity at 5 min post-dose, ~34% at 80 min post-dose, and ~22% at 130 minutes post-dose • 2 metabolites detected in HPLC/methanol-soluble fraction (Fraction 3) accounted for 30% to 35% of plasma radioactivity ≥80 min post-dose • SUVr based on acquisitions between 80 and 100 min post-dose correlated well with 2TC distribution volume ratio (r^2 and slope approaching 1) using PERSI reference region

Abbreviations: 2TC = 2-tissue compartmental model; Aβ = beta amyloid; AD = Alzheimer's disease; CI = confidence interval; CN = cognitively normal; CT = computerised tomography; ECG = electrocardiogram; HPLC = high-performance liquid chromatography; HV = healthy volunteer; ICC = intra-class correlation; ID = identifier; IV = intravenous; MCI = mild cognitive impairment; MMSE = Mini Mental State Examination; MRI = magnetic resonance imaging; NIA = National Institute on Aging; OCN = older cognitively normal controls; PERSI = parametric estimate of reference signal intensity; PET = positron emission tomography; PK = pharmacokinetics; SD = standard deviation; SUVr = standard uptake value ratio; WBD = whole-body distribution; YCN = younger cognitively normal controls.

Overall, the described analytical methods are considered adequate.

As regards pharmacokinetic data analysis in study A10 tracer kinetics were evaluated for the purpose of quantifying specific binding of flortaucipir and to compare the SUVr method of quantitation to full tracer kinetic modelling to determine the applicability of the simplified method to large-scale trials. Overall, the described approach is considered adequate, but could not be adequately assessed due to very limited information on model building and evaluation.

Furthermore, concerning the evaluation and qualification of models, study A10 evaluated tracer kinetic models for quantifying specific binding of flortaucipir. Subjects with AD and healthy control subjects

were equally distributed across 2 cohorts; healthy controls were included to study the normal, nonspecific binding patterns of flortaucipir. One of the primary objectives of this study was to evaluate tracer kinetic models for quantifying specific binding of flortaucipir in cross-sectional and longitudinal applications. In addition, pharmacokinetic modelling was performed to understand the *in vivo* kinetics of the tracer. This was followed by an assessment of whether SUVr can be used as a surrogate to binding potential and potential biases of the SUVr quantitation approach. Regression analyses were used to compare the results of the kinetic models to each other and to the SUVr estimate of tracer retention. For kinetic modelling the time–activity curves from the pre-specified regions were fitted using the 2-tissue compartment (2TC) model and the metabolite-corrected plasma input function, with blood volume fraction (VB) as an additional fit parameter. Standard nonlinear regression fitting to these models was performed for each ROI. Macro parameters, such as BPND, VT, and DVRs (DVR = ratio of target to reference region) were calculated for each reference region and for the composite region for the model(s) of interest. However, the A10-01 Technical Report does not include sufficient information and the model performance remains unclear. Furthermore, PMOD software was used which is meant for research use only and must not be used for diagnosis or treatment of patients. Therefore, due to the limitations the results from the analyses were not used for any conclusions related to the marketing authorisation and were not mentioned in the SmPC.

The statistical methods used in the relevant studies are considered adequate.

Absorption

Flortaucipir is administered by IV injection and therefore is fully bioavailable. Thus, no studies on bioavailability, comparative bioavailability, bioequivalence, or food-interaction were performed.

Distribution

Following an intravenous (IV) injection, flortaucipir showed a quick uptake in various body organs and tissues including brain. As reported in the application, flortaucipir is rapidly metabolised, with 86% of parent flortaucipir present in the plasma at 5 minutes post-administration and approximately 34% and 22% of the not-metabolised flortaucipir at 80 minutes and 130 minutes post-administration, respectively (whereas 4%, 66%, and 78% account for the metabolites of flortaucipir at the corresponding time-points). The applicant does not consider the rapid metabolism to be an issue due to an almost immediate uptake of flortaucipir in the brain, which allows for binding with target pathology before an adverse effect of the metabolism. The highest uptake in the brain was observed within the first 10 minutes post-administration. The fraction of the metabolised flortaucipir is notable already several minutes post-injection. The biokinetics of the metabolised flortaucipir potentially differs from that of the parent flortaucipir. This complicates the biokinetic modelling, since PET scanner detects only the activity of 18F regardless whether it is bound to the parent flortaucipir or comes from the metabolised flortaucipir. It is stated in the application that no clinically relevant changes in the pharmacokinetics of the metabolised flortaucipir compared to the parent flortaucipir are expected and no specific recommendations or limitations with respect to drug interactions have been made in the labelling.

Phase 2 study A10 evaluated the blood clearance of flortaucipir. The applicant provided a table with relative proportions of ^{18}F activity in blood and plasma attributed to the parent flortaucipir and metabolites thereof at 9 time-points (5 minutes up to ca. 2.2 hours). Phase 1 kinetic study A01 showed a rapid clearance of flortaucipir from nontarget reference regions and a slower wash-out from tau containing target regions. In clinically normal and amyloid-negative cognitively impaired subjects flortaucipir was rapidly eliminated from the brain, whereas in amyloid-positive cognitively impaired subjects a regionally distinct neocortical distribution of an increased tracer retention was observed.

Figure 3 below shows representative maximum intensity projection images of 18F activity distribution at 10 time-points after an iv administration of approximately 370 MBq (10 mCi) flortaucipir to a 58-year-old male subject (study A01).

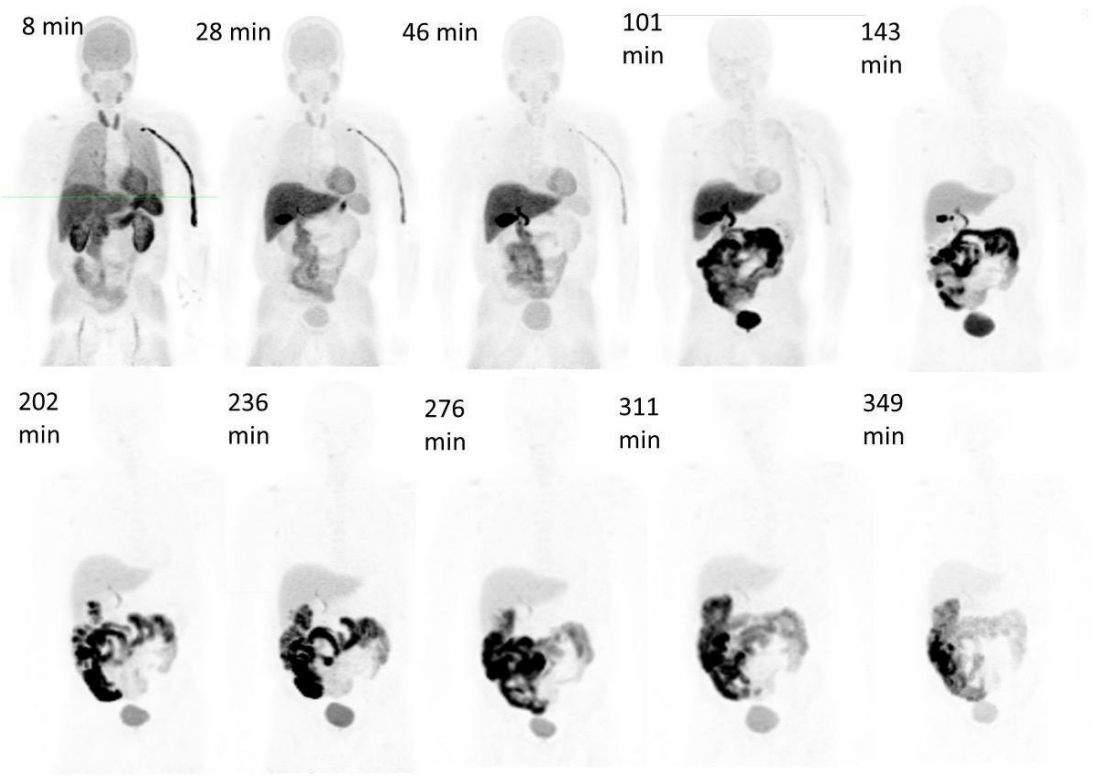


Figure 3: Maximum intensity projection images showing representative radioactivity biodistributions after intravenous administration of approximately 370 MBq (10 mCi) flortaucipir 18F to a 58-year-old male subject

The applicant has provided the time-activity data up to 343 min p.i. (not decay-corrected) for the following regions of one subject: brain, gall bladder, heart, intestines, kidneys, liver, lungs, urinary bladder and body remainder. These data for the region "intestines" represent the total intestinal excretion, as stated in the applicant's response. Additionally, the individual time-integrated activity coefficients (TIACs) for all considered source regions were reported. Initially, the TIACs in each segment of the intestines (upper large intestine (ULI) and lower large intestine (LLI)) were determined using the gastrointestinal (GI) tract model of OLINDA-EXM (version 1.1). Subsequently the Human Alimentary Tract Model (HATM) of the ICRP Publication 100 was applied to recalculate the TIACs obtained for ULI and LLI to the TIACs for right colon, left colon and recto-sigmoid colon in order to be conform with the GI structures of the anthropomorphic reference phantoms. Besides the individual TIACs, the mean values along with the standard deviations were computed across the subjects to illustrate inter-individual variability of the TIACS. The following source regions were considered in dosimetry calculations for flortaucipir: brain, gallbladder contents, left colon, small intestine, right colon, rectum, heart wall, kidneys, liver, lungs, urinary bladder contents and body remainder. The activity of 18F in the blood or plasma was not measured within the conducted biodistribution study A01. According to the applicant, this activity cannot be derived from the images either. Thus, blood/plasma was not considered as an explicit source region in dosimetry. The activity of ^{18}F in blood was attributed to the source region "remainder". Although such methodology may potentially underestimate the absorbed doses for highly vascularised organs, it can be considered acceptable. The clinical study report A10 provides relative proportions of the parent vs. metabolised flortaucipir in blood and plasma. It also contains a column "Activity in Plasma", but without the units of the reported

values. From these data it is difficult to assess the clearance of the ^{18}F activity from blood/plasma, since neither the units of the reported "Activity" nor the volumes of the collected samples are given. As reported in the application, flortaucipir was eliminated from the body via both, a hepatobiliary and a renal excretion routes. The applicant has specified that approximately 22% of flortaucipir were eliminated via the hepatobiliary tract, and the remaining 78% - via the renal excretion route.

Information on the volume of distribution has not been reported.

Protein binding of flortaucipir has been established in *in vitro* studies. In plasma it was 94.7% and in normal physiologic solutions of HSA (4%) 88.8%, which was similar to human plasma protein binding, suggesting that flortaucipir is primarily bound to albumin.

Radiation dosimetry

Distribution of ^{18}F radioactivity from iv flortaucipir administration was evaluated by whole-body PET imaging in Cohort 2 of study A01. Following completion of Study A01 procedures, in study A15 urinary excretion data were collected from 6 additional HVs enrolled in the Phase 2/3 Study A05 to consider the urinary excretion in radiation dosimetry.

As follows from the study report A01, approximately 10 PET scans were done from the vertex of the head to the thigh for each subject from the time of flortaucipir administration over a period of about 6 hours. Volumes of interest were delineated on PET images for brain, liver, lung, heart wall, vertebrae, gallbladder, urinary bladder, and large and small intestine to define the activity concentration of 18F. According to the applicant's response, the volumes of interest drawn on the images completely encompassed each entire source region.

The time-activity data derived from PET images were fit to exponential functions using OLINDA/EXM. It is stated in the applicant's response that for organs that did not follow the exponential function at the early time-points, the integration for the early time-points was done using the Riemann Sum and for the later time-points – the analytical integration of the exponential function. Otherwise, OLINDA/EXM was employed to calculate the TIACs in source regions. As mentioned above, the TIACs for right colon, left colon and recto-sigmoid colon were calculated from those for ULI and LLI using the Human Alimentary Tract Model (HATM) of the ICRP Publication 100. The TIACs for most of the source regions were derived within the study A01. As stated in the applicant's response, these values were computed under the assumption of no voiding. To account for the activity of flortaucipir excreted in urine, the mean value of TIAC for excreted urine derived within the study A15, was subtracted from the TIACs for body remainder of each subject. As recommended, the applicant has provided a table with the individual TIACs of each subject used in the final dose calculations. An inconsistency in the TIAC for brain of the subject 1002-006 (3.30E-02 MBq-h/MBq) was observed when compared to the corresponding value and in the previously submitted data (3.35E-02 MBq-h/MBq). The applicant has checked the entries and recognised that there was a transcription error when entering the TIAC data for the updated dose calculations for Subject 1002-006. Brain TIAC was thus corrected to 3.35E-02 MBq-h/MBq. Moreover, the applicant changed the name of the last source region listed, originally incorrectly indicated as "Remainder", to "Total Body".

As requested, the applicant has recalculated the absorbed organ dose coefficients using the anthropomorphic reference adult phantoms and computed the coefficient of the effective dose applying the weighting factors of the ICRP Publication 103 (2007)[1]. The resulting dose coefficients derived using the individual TIAC sets were reported for reference adult in male and female, respectively. The mean values of the absorbed organ dose coefficients for reference male and female along with the coefficient of the effective dose were summarised. Slight discrepancies between the mean values of the absorbed dose coefficients for rectum, ovaries, osteogenic cells and uterus of reference adult female reported

[1] https://www.icrp.org/publication.asp?id=ICRP%20Publication%20103

(1.88E-02, 1.64E-02, 9.69E-03 and 1.92E-02, respectively) and the corresponding values (1.83E-02, 1.62E-02, 9.67E-03 and 1.87E-02, respectively) were accounted for and corrected.

The organs that receive the highest absorbed dose coefficients as calculated by the applicant were right colon (0.1303 and 0.1298 mGy/MBq for adult male and female, respectively), small intestine (0.1039 and 0.1243 mGy/MBq for adult male and female, respectively), and liver (0.0620 and 0.0767 mGy/MBq for adult male and female, respectively). The effective dose coefficient for flortaucipir was calculated according to the formalism and the tissue weighting factors of the ICRP Publication 103 (2007). This effective dose coefficient amounted to 2.60E-02 mSv/MBq. The effective dose from the administered typical activity of flortaucipir of 370 MBq amounts, thus, to 9.6 mSv.

The applicant has proposed to keep in the EU label the dosimetry conducted using the mathematical phantoms as implemented in the software OLINDA/EXM version 1.1 and used in other global labels. EU Directive 2013/59/Euratom (paragraph 11 page 2) explicitly states that for internal exposures, the dose coefficients computed according to the earlier ICRP recommendations "will be superseded by values based on the radiation and tissue weighting factors and phantoms laid down in ICRP Publication 103." The ICRP Publication 103 clearly specified (see e.g. paragraph 130 page 69) that the evaluation of organ doses is to be based on the use of anthropomorphic computational phantoms (voxel phantoms of adult Reference Male and adult Reference Female). Thus, the EU Directive 2013/59/Euratom is to be followed. Moreover, the mathematical phantoms represent human anatomy in a simplified stylised way. The usage of the stylised mathematical phantoms is not the current state-of-the-art in nuclear medicine dosimetry. The section 11 "Dosimetry" of the SmPC has been updated to report the dose coefficients for organ absorbed doses and the effective dose derived using the anthropomorphic reference phantoms and following the recommendations of the ICRP Publication 103.

Elimination

Flortaucipir metabolism has been investigated in *in vitro* studies. Based on these data, the proposed primary metabolic pathways of flortaucipir are consistent with CYP- and non-CYP-mediated oxidation and direct glucuronidation. Identities of the metabolites have not been provided.

In vivo flortaucipir was rapidly eliminated from blood; plasma radioactivity (including parent flortaucipir and all its metabolites) fell below 10% of the theoretical maximum concentration by 5 minutes post-dose. Results suggest that flortaucipir is gradually metabolised with parent flortaucipir accounting for approximately 86% of plasma radioactivity at 5 minutes, 34% at 80 minutes, and 22% at 130 minutes post injection. In addition, 2 significant metabolite peaks were detected, accounting for approximately 30% to 35% of plasma radioactivity at or beyond 80 minutes postdose.

Dose proportionality and time dependencies

N/A

Special populations

No data on the pharmacokinetics of flortaucipir in special populations have been provided. Due to the fast uptake into the brain (peak 2 - 3 minutes), the fact that only a micro-dose of drug substance is administered to patients and that therefore there is no scientific rationale for performing additional pharmacokinetic studies in special populations this is considered acceptable. For further information reference is made to "Safety in special populations".

Pharmacokinetic interaction studies

Drug-drug interaction (DDI) of flortaucipir has been investigated in *in vitro* studies. In study XT185100 flortaucipir did not show clinically relevant drug interactions. This study used Dextromorphan O-demethylation as the marker reaction for CYP2D6 which is considered acceptable.

CYP appears to be a relatively minor contributor to the metabolism of flortaucipir; the major enzyme responsible for hepatic metabolism is AO, with some contribution by UDP-glucuronosyltransferase. Clinically relevant DDIs are not expected with AO inhibitors. Flortaucipir is not a P-gp substrate. As regards the effect of flortaucipir on the pharmacokinetics of other drugs flortaucipir is not expected to cause clinically significant inhibition of the clearance of drugs metabolised by CYP enzymes and it is unlikely that flortaucipir inhibits P-gp.

Flortaucipir is structurally similar to known MAO-A ligands and another putative tau binding tracer has been shown to bind to MAO-B. PET signal from a MAO-B-binding agent could potentially show a similar distribution to signal from a tau tracer as MAO-B is found in reactive astrocytes associated with NFTs and neuritic plaques. However, the binding and the *in vivo* data taken together suggest that neither MAO-A nor MAO-B binding contributes significantly to the flortaucipir PET signal in cortical target areas.

2.6.2.2. Pharmacodynamics

Three human studies evaluated the performance of flortaucipir as a tau PET tracer. The phase 0 study T807000, titled 'A Phase 0, Open Label, Non-Randomized, Multi-Centre, Exploratory and Safety Study of [F-18]T807', evaluated a small number of subjects with the aim of showing differential tracer retention between subjects clinically diagnosed with probable AD dementia versus CN subjects.
The phase I study A01, titled 'An Exploratory Evaluation of the Tau Protein Binding Properties, Whole-Body Biodistribution and Safety of 18F-AV-1451 Injection in Healthy Volunteers and Cognitively Impaired Subjects' was conducted to expand the preliminary results of study T807000 in a larger cohort.

The exploratory part of the phase II/III study A05, titled 'An Open-Label, Multicenter Study, Evaluating the Safety and Imaging Characteristics of 18F-AV-1451 in Cognitively Healthy Volunteers, Subjects with Mild Cognitive Impairment, and Subjects with Alzheimer's Disease', had both cross-sectional and longitudinal components. The cross-sectional component was to confirm the results of Studies T807000 and A01 in a larger sample (n=223) of subjects with clinically defined AD and MCI, as well as cognitively normal controls. The longitudinal component obtained scans at 9 and 18 months post-baseline to observe changes in cognitive and functional performance, as well as potential increases in flortaucipir PET signal over time.

Furthermore, the phase 1 study A03, titled 'Test-Retest Reproducibility of 18F-AV-1451 Injection for Brain Imaging of Tau in Healthy Volunteers and Cognitively Impaired Subjects' evaluated the test-retest performance of flortaucipir PET.

For a tabular summary of the studies evaluating PD and dose-response of flortaucipir please see Table 6 below. For details regarding study A01 please see 'Pharmacokinetics' above.

Table 6: Summary of clinical pharmacodynamic studies

Study ID	Study Drug Dose, Route, Frequency, Image Acquisition	Study Objectives	Subjects Entered and Cohort Populations	Inclusion Criteria	Clinical Pharmacology Results
T807000	370 MBq, IV, single dose Dynamic images collected from 0 to 60, and 80 to 100 minutes post-dose	1. To monitor/assess safety of IV administration of flortaucipir ^{18}F 2. To evaluate biodistribution, metabolism, and radiation dosimetry in participants with low probability of AD using PET/CT whole-body imaging 3. To evaluate uptake and signal/background information in brain PET/CT imaging of participants with a high probability of AD and age-matched low-probability controls 4. To begin collecting baseline flortaucipir ^{18}F PET/CT imaging data and gain information to improve study design for the conduct of future trials. (This study was ended before radiation dosimetry scans were performed.)	N=11 n=3 HV with low probability of AD n=8 subjects with high probability AD	HV, low probability of AD: Age ≥55 years, MMSE ≥28 High probability of AD: Age ≥55 years, MMSE <17 (original protocol), ≤25 (protocol amendment 1)	• Flortaucipir ^{18}F administration appeared to be well tolerated • No serious or treatment-emergent adverse events • No clinically significant changes in other safety parameters (vital signs, ECGs, physical examinations) • Flortaucipir ^{18}F had favourable kinetics for a tau PET imaging agent and provided preliminary evidence for potential to distinguish participants with a high probability of AD dementia vs. participants with a low probability of AD dementia
18F-AV-1451-A05E Exploratory Phase	370 MBq, IV, single dose Images collected at 80 to 100 minutes post-dose Initial PET scan at baseline First follow-up PET scan at 9 (±2) months Second follow-up PET scan at 18 (±2) months	• To compare flortaucipir ^{18}F imaging results among subjects with clinically diagnosed AD to subjects with MCI and cognitively healthy older individuals • To assess the rate of change of tau deposition as measured by flortaucipir ^{18}F uptake over time • To establish a database of cognitively healthy individuals to show the spectrum of flortaucipir imaging results in cognitively healthy individuals across a range of age strata • To explore associations between changes in flortaucipir F 18 uptake in the brain and clinical and functional measures	N=223 n=51 AD n=98 MCI n=58 OCN n=16 YCN	YCN: Age ≥20 to ≤40 years, no evidence of cognitive impairment; MMSE ≥29 OCN: Age ≥50 years, no evidence of cognitive impairment MCI: Age ≥50 years, had MCI consistent with NIA Working Group guidelines for MCI (Albert et al. 2011); MMSE ≥24 AD: Age ≥50 years, had possible or probable AD dementia consistent with NIA working group guidelines for AD; MMSE ≥20 (60 subjects) and MMSE >10 and <20 (10 subjects) (McKhann et al. 2011)	• Cognitively impaired, Aβ+ subjects had significantly higher flortaucipir SUVr than corresponding Aβ- subjects both quantitatively and visually • There was a significant relationship between baseline flortaucipir signal and baseline cognitive performance • PET signal increased longitudinally in Aβ+ but not Aβ- subjects, and the magnitude of the change was predicted most strongly by the baseline flortaucipir PET SUVr value • Apparent off-target binding has little or no impact on quantitation or visual reads
18F-AV-1451-A03	370 MBq, IV, 2 doses not more than 4 weeks apart Images collected 80 to 100 minutes and 110 to 130 minutes post-dose	• To evaluate test-retest reproducibility of flortaucipir ^{18}F for brain imaging of aggregated tau in HV and subjects with cognitive impairment	N=24 n=6 HV n=8 MCI n=10 probable AD dementia	HV: Age ≥50 years, MMSE ≥29 MCI/AD: Age ≥50 years of age who met the criteria of a diagnosis for either MCI or probable AD, MMSE ≥24	• The ICC and 95% CI for the combination region for all 24 subjects in the study was 0.971 (0.935, 0.988) for 80 to 100 minutes post-dose and 0.968 (0.926, 0.986) for 110 to 130 minutes post-dose, indicating substantial test-retest agreement.

Abbreviations: 2TC = 2-tissue compartmental model; Aβ = beta amyloid; AD = Alzheimer's disease; CI = confidence interval; CN = cognitively normal; CT = computerised tomography; ECG = electrocardiogram; HPLC = high-performance liquid chromatography; HV = healthy volunteer; ICC = intra-class correlation; ID = identifier; IV = intravenous; MCI = mild cognitive impairment; MMSE = Mini Mental State Examination; MRI = magnetic resonance imaging; NIA = National Institute on Aging; OCN = older cognitively normal controls; PERSI = parametric estimate of reference signal intensity; PET = positron emission tomography; PK = pharmacokinetics; SD = standard deviation; SUVr = standard uptake value ratio; WBD = whole-body distribution; YCN = younger cognitively normal controls.

Mechanism of action

The mechanism of action of flortaucipir has been established in non-clinical studies. According to these data, the 18F-lablelled flortaucipir binds to hyperphosphorylated PHF-tau enriched within tau neurofibrillary tangles (NFTs) thus enabling an assessment of their presence and distribution. In contrast to other approved PET imaging agents for the evaluation of AD, flortaucipir does not bind to beta amyloid (Aβ).

Primary and secondary pharmacology

As regards the primary pharmacology, the available pharmacodynamic data indicate that there is a pattern of flortaucipir retention unique to patients with AD pathology. It is typically not found in Aβ- subjects and it increases in intensity and extent with disease severity in Aβ+, but not Aβ- subjects. Within the group of Aβ+ subjects, on average flortaucipir global SUVr was higher in subjects with clinically more advanced disease.

However, potentially off-target flortaucipir-binding sites are important to note. In addition to the amyloid-associated neocortical retention attributed to binding to AD NFTs, age-related retention of flortaucipir was seen in amyloid-negative normal controls from study A05E.

As regards the test-retest performance of flortaucipir and the selection of imaging time period, the analysis of the primary efficacy variable, test-retest reproducibility using intra-class correlation (ICC), showed high test-retest agreement (ICC>0.9) for the whole cohort and across diagnostic groups for most individual brain regions and for the composite region of interest (ROI).

Analysis of variance in percent change showed low variability between scans for both the 80- to 100-minute and the 110- to 130-minute imaging windows, but with approximately 20% less variability in the standard deviation of the percent change for the 80- to 100-minute window compared to the 110- to 130-minute window. As imaging during the 80- to 100-minute window may be more convenient and results in similar or slightly greater reliability compared to the 110- to 130-minute imaging window, the 80- to 100-minute acquisition period was used in the pivotal phase 3 autopsy study A16 and is recommended in the proposed clinical labelling.

No clinical data on secondary pharmacology and on genetic differences in PD response have been provided and no relevant pharmacodynamic interaction studies have been performed in humans; this is considered acceptable.

The exposure-response relationships for flortaucipir have not been studied systematically. This is acceptable given the microdose range of the recommended mass dose of flortaucipir.

2.6.3. Discussion on clinical pharmacology

Biodistribution, plasma levels, and metabolism of flortaucipir were studied *in vitro* in human biomaterials and *in vivo* in animal models and in humans; the *in vitro* pharmacology studies are summarised in the non-clinical modules.

Overall, the analytical and statistical methods, as well as pharmacokinetic data analysis described by the applicant are considered adequate. However, with respect to the pharmacokinetic modelling, sufficient information has not been provided and the model performance remains unclear. Therefore, the results from the analyses of the pharmacokinetic modelling were considered inadequate and were not used for any conclusions to support the marketing authorisation.

The applicant has re-calculated dosimetry on request of the CHMP. The absorbed dose calculations for the reference adult male and female take into account the weighting factors (radiation and tissue) according to the recommendations of ICRP Publication 103. Due to a difference in the gastrointestinal tract structures, the time-integrated activity in the gastrointestinal structures was determined using the ICRP Publication 100 GI tract model in OLINDA. Time-integrated activity curves for all other source organs remained unchanged. Reference person equivalent doses were used to compute a reference person effective dose using equation B.3.9. The updated dosimetry calculation is acceptable. The calculated effective dose resulting from the administration of a (maximal recommended) activity of 370 MBq for an adult weighing 70 kg is about 9.6 mSv. For an administered activity of 370 MBq the typical

radiation dose to the target organ [brain] is 3.1 mGy and the typical radiation dose/doses to the critical organ/organs [right colon, small intestine, liver] are 48.2 mGy, 38.4 mGy, and 23.0 mGy respectively. Information about the effective dose, as well as the warning regarding the generic risks of radiation exposure have been included in the product information.

Furthermore, no data on the pharmacokinetics of flortaucipir in special populations have been provided. Due to the fast uptake into the brain and since only a micro-dose of drug substance is administered to patients not performing additional pharmacokinetic studies in special populations is considered acceptable.

The mechanism of action of flortaucipir has been established in non-clinical studies. According to these data, the ^{18}F-lablelled flortaucipir binds to hyperphosphorylated PHF-tau enriched within tau neurofibrillary tangles (NFTs) thus enabling an assessment of their presence. In contrast to other approved PET imaging agents for the evaluation of AD, flortaucipir does not bind to beta amyloid (Aβ). As regards the primary pharmacology, the available pharmacodynamic data indicate that there is a pattern of flortaucipir retention unique to patients with AD pathology. However, potentially off-target flortaucipir-binding sites are important to note.

No clinical data on secondary pharmacology and on genetic differences in PD response have been provided and no relevant pharmacodynamic interaction studies have been performed in humans; this is considered acceptable. Furthermore, the exposure-response relationships for flortaucipir have not been studied systematically which is acceptable given the microdose range of the recommended mass dose of flortaucipir.

2.6.4. Conclusions on clinical pharmacology

Data package on clinical pharmacology is considered acceptable.

2.6.5. Clinical efficacy

Initially targeted indication contained 2 sub-indications:

Indication I addressing general diagnostic capabilities of the Tauvid PET ("to estimate the density and distribution of the aggregated tau neurofibrillary tangles (NFTs) of Alzheimer's disease (AD) in adult patients who are being evaluated for AD." Flortaucipir (18F) was stated "to help establish a diagnosis of AD" and "to confirm the presence of AD neuropathology"), while the **Indication II** claimed its prognostic ability (…and to inform risk for progression of cognitive and functional impairment due to AD.). Flortaucipir (18F) is an adjunct to clinical and other diagnostic evaluations.

With the responses to the day 120 LoOI the applicant withdrew the second indication and adapted the wording of the first indication restricting it to the adult population with cognitive impairment.

With the day 180 LoOI further claims regarding density, distribution of NFTs throughout the brain and ability to confirm the presence of AD neuropathology were removed on request of the CHMP.

The updated claimed "Diagnostic" indication is:

> *"This medicinal product is for diagnostic use only.*
>
> *Flortaucipir (^{18}F) is a radiopharmaceutical indicated for positron emission tomography (PET) imaging of the brain to assess the neocortical distribution of aggregated tau neurofibrillary tangles (NFTs) in adult patients with cognitive impairment who are being evaluated for Alzheimer's disease (AD). Flortaucipir (^{18}F) is an adjunct to clinical and other diagnostic evaluations.*

For limitations of use, see sections 4.4 and 5.1.".

Of note, throughout this assessment "Indication I" and "Diagnostic indication" are used interchangeably and represent the current/updated claimed indication.

Indication I - Diagnostic claim

The primary support for the use of flortaucipir to characterise AD NFT pathology comes from the pivotal image-to-autopsy Study 18F-AV-1451-A16 (A16). Supportive evidence is provided from the related repeated read Study 18F-AV-1451-FR01 (FR01). The latter study was submitted as pivotal, but was not accepted as such. The two studies tested the relationship of a specific pattern of flortaucipir retention, that is, elevated retention in the posterior lateral temporal lobe, occipital lobe, or parietal lobe, with or without frontal lobe involvement (a τAD pattern [tAD+ or tAD++] as determined by visual evaluation of PET images), with NFT pathology at autopsy. Additionally, inter and intra-reader agreement and data on quantitative measurements were collected.

Studies evaluating impact of flortaucipir (18F) PET on patient management and diagnostic thinking have not been conducted, but published evidence has been provided.

Indication II - Prognosis claim

A second series of studies tested the relationship between a specific AD pattern of flortaucipir PET retention, that is, an advanced AD flortaucipir pattern (referred to during clinical development as a τAD++ pattern), and risk for near-term (18-month) cognitive deterioration. The primary analysis of the confirmatory cohort of Study 18F-AV-1451-A05 (A05C) explored whether the group of patients with an advanced flortaucipir pattern by majority read had an increased risk of a pre-defined threshold of clinical deterioration (as assessed by CDR-SB; a categorical endpoint), by comparison to the patients whose scans were not advanced. The association between advanced AD flortaucipir pattern and risk of cognitive and functional decline was also explored in a second study, Study 18F-AV-1451-PX01 (PX01). Further supportive evidence was provided from the exploratory cohort of Study A05 (Study A05E), the tau imaging addendum of Expedition 3, i.e. solanezumab study H8A-MC-LZAX (referred to as study TZAX), imaging Follow-up Study 18F-AV-1451-A18 (A18) from A05C and two supportive longitudinal trials 18F-AV-1451-A04 (A04) and 18F-AV-1451-A08 (A08). Multiple published studies including work conducted by academic investigators have also been submitted. However, as stated above, the second claim for prognosis was withdrawn.

Table 7 below contains the pivotal studies and most significant supportive studies for efficacy (for both initially proposed indications). The full list of studies contributing to safety and efficacy analyses is summarised in tabular form in documentation.

Table 7. Summary of completed studies - flortaucipir development plan (more relevant efficacy studies)

Protocol No. Study Start – Stop Date	**Study Drug Dose, Route & Frequency**	**Study Objective**	**No. of Subjects**
Pivotal trials: AD Neuropathologic Correspondence Studies for Diagnostic Performance			
^{18}F-AV-1451-A16	370 MBq, IV	Relationship between antemortem flortaucipir imaging and post-mortem tau NFT and AD neuropathologic change.	Scanned 156; Autopsy: 67 Supplemental Academic Autopsies: 16
^{18}F-AV-1451-FR01	No new doses	Relationship between antemortem flortaucipir imaging and post-mortem tau NFT and AD neuropathologic change. Assess inter-reader reliability.	Autopsy cohort: 83 cases from Study A16 Target Population Cohort:

Protocol No. Study Start – Stop Date	Study Drug Dose, Route & Frequency	Study Objective	No. of Subjects
			159 MCI and AD from Study A05C
Pivotal trials: Prognosis Studies			
^{18}F-AV-1451-A05C	370 MBq, IV	Relationship between flortaucipir PET visual interpretation and 18-month cognitive deterioration.	Enrolled 160 Scanned 159 97 MCI; 62 AD
^{18}F-AV-1451-PX01	No new doses 240 MBq, IV in Study AZES	Visual interpretation of flortaucipir PET and risk of meaningful deterioration on the CDR-SB in Study AZES.	205 141 AD, 64 MCI
Supportive trials: Prognosis Studies			
^{18}F-AV-1451-A05E	370 MBq, IV	Flortaucipir PET in AD, MCI, CN Longitudinal change in flortaucipir Preliminary relationship between flortaucipir signal and longitudinal cognitive change.	Enrolled 223 Scanned 222 51 AD, 98 MCI, 58 OCN (>50); 16 YCN (≥20 to ≤40 years of age) 1 MCI did not get scanned
^{18}F-AV-1451-TZAX	240 MBq, IV Baseline, 9 and 18 months	Visual interpretation of flortaucipir PET and risk of meaningful deterioration on the CDR-SB in Study LZAX Phase 3 solanezumab treatment study.	206 AD 204 evaluable images
Other Supportive Longitudinal Trials			
^{18}F-AV-1451-A04	370 MBq, IV	Longitudinal change in flortaucipir PET signal.	Enrolled 44 Scanned 37 (5 AD, 10 MCI, 1 ODD, 28 CN)
^{18}F-AV-1451-A08	240 MBq, IV at baseline and 12 months	Flortaucipir PET signal in subjects with MCI, AD, SMC, and CN.	Enrolled 89 Scanned 86 (5 AD, 11 MCI 45 SMC, 25 CN)
^{18}F-AV-1451-A18	370 MBq, IV	Longitudinal change of tau deposition as measured by flortaucipir F18 uptake over time: Follow-up of subjects in A05C	Enrolled 79: 29 AD; 54 MCI

Abbreviations: AD = Alzheimer's disease; AZES = I8D-MC-AZES; CDR-SB = Clinical Dementia Rating Scale – Sum of Boxes; CN = cognitively normal; IV = intravenous; MCI = mild cognitive impairment; NFT = neurofibrillary tangles; OCN = old cognitively normal; PET = positron emission tomography; YCN = young cognitively normal.

2.6.5.1. Dose response studies

Proper dose-finding has not been performed. The majority of the studies including the pivotal studies (18F-AV-1451-A16 and 18F-AV-1451-A05C) applied 370 MBq (10 mCi) flortaucipir (18F) with acquisition of 20 min duration started at 80 min post-dose. Smaller dose of 240 MBq and different times for image acquisition (a 30-minute PET acquisition after approximately 75 minutes following i.v. administration of flortaucipir 18F) were utilised in few studies, including the pivotal study 18F-AV-1451-PX01 and reached similar uptake values for flortaucipir (18F), but with higher variability as measured by the quantitative measurement methods, and higher dosing regimen of 370 MBq with image acquisition at 80-100 min post-dose has been recommended in the product information.

Supportive evidence for the use of an 80- to 100-minute post-dose PET scan acquisition period comes also from the kinetic modelling Study A10. According to the applicant, a 2-tissue compartment kinetic model provided a good fit to the flortaucipir time-activity curves and imaging data, and adequately separated subjects with AD from CN subjects. SUVr values (cerebellum reference region) tended to underestimate the kinetically modelled DVR at short uptake/washout times (40 to 60 minutes post-

dose) and overestimate DVR at long uptake/washout times (110 to 130 minutes post-dose), with the best fit at 80 to 100 minutes post-dose.

The proposed 370 MBq dose (effective dose 9.6 mSv)/20-minute acquisition exposes patients to a dose of radiation roughly similar to that for other 18F radiopharmaceuticals. The shorter 20-minute scan period reduces the likelihood of motion while on the scanner and is therefore recommended in labelling. The total mass dose of the cold product is limited to 20 µg and the volume of the injected product to 10 mL.

2.6.5.2. Main studies - Indication I (updated claimed indication)

Study 18F-AV-1451-A16 (A16)

Methods A16

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

This was a Phase III prospectively planned clinico-pathological multicentre study that evaluated the correspondence between flortaucipir PET imaging and post-mortem assessment of tau pathology.

EudraCT or clinicaltrials.gov numbers not available.

Enrolled subjects received flortaucipir (18F) PET and were followed up for autopsy during 9 months post PET-imaging. If death did not occur within this period, cognitively impaired subjects were given the opportunity to undergo a second flortaucipir F 18 PET scan and continue in the protocol at the Sponsor's discretion. No such scan was required in the cognitively normal (CN) subjects. If a subject received more than 1 scan, the scan proximal to death was used in the study analyses.

Brain samples after patients' death were collected and analysed for presence of tau/TNF and amyloid. Acquired PET images underwent blinded read by 5 readers and the outcomes were compared to the findings from the histology evaluation. The reader part of the study was blinded.

Supplemental autopsy cohort: Anticipating the possibility that enrolment of CN or non-AD dementia end-of-life patients might be challenging, the study also included data from a supplemental autopsy cohort (SAC) created in collaboration with external investigators who had independently collected flortaucipir PET scans and autopsy brain tissue.

Study Participants A16

Subjects were terminally ill patients >50 years of age with a projected life expectancy of <6 months, and cognitive status ranging from clinically normal through dementia, including subjects with both AD and non-AD dementia diagnoses. No specific inclusion criterion regarding cognitive status was defined and enrolment was not formally stratified. Subjects were classified by their neurological history by the referring physician at the time of entry into the study. No formal neurological diagnosis was done as part of the study. No biomarkers were used to assess dementia criteria.

Subjects with suspected encephalopathy due to alcoholism or end-stage liver disease, or with structural brain lesion, having clinically significant infectious disease (hepatitis, HIV, prion disease), receiving investigational, or aggressive treatment (e.g., for chemotherapy, or on respirator), or those who had ever participated in an experimental study with an amyloid or tau targeting agent (for example, immunotherapy, secretase inhibitor), or increased risk of developing long QT were excluded.

Treatments A16

Flortaucipir F 18 (18F-AV-1451), 370 MBq [10 mCi] with a maximum human dose limited to 20 µg of compound by mass was applied as an intravenous (IV) bolus prior to the first PET scan. The target dose prior to the second PET scan was the same as for the first scan. No comparator was used.

Objectives A16

The primary objective of the study was to test the relationship between ante-mortem flortaucipir F 18 positron emission tomography (PET) imaging and tau neurofibrillary pathology associated with Alzheimer's disease (AD), as measured at autopsy. It was hypothesised that a τAD pattern (τAD+, moderate AD flortaucipir pattern, or τAD++, advanced AD flortaucipir pattern) of flortaucipir retention would correspond to a B3 pattern of tau accumulation at autopsy and furthermore would occur selectively in the presence of high amyloid burden such that cases with a τAD pattern of flortaucipir retention would meet both the NFT tau and amyloid burden criteria for high ADNC at autopsy.

The secondary objective was to assess performance/reliability of independent, blinded readers.

Outcomes/endpoints A16

Only 4 endpoints (2 primary and 2 secondary) have been defined in the protocol.

Two co-primary endpoints were:

Primary Analysis 1: The diagnostic performance (sensitivity/specificity) of 5 independent readers' interpretations of ante-mortem flortaucipir PET images for detection of a pattern of flortaucipir neocortical uptake that corresponded to neurofibrillary tangles (NFT) scores of B3 (Hyman et al. 2012; Montine et al. 2012) as measured at autopsy.

Primary Analysis 2: The diagnostic performance (sensitivity/specificity) of 5 independent readers' interpretation of ante-mortem flortaucipir images for detection of a pattern of flortaucipir neocortical uptake that corresponds to high levels of AD neuropathologic change (ADNC) as defined by National Institute on Aging Alzheimer's Association (NIA-AA) criteria (Hyman et al. 2012) as measured at autopsy were evaluated (reference is made to section 2.1.2).

In order for the study to be considered a success, at least 3 of 5 readers had to achieve a lower bound of the 95% CI for both sensitivity and specificity of ≥50% (that is, statistically significantly greater than chance at a 2-sided significance level of 0.05). Setting a threshold of the lower bound at 50% was meant to rule out chance agreement (that is, guessing) as the explanation for correct responses. Requiring the same reader to have both lower bounds ≥50% for sensitivity and specificity was targeted to lower the likelihood that the result is due to chance.

The Primary Analysis 2 was only tested if the Primary Analysis 1 was met and the trial was to be considered successful if both co-primary endpoints were met. If a majority of readers determined that a PET scan was not evaluable (for example, due to motion or other issue), it was excluded from the analysis.

The two secondary endpoints were:

- Diagnostic performance of ante-mortem flortaucipir PET imaging, based on majority of interpretation of 5 independent readers (the outcome being defined by the interpretation that is chosen by the majority of the readers), for detection of a pattern of flortaucipir neocortical uptake that corresponded to NFT scores of B3 at autopsy; and for detection of a pattern of flortaucipir F 18 neocortical uptake that corresponded to high levels of ADNC as defined by NIA-AA criteria at autopsy.

- Agreement among readers in image interpretation.

Additionally, quantitative analysis of flortaucipir PET scans (standard uptake values ratio - SUVr) was performed. The goal of these analyses was to determine the SUVr optimal cut-off value to distinguish subjects with AD from subjects without AD using computer-based processing (SUVr) of PET images.

Sample size A16

Assuming sensitivity/specificity of approximately 80%, a minimum of 14 autopsy cases in each of the truth standard groups (pathological positive or negative) is required for the lower bound of the two-sided 95% CI to be > 50% for each individual reader. In order to ensure that an adequate number of subjects were recruited and came to autopsy for each truth standard group, the subjects diagnosed on clinical grounds as having dementia due to AD were considered truth standard (TS) positive and subjects diagnosed as having no cognitive impairment (cognitively normal) TS negative. Based on these assumptions, the study sample size requirement and trial stopping criteria were deemed to have been met when at least 14 subjects meeting the criteria of either TS positive and TS negative came to autopsy.

Randomisation and blinding (masking) A16

This was a single-arm study. No randomisation was done.

The study itself, was an open-label study, i.e., patients and doctors were aware of the investigational product applied. Image reads were performed in fully blinded fashion.

The sponsor's (Avid) personnel were blinded to all flortaucipir PET images after the 3 front-runner cases. The first 3 subjects (front-runners) completing the study were analysed in an open-label manner as a pilot set. These data were not included in the primary analysis cohort. A single Avid member (who was not aware of the clinical status of the subjects/images reviewed) performed periodic QA assessments by reviewing a listing of the PET image header data (not to exceed 20% unless issues identified that warrant additional review). These QA reviews served to ensure image quality control processes at the imaging core laboratory were adequately performed.

The first 30 PET scans collected were reviewed by an independent Data monitoring board (DMB) to ensure that scans in the autopsy population did not appear significantly different than scans obtained in early phase studies, where subjects were not at end-of-life. The DMB could recommend unblinding the Sponsor that would have led to exclusion of these cases from the efficacy analysis, or the study may have been terminated.

SUVr analysis was done blinded to read results and pathology data. All other Avid personnel were blinded to all flortaucipir read results, pathology data, and SUVr results until database lock.

Statistical methods A16

For the primary analysis of the first primary efficacy endpoint, the hypothesis tested was that for at least the same 3 of 5 independent readers, the lower bound of the 2-sided 95% confidence interval (CI) for both sensitivity and specificity of flortaucipir PET reading interpretations was ≥50%. For the primary analysis of the second primary efficacy endpoint, the hypothesis tested was also that for at least the same 3 of 5 independent readers, the lower bound of the 2-sided 95% CI for both sensitivity and specificity of flortaucipir PET reading interpretations was ≥50%. This trial was considered as successful if both Primary Efficacy Analyses reach statistical significance at a 2-sided significance level

of 0.05. All other efficacy endpoints were tested without multiplicity adjustments, and nominal p-values were provided.

Although nominal p-values were provided for the secondary efficacy endpoints, the hypothesis to be tested for these endpoints were:

- The majority read results had the lower bound of the 2-sided 95% CI ≥55% for both sensitivity and specificity for detection of tau neurofibrillary pathology that corresponded to NFT score of B3
- The majority read results had the lower bound of the 2-sided 95% CI ≥55% for both sensitivity and specificity for detection of a pattern of flortaucipir F 18 neocortical uptake that corresponded to high levels of AD neuropathologic change as defined by NIA-AA criteria

The observed kappa values were ≥0.64 and the lower bound of the 2-sided 95% CIs were ≥0.55 for the inter-reader agreement among readers.

No statistical testing was performed for disposition and withdrawal, demographic or other baseline characteristics, medical/surgical history, concurrent diseases/conditions, prior and concomitant medications, exposure to flortaucipir F 18, other efficacy endpoints, and all safety outcomes.

The analysis sets were defined as follows:

- Safety analysis set **(SAF)**: all patients, who received flortaucipir injection.
- Modified Intent-to-treat analysis set **(mITT)**: all SAF subjects with available PET scan visual interpretation data (regardless of PET scan image validity and evaluability).
- Full analysis sets **(FASs)**: all treated subjects who came to autopsy and had evaluable PET image and valid and interpretable autopsy specimen.
- Efficacy analysis sets 1 and 2 **(EEF1 and EEF2 for the two primary analysis respectively)**: all treated subjects who came to autopsy and had evaluable PET image and valid and interpretable autopsy specimen (for NFT and ADNC respectively) excluding the first 3 subjects who came to autopsy.
- The final full analysis set abbreviated as **SACFAS** in tables and figures (not defined in the protocol) included the patients from FAS and those from supplemental data set.

Methodologies applied for visual reads and autopsy analyses

Visual read was performed by 5 blinded readers independently. Increased neocortical flortaucipir F18 activity was to be described for 6 regions in each hemisphere of brain: anterior lateral temporal (ALT), posterior lateral temporal (PLT), occipital, parietal, precuneus and frontal. First the presence of increased uptake was evaluated. Subsequently the readers interpreted flortaucipir baseline scans as either consistent of not consistent with an AD pattern and dichotomised the reads for analysis applying the criteria in

Table 8 below:

Table 8. Visual interpretation criteria

Read Outcome		**Objective Image Features**
Not consistent with AD pattern (**τAD-**/ **Negative** flortaucipir pattern)		No increased neocortical activity or increased neocortical activity isolated to the mesial temporal, anterolateral temporal, and/or frontal regions.
Consistent with AD – **Positive** pattern (**τAD**)	**τAD+**/ Moderate AD flortaucipir pattern	In either hemisphere, increased neocortical activity in the posterolateral temporal (PLT) or occipital region(s).
	τAD++/ Advanced AD flortaucipir pattern	In either hemisphere, increased neocortical activity in the parietal/precuneus region(s), or frontal region(s) with increased uptake in the PLT, parietal, or occipital region(s).

Abbreviations: AD = Alzheimer's disease; PLT = posterolateral temporal.

All neuropathological measurements on brain tissue were evaluated in a standardised and blinded (with respect to PET image results and subject identifying/clinical information) fashion in a qualified laboratory, as per rules defined in the Technical Autopsy Manual and a Neuropathology Analysis Plan.

Neuropathological assessment was derived from the NIA-AA guidelines (Hyman et al. 2012; Montine et al. 2012). Sections from brain regions recommended for pathological assessment of AD, as well as additional, prespecified neocortical regions were sampled. Staining followed the procedures recommended in the NIA-AA practical approach guidelines (see Table 9 below). In addition to the stains listed below, regions were also stained for incidental/general pathology with H & E (haematoxylin and eosin stain), as recommended by the NIA-AA consensus recommendations.

Table 9. Regions of interest and respective staining methods in autopsy – primary analysis

		AD Neuropathologic Change		
		A	**B**	**C**
ROI number	**Region**	**IHC stain for Aβ/amyloid plaques**	**Stain for NFTs/tau***	**Silver stain for NPs**
1	Midbrain including SN and red nucleus	**Yes**	**No**	**No**
2	Cerebellar cortex/dentate nucleus	**Yes**	**No**	**No**
3	Basal ganglia at the level of the AC with basal nucleus of Meynert	**Yes**	**No**	**No**
4	Hippocampus/Entorhinal Cortex/Fusiform gyrus	**Yes**	**Yes**	**No**
5	Amygdala	**No**	**Yes**	**No**
6	Middle frontal gyrus	**Yes**	**Yes**	**Yes**
7	Superior and middle temporal gyri	**Yes**	**Yes**	**Yes**
8	Inferior parietal lobule	**Yes**	**Yes**	**Yes**
9	Occipital cortex (BA 17 & 18)	**No**	**Yes**	**No**

SN= substantia nigra; AC= anterior commissure; BA= Brodmann area ; NP= neuritic plaques

*For the front-runner cases, NFTs will be stained with both IHC (AT8) and Bielschowsky silver stain. Cases from the Efficacy Cohort will be stained for NFTs with only AT8 as IHC was determined to be the stain used for the primary pathology endpoints.

Twenty-six regions in the brain (13 per brain half – left and right side): transentorhinal cortex, entorhinal cortex, hippocampus, fusiform gyrus, amygdala, entorhinal at amygdala, transentorhinal at amygdala, middle temporal gyrus, superior temporal gyrus, middle frontal gyrus, inferior parietal lobule, peristriate occipital cortex, striate cortex, were evaluated and described for slide integrity, presence and density of NFT/NT (categories: none, rare, sparse, moderate, frequent, not assessed) and presence of atypical tau pathology (categories: present, absent, not applicable).

Consensus Panel agreed on summary scores for the following parameters and for each hemisphere: Thal Phase (0-5, not evaluable), A: Aβ plaque (0-3, not evaluable), Braak stage (0, I-VI, not evaluable), B: NFT stage based on NFTs typical for AD (0-3, not evaluable), Atypical tau (presence, absence, not evaluable), C: Neuritic plaque score (modified Consortium to Establish a Registry for Alzheimer's Disease (CERAD) score, any score, or not evaluable), NIA-AA ADNC (Not AD, low ADNC, intermediate ADNC, high ADNC, not evaluable).

The results of the pathology assessment were dichotomised for NTF and ADNC separately to define the TS for Primary Analyses 1 and 2, respectively. Positive and negative PET and TS results were then matched to define four categories of findings: "true positive (TP), "true negative" (TN), "false positive" (FP) and false negative" (FN) as shown in Table 10 below.

Table 10. Autopsy NFT score and NIA-AA autopsy diagnosis (ADNC level) truth standard (primary analysis 1 and primary analysis 2, respectively)

Physician Reader Interpretation	**Autopsy NFT Score/ADNC level - Truth Standard (TS)**			
	NFT Score B3	**High ADNC**	**NFT Score B0, B1, or B2**	**ADNC None, Low, or Intermediate**
	Truth Positive		**Truth Negative**	
Flortaucipir neocortical uptake; AD pattern **(τAD+/τAD++)**	True positive		False positive	
Flortaucipir neocortical uptake not consistent with AD **(τAD-)**	False negative		True negative	

Abbreviations: AD = Alzheimer's disease; NFT = neurofibrillary tangle; PET = positron emission tomography.
Source: Protocol Amendment 3.

Quantitative measurement of flortaucipir F 18 standard uptake value ratio (SUVr)

Descriptive statistics were presented for the quantitative analysis result of flortaucipir PET scan (SUVr) by most recent neurological disease diagnosis based on the FAS. The reference region for SUVr values was a selected white matter region derived using a parametric estimation of signal intensity method (PERSI) (Southekal et al. 2018). The global SUVr assessment was computed using a data-driven target region using multiblock barycentric discriminant analysis (MUBADA) method (Devous et al. 2018). The target region for regional SUVr were based on automated anatomical labelling atlas.

The objective of the exploratory analysis was to find a cut-off value for MUBADA SUVr for each pathological TS, to define a positive (corresponding to pathological positive) and negative (corresponding to pathological negative) scan using MUBADA SUVr values. For each pathological TS, a logistical regression model was applied using pathological TS as the dependent variable and MUBADA

SUVr value as the independent variable. Receiver operating curves (ROC) were generated, and a Youden's Index method was used to select the best cut-off values. The ROC figure and area under the curve (AUC) of the ROC was provided as well as its p-value and the SUVr cut-off value.

Results A16

Participant flow A16 (

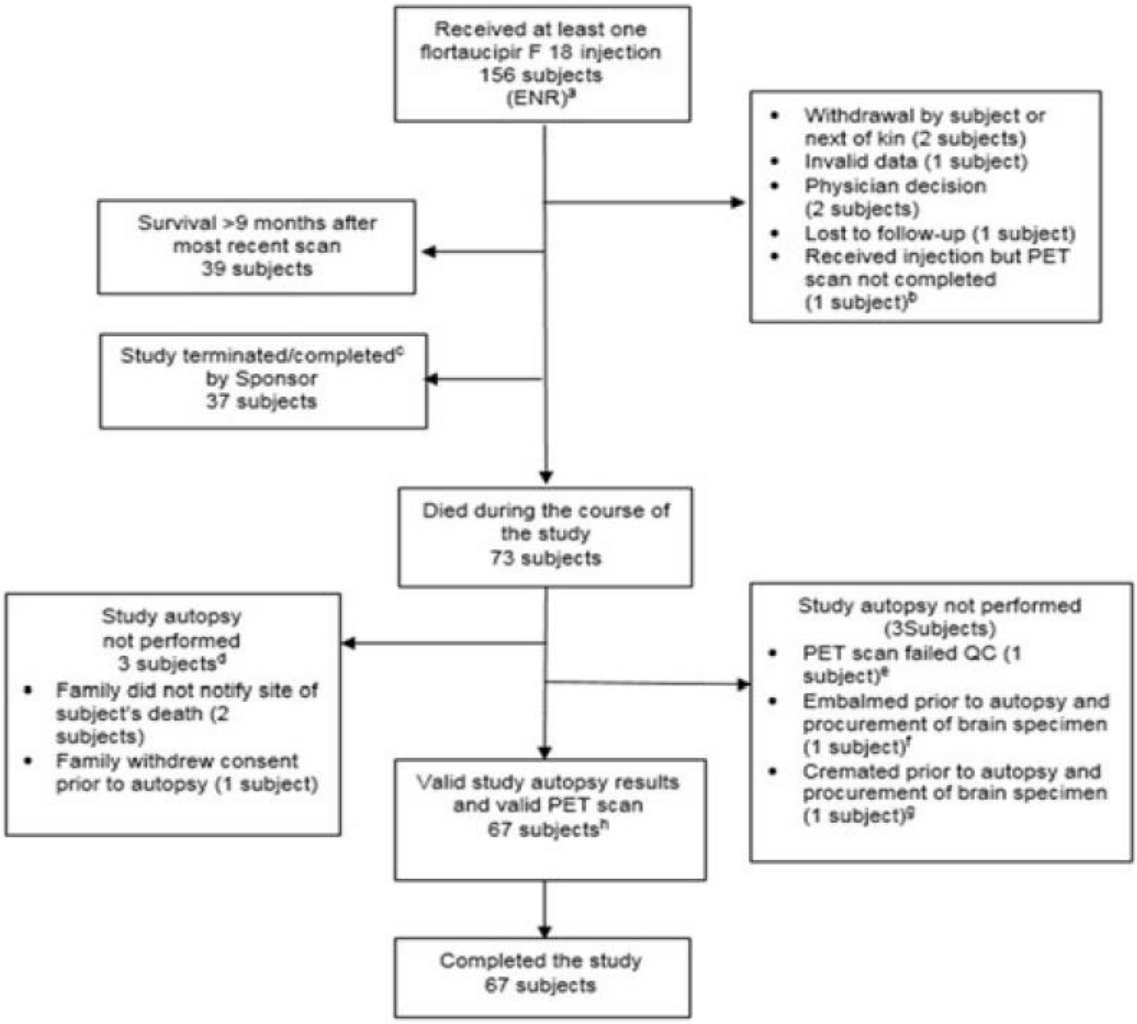


Figure 4)

A total of 156 subjects were enrolled: 103 subjects with dementia, 3 subjects with MCI, and 50 CN subjects.

Of the 70 subjects who died during the study, 3 did not come to autopsy. Thus, 67 subjects completed the study (including 3 frontrunners), defined as subjects who died, and who had valid PET scan and valid autopsy results:

- 52 subjects with dementia,
- 1 subject with MCI, and
- 14 CN subjects.

As the study neared completion, PET scan visual reads were performed on all subjects who came to autopsy and the surviving subjects who were still eligible to be followed for autopsy (n=105). This population was used for inter-reader reliability analyses.

The primary efficacy analysis population consisted of N=64 subjects (67 subjects minus the 3 frontrunners) with both a valid autopsy and flortaucipir PET scan when the study was completed:

- 49 subjects with dementia,
- 1 subject with MCI, and
- 14 CN subjects.

Thus, the final full analysis set of available cases with valid PET and autopsy (abbreviated as SACFAS in tables and figures) included:

- the primary efficacy analysis population (n=64)
- 2 evaluable frontrunners (1 frontrunner was judged unevaluable by the readers due to a motion artefact and excluded), and
- additional 16 scans from the supplemental cohort (n=82).

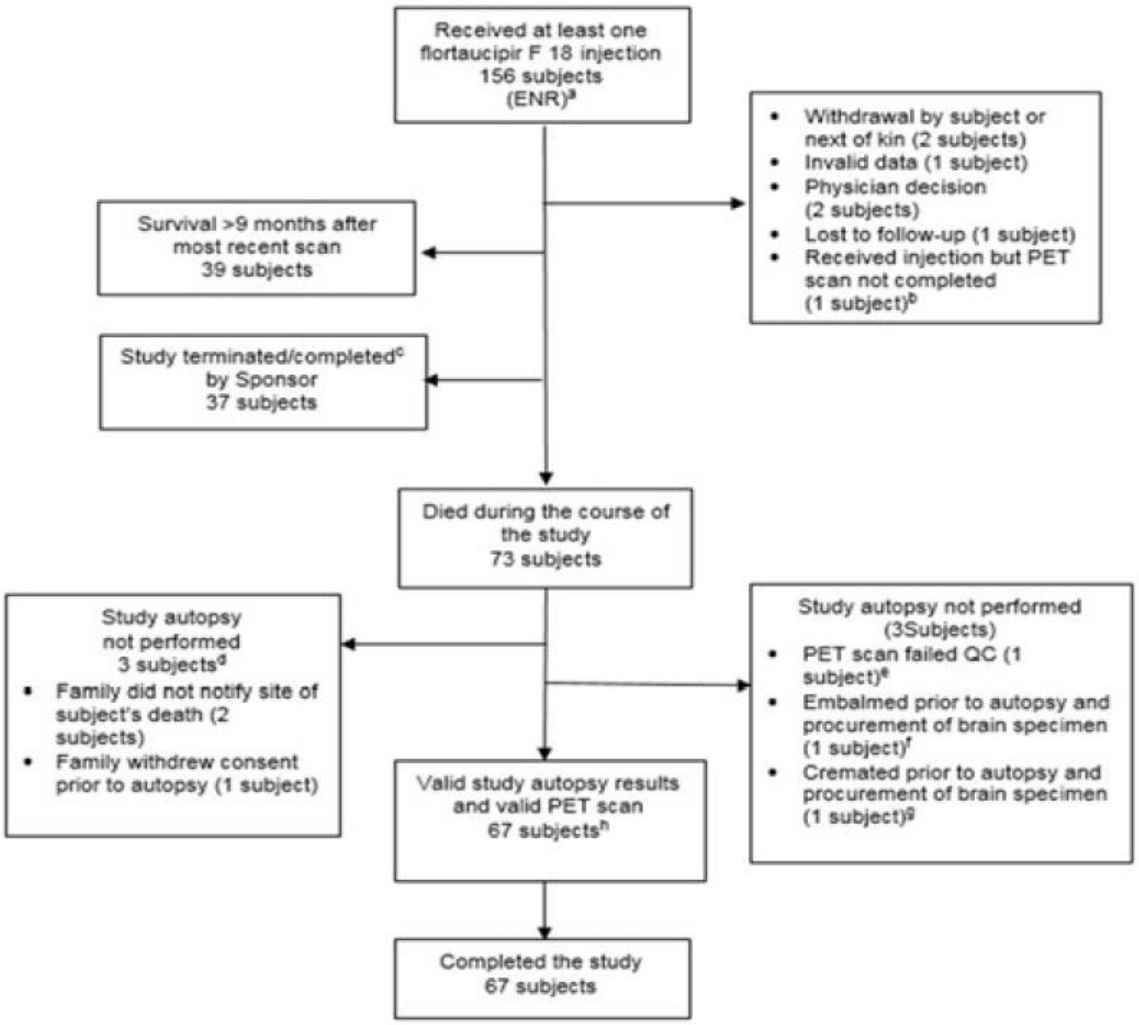


Figure 4 Participant flow

Recruitment A16

The study was conducted in 28 sites and 2 countries (USA – 27 sites and Australia – 1 site with 5 patients). The participating sites included clinical research centres, university and specialised (neurology, psychiatry, Alzheimer's disease, memory centres) hospitals and clinics, private practices. All but 1 site were located in the USA.

Date of first subject enrolled: 27 October 2015

Date of last subject completed: 13 June 2018

18F-AV-1451-A16-Supplemental Autopsy Cohort (SAC)

First case received: 21 June 2018

Last case received: 12 September 2018

Start and end dates of the autopsy and PET evaluations could not be located in the study documentation.

Conduct of the study A16

The protocol for this study was approved on 04 June 2015 and was amended three times. The Statistical Analysis Plan (SAP), which superseded the statistical plans described in the protocol, was approved on 16 August 2018. The reporting database (electronic data capture database for clinical data collected) was validated and locked for analysis on 31 July 2018. The main A16 study database including imaging reads, imaging quantitative analysis data, and pathological TS data was locked subsequently on 20 August 2018.

Key changes introduced via the three protocol amendments were:

Amendment 1 (22 October 2015): Change in the description of the primary analysis 1 and primary analysis 2, as well as addition of an enhanced secondary objective, Changes to Statistical Methods for Primary Analyses Addition of optional additional research and genetic testing, Change to sample size estimation.

Amendment 2 (18 December 2017): Increased the planned number of autopsies from 60 to 80 to contribute post-mortem histological data, Primary analysis changed from a nested sequential analysis to co-primary analyses, Neuropathology Evaluation Table 1: Change in categorisation of Truth Standard relative to Braak Stage and NFT Score, Neuropathology Evaluation Table 2: Change in categorisation of Truth Standard relative to level of AD neuropathologic change present, Removal of 54 autopsy requirement to stop enrolment, Addition of Australia as a study country, Changes to evaluation of imaging, Addition of sample size calculation and trial stopping criteria to statistical methods.

Amendment 3 (03 April 2018): Analysis Populations section added, Sample Size Calculation and Trial Stopping Criteria: added a detailed description of the consensus panel and process for determining if the stopping criteria have been met, Addition of supplemental autopsy cohort, Addition of SUVr analysis at Avid.

Originally, B2/B3 was considered truth positive for NFT endpoint and intermediate/high were considered truth positive for the ADNC endpoint in line with Hyman et al. (2012) that states, that "*For individuals with cognitive impairment at the time tissue was obtained, "Intermediate" or "High" level of AD neuropathologic change should be considered adequate explanation of cognitive impairment or*

dementia. When "Low" level of AD neuropathologic change is observed in the setting of cognitive impairment, it is likely that other diseases are present."

However, as an intermediate ADNC classification requires both a B2 level of NFTs and a significant amyloid burden, a flortaucipir scan read method associated with a B2/intermediate ADNC would have limited diagnostic utility in the absence of amyloid assessment by PET or CSF and thus, after consultation with pathology experts, the definition of positive TS was changed in mid-study from B2/B3 to B3. This decision was made before unblinding of pathology data or any flortaucipir PET scan reads occurred.

A number of amendments were also made to other relevant documents, but exact changes made to these documents and their potential impact on study outcomes was not discussed. These documents were finalised with the following version numbers and at the following dates:

Imaging review charter (Version 2.0; dated 16.05.2018), **technical operations manual** (Version 4.0; dated 29.10.2015), **blinded review manual** (Version 2.0; dated 15.05.2018), **data monitoring board** (DMB) **charter** (Version 1.0; dated 18.03.2016), **image analysis plan** (for quantitative measurements; Dated 28.06.2018), **neuropathology analysis plan** (Version 5; dated 14.05.2018), **technical autopsy manual** (Version 2.1; dated 15.09.2016), **post-mortem technical manual** (for submission of supplemental brain tissue; Version 1; dated 30.04.2018).

Conduct of the study was monitored by IQVIA (CRO) that performed routine monitoring visits and 100% source document verification at the study site.

Avid conducted audits at the study sites. Audit Certificates have been provided and do not reveal any findings of critical importance.

Two subjects (1.3%) were reported with critical protocol deviations, both related to informed consent (ICF) (signed ICF was discarded in error, and the replacement ICF was backdated; study team member (PET technologist) signed the ICF as an impartial witness).

A total of 25 subjects (16.0%) were reported with 26 major protocol deviations concerning: Study procedures criteria (6 subjects [3.8%]), Eligibility and entry criteria (5 subjects [3.2%]), Informed consent (5 subjects [3.2%]), Concomitant medication criteria (3 subjects [1.9%]), Investigational product compliance (3 subjects [1.9%]), Laboratory assessment criteria (2 subjects [1.3%]), Other criteria/GCP violations (2 subjects [1.3%]; One subject was consented with ICF v1.0, even though v2.0 was available and one subject had a cardiac arrest 1 day after the PET scan; the site was informed the same day, but reported the SAE 2 days later).

The inclusion/exclusion criteria were not met for 2 subjects (One cognitively normal subject with a life expectancy >6 months was enrolled and one subject with dementia previously participated in another clinical trial).

These protocol deviations were not likely to have affected the analyses or conclusions presented in this report.

Baseline data A16

Overview of the status of cognition, and information on baseline demography was not provided for supplemental set (n=16).

Of the 64 subjects in EFF1 (Table 11), 62 subjects (96.9%) were white, and 34 subjects (53.1%) were female. The mean age of the EFF1 was 82.5 years (range 55 to 100 years). Majority of the subjects in the CN and dementia subgroups had high school or university level of education.

Table 11 Demographics and other baseline characteristics (EFF1)

Characteristic Category	Statistic	Most Recent Neurological Disease Diagnosis[a]			Total (N = 64)
		Normal (N = 14)	Mild Cognitive Impairment (N = 1)	Dementia (N = 49)	
Age (years)	n	14	1	49	64
	Mean	78.6	76.0	83.8	82.5
	SD	12.06	-	8.63	9.59
	Median	78.5	76.0	84.0	83.5
	Min, Max	55, 97	76, 76	59, 100	55, 100
<65 years	n (%)	2 (14.3)	0	1 (2.0)	3 (4.7)
≥65 to <75 years	n (%)	4 (28.6)	0	5 (10.2)	9 (14.1)
≥75 years	n (%)	8 (57.1)	1 (100.0)	43 (87.8)	52 (81.3)
Birth gender					
Male	n (%)	8 (57.1)	1 (100.0)	21 (42.9)	30 (46.9)
Female	n (%)	6 (42.9)	0	28 (57.1)	34 (53.1)
Race					
Asian	n (%)	0	0	1 (2.0)	1 (1.6)
Black or African American	n (%)	0	0	1 (2.0)	1 (1.6)
White	n (%)	14 (100.0)	1 (100.0)	47 (95.9)	62 (96.9)
American Indian or Alaska native	n (%)	0	0	0	0
Native Hawaiian or other Pacific Islander	n (%)	0	0	0	0
Ethnicity					
Hispanic or Latino	n (%)	2 (14.3)	0	1 (2.0)	3 (4.7)
Not Hispanic or Latino	n (%)	12 (85.7)	1 (100.0)	48 (98.0)	61 (95.3)
Highest level of education completed					
Elementary school	n (%)	0	0	0	0
Middle school	n (%)	0	0	2 (4.1)	2 (3.1)
High school	n (%)	6 (42.9)	0	21 (42.9)	27 (42.2)
College/university	n (%)	6 (42.9)	1 (100.0)	13 (26.5)	20 (31.3)
Post graduate school	n (%)	2 (14.3)	0	13 (26.5)	15 (23.4)

Abbreviations: Max = maximum; Min = minimum; N and n = number of subjects; PET = positron emission tomography; SD = standard deviation.

[a] Most recent neurological diagnosis collected prior to subject's most recent flortaucipir F 18 PET scan. Subjects were classified by their neurological history by the referring physician at the time of entry into the study. No formal neurological diagnosis was done as part of the study.

Note: Percentages were based on the number of subjects in the Efficacy Analysis Set reported in each column.

There were no notable differences among the subjects based on their most recent neurological disease diagnosis.

Of 49 subjects with dementia 16 subjects (32.7%) had the diagnosis of non-AD dementia (EFF1), including those related to vascular pathology (n; %) (2; 4.1%), frontotemporal disease (FTD; 2; 4.1%), Lewy Body disease (LBD; 5; 10.2%), mixed dementia (AD plus another neurological disease; 3; 6.1%), and other (not specified; 4; 5.2%).

Findings in histology (Table 12) showed that all but 1 subject with Braak stage V/VI had amyloid burden by CERAD of level 2/3 and all but 2 subjects with Braak V/VI stage had Thal scores of 4/5 and high level ADNC. All subjects with high level ADNC had high degree of Braak (V/VI), but high scores of CERAD and Thal (2/3 and 4/5 respectively) were observed also in the subjects with lower Braak stages.

Table 12 Distribution of pathology (FAS)

Braak Stage	CERAD		THAL		ADNC	
	2/3	<2	4/5	<4	High	No/Low/Intermediate
1	1	1	1	1	0	2
2	4	2	1	5	0	6
3	2	3	0	5	0	5
4	10	2	5	7	0	12
5	14	1	13	2	13	2
6	27	0	27	0	27	0
Total	58	9	47	20	40	27

Abbreviations: ADNC = Alzheimer's disease neuropathologic change; CERAD = Consortium to Establish a Registry for Alzheimer's Disease.

Relevant family neurological disease history: In the SAF, 63 subjects (61.2%) with dementia and 16 subjects (32.0%) with normal cognition had at least 1 relevant family neurological disease history, with AD in a parent most commonly reported. Similar findings were observed in the EFF1.

Neurological disease history: In the SAF, the median time (range) since neurological disease symptoms onset was 84.0 months (11, 265) for 103 subjects with dementia and 60.0 months (60, 119) for 3 subjects with MCI; and median time (range) since neurological disease diagnosis was 60.4 months (5, 205) for 103 subjects with dementia and 60.0 months (48, 89) for 3 subjects with MCI. Similar results were observed in the EFF1.

Cognition assessments at screening: For the SAF, the mean (SD) total MMSE score for 50 subjects with normal cognition was 27.3 ± 3.44; 23.0 ± 2.65 for 3 subjects with MCI, and 11.8 ± 9.47 for 32 subjects with dementia (MMSE results were not available for the remaining 71 subjects in this group). The IQCODE score for 27 subjects with normal cognition was 3.25 ± 0.526; 3.54 ± 0.402 for 3 subjects with MCI; and 4.85 ± 0.391 for 102 subjects with dementia. Similar results were observed for the EFF1.

The most frequent (20% or higher) conditions in the medical history (SAF – all treated patients) were reported in the following 3 SOCs: infections and infestations (36 subjects [23.1%]; 26 with dementia and 10 with normal cognition), nervous system disorders (33 subjects [21.2%]; 25 with dementia, 1 with MCI, and 7 with normal cognition) and neoplasms benign, malignant and unspecified (31 subjects [19.9%]; 22 with dementia and 9 with normal cognition).

Concurrent medications: all but one subject with dementia were taking at least 1 concomitant medication. Many of the concomitant medications are those commonly prescribed for older patients, such as antithrombotic agents, beta-blocking agents, drugs for constipation, drugs for peptic ulcer and gastroesophageal reflux, lipid-modifying agents, and drugs for pain management (opioids and other analgesics and antipyretics).

Causes of death in the subjects undergoing autopsy: The most frequent cause of death in cognitively normal subjects was cardiac condition; the most frequent causes of death in the subjects with dementia were dementia-related death and various types of cancer.

Treatment compliance was not assessed.

Numbers analysed A16

The following number of subjects were included and analysed in the study (Table 13):

Table 13 Subject numbers in the analyses sets

	CN (n=50)	MCI (n=3)	Dementia (n=103)
SAF/ENR (n=156)	50	3	103
mITT (n=105)	47	1	57
FAS (n=66)	14	1	51
EFFs/SENS (n=64)	14	1	49
SACFAS (n=82)	UK	UK	UK

Abbreviations: SAF – safety analysis set; ENR – enrolled population set; mITT – modified intent-to-treat; FAS – full analysis set; EFF – efficacy analysis set; SENS – sensitivity analysis set; CN – cognitively normal; MCI – mild cognitive impairment; UK - unknown

Only 2 of the 3 patients were added in some of the analyses sets, as one had unintelligible PET image.

Outcomes and estimation A16

Primary Efficacy Analysis 1: B3 NFTs

The pre-specified study success criteria were met for both co-primary endpoints:

- Primary efficacy cohort (EFF1; n=64): Lower confidence limits of sensitivity and specificity were ≥50% for the same 3 of 5 flortaucipir PET scan readers. The sensitivity (95% CI) ranged from 92.3% (79.7, 97.3) to 100.0% (91.0, 100.0) across readers and specificity from 76.0% (56.6, 88.5) to 92.0% (75.0, 97.8) for these 3 readers.

Table 14 Primary efficacy analyses 1 (vs. B3), diagnostic performance parameters per reader, primary efficacy cohort (n=64)

Read. Num	Sensitivity (%) (95% CI)	Specificity (%) (95% CI)	PPV (%) (95% CI)	NPV (%) (95% CI)	Accuracy (%) (95% CI)	LR+ (95% CI)*	LR- (95% CI)*
1	97.4 (86.8, 99.5)	68.0 (48.4, 82.8)	82.6 (69.3, 90.9)	94.4 (74.2, 99.0)	85.9 (75.4, 92.4)	3.04 (1.85, 9.48)	0.04 (0.0, 0.19)
2	92.3 (79.7, 97.3)	92.0 (75.0, 97.8)	94.7 (82.7, 98.5)	88.5 (71.0, 96.0)	92.2 (83.0, 96.6)	11.54 (3.71, 151.59)	0.08 (0.01, 0.22)
3	92.3 (79.7, 97.3)	88.0 (70.0, 95.8)	92.3 (79.7, 97.3)	88.0 (70.0, 95.8)	90.6 (81.0, 95.6)	7.69 (3.08, 54.66)	0.09 (0.01, 0.23)
4	92.3	76.0	85.7	86.4	85.9	3.85	0.10

	(79.7, 97.3)	(56.6, 88.5)	(72.2, 93.3)	(66.7, 95.3)	(75.4, 92.4)	(2.02, 10.95)	(0.02, 0.27)
5	100.0	52.0	76.5	100.0	81.3	2.08	0.00
	(91.0, 100.0)	(33.5, 70.0)	(63.2, 86.0)	(77.2, 100.0)	(70.0, 88.9)	(1.46, 3.80)	(0.00, 0.17)

Abbreviations: CI = confidence interval; LR – likelihood ratio; NPV = negative predictive value; PPV = positive predictive value.
Note: The 95% CI was based on the Wilson score method.
*95% CIs are based on exact unconditional confidence limits. Provided as post hoc analysis.

Primary Efficacy Analysis 2: High ADNC

- Primary efficacy cohort (EFF2; n=64): The study met the success criterion; flortaucipir demonstrated sensitivity and specificity significantly greater than 50%, for detecting a high level of AD neuropathologic change, as determined by the interpretations of at least 3 of the 5 readers. The sensitivity (95% CI) ranged from 94.7% (82.7, 98.5) to 100.0% (90.8, 100.0) across readers. For specificity, the lower bound of the 95% CI was ≥50% for 3 of the 5 readers; for these 3 readers, the specificity (95% CI) ranged from 76.9% (57.9, 89.0) to 92.3% (75.9, 97.9).

Table 15 Primary efficacy analyses 2 (vs. High ADNC), diagnostic performance parameters per reader, primary efficacy cohort (n=64)

Reader Number	Sensitivity (%) (95% CI)	Specificity (%) (95% CI)	PPV (%) (95% CI)	NPV (%) (95% CI)
1	97.4	65.4	80.4	94.4
	(86.5, 99.5)	(46.2, 80.6)	(66.8, 89.3)	(74.2, 99.0)
2	94.7	92.3	94.7	92.3
	(82.7, 98.5)	(75.9, 97.9)	(82.7, 98.5)	(75.9, 97.9)
3	94.7	88.5	92.3	92.0
	(82.7, 98.5)	(71.0, 96.0)	(79.7, 97.3)	(75.0, 97.8)
4	94.7	76.9	85.7	90.9
	(82.7, 98.5)	(57.9, 89.0)	(72.2, 93.3)	(72.2, 97.5)
5	100.0	50.0	74.5	100.0
	(90.8, 100.0)	(32.1, 67.9)	(61.1, 84.5)	(77.2, 100.0)

Abbreviations: ADNC = Alzheimer's disease neuropathologic change; CI = confidence interval; NPV = negative predictive value; PPV = positive predictive value.
Note: The 95% CI was based on the Wilson score method.

Secondary and exploratory efficacy endpoints

Majority read analysis vs autopsy TS

The study met both secondary objectives (Table 16): flortaucipir majority read demonstrated statistically significant sensitivity and specificity for detecting tau neurofibrillary pathology in the majority reading results. In both analyses, the lower bounds of the 2-sided 95% CI was ≥55%.

Table 16 Secondary analyses (majority read) results

	Majority PET scan interpretation versus autopsy NFT score TS (EFF1)	Majority PET scan interpretation versus NIA-AA autopsy diagnosis TS (EFF2)
Sensitivity (95% CI)	92.3% (79.7, 97.3)	94.7% (82.7, 98.5)
Specificity (95% CI)	80.0% (60.9, 91.1)	80.8% (62.1, 91.5)
Accuracy (%) (95% CI)[a]	87.5 (77.2, 93.5)	89.1 (79.1, 94.6)
PPV (%) (95% CI)[a]	87.8 (74.5, 94.7)	87.8 (74.5, 94.7)
NPV (%) (95% CI)[a]	87.0 (67.9, 95.5)	91.3 (73.2, 97.6)
LR+	4.62	4.93
LR-	0.10	0.07

Abbreviations: CI = confidence interval; LR- = likelihood ratio negative; LR+ = likelihood ratio positive; NIA-AA = National Institute on Aging–Alzheimer's Association; NPV = negative predictive value; PET = positron emission tomography; PPV = positive predictive value; TS = truth standard.

[a] The 95% CI was based on the Wilson score method.

Similar results related to NFT scores were observed for the FAS1 and for the SENS; and likewise, for the results related to NIA-AA autopsy diagnosis.

Inter-reader agreement

Inter-reader reliability assessed agreement among the readers of the PET scans using Fleiss' kappa statistics for overall agreement and Cohens kappa for comparing pairs of readers across all PET scans (n=105) that were interpreted in the main portion of Study A16. There was a high degree of agreement, both in the overall percentage agreement (89.9%) and when comparing all reader pairs (range 81.9% to 98.1%). The Fleiss' kappa was also high: 0.80 (0.74, 0.86) for overall agreement and when comparing among all readers (range 0.65 to 0.96); p<0.001 for all comparisons.

Quantitative analysis of PET scans versus Braak stage

The goal of these analyses was to determine the SUVr optimal cut-off value to distinguish subjects with AD from subjects without AD using a nonsubjective, quantitative measurement (SUVr) of the PET images. Based on the receiver operating characteristics (ROI), an optimal cut-off point (1.126) was

identified using the Youden index.

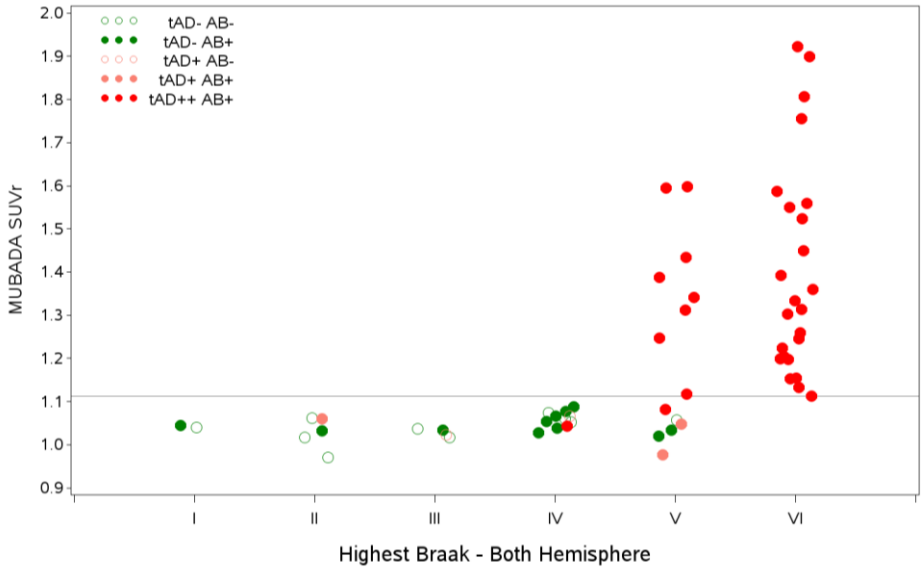


Figure 5 below presents a scatter plot of SUVr data by Braak-stage scores and visual-read categories. There was good agreement between the visual read and AD-signature VOI quantitation.

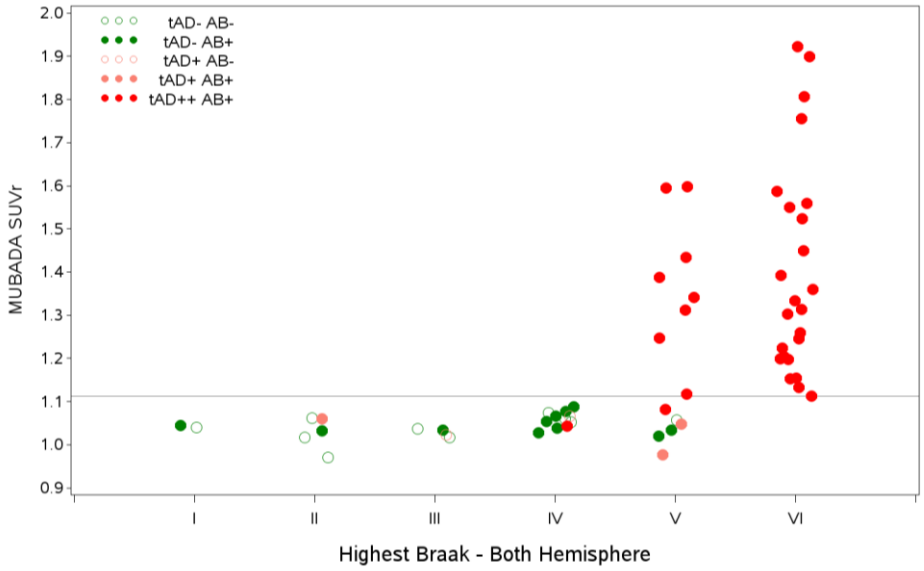


Figure 5 Scatter plot of SUVr versus NFT truth standard (Study A16)

Abbreviation: AD = Alzheimer's disease; MUBADA = multiblock barycentric discriminant analysis; NFT = neurofibrillary tangle; SUVr = standardised uptake value ratio; VOI = volume of interest.
Note: MUBADA refers to AD-signature VOI.

Ancillary analyses A16

Primary analyses by age and gender
Most subjects (52) in the EFF1 were ≥75 years of age, with 9 subjects in the ≥65 to <75 years and 3 subjects <65 years (which were too few for this statistical analysis).

For the ≥65 to <75 years group (n=9) and TS of NFT, sensitivity was 100% for all 5 readers (5 true positive readings, no false negatives), and specificity was 100% for 4 of the 5 readers (1 reader had 3 true negatives and 1 false positive). Although the sensitivity remained high (90.0% to 100%) for the ≥75 years group (n=52), the specificity decreased (that is, there were more false positives) and ranged from 47.4% to 89.5%.

Similar results were observed in the analysis of individual flortaucipir PET scan interpretation versus autopsy NIA-AA diagnosis TS by age group.

For the TS – NFT, the sensitivity was 100% for all 5 readers when interpreting scans from the 34 females in EFF1 and ranged from 83.3% to 100% for the 30 males in the analysis set. Specificity ranged from 53.8% to 92.3% for the female cohort and 50.0% to 91.7% for the male cohort.

Similar results were observed in the analysis of individual flortaucipir PET scan interpretation versus autopsy NIA-AA autopsy diagnosis TS by birth gender.

Primary analysis by Most Recent Neurological Diagnosis

Subjects were classified by their neurological history by the referring physician at the time of entry into the study. No formal neurological diagnosis was done as part of the study.

There was an insufficient number of subjects (n=1) with MCI for an analysis of that population. Among the 14 subjects with normal cognition, 1 subject had a NFT score of B3 on autopsy. This subject was scored as τAD- by 4 of the 5 readers, so their sensitivity scores were zero. The specificity scores ranged from 69.2% to 92.3% for the remaining 13 subjects with normal cognition. Among the 49 subjects with dementia, 37 (75.5% of the patients with dementia) had NFT scores of B3. The sensitivity percentages for the 5 readers were high (94.6% to 100%). The specificity percentages ranged from 33.3% to 91.7%.

Similar results were observed in the analysis of individual flortaucipir PET scan interpretation versus NIA-AA autopsy diagnosis TS by most recent neurological disease diagnosis.

Primary analysis by presence of non-AD tau

Clinical diagnosis of atypical or non-AD dementia was reported at enrolment in 16 subjects (see Table below). In histopathology 8 patients showed presence of NFT (B3) and all but one of these patients had high degree ADNC. Among the patients with positive NFT and ADNC (n=7) 6 (1 – frontotemporal dementia, 2 – Lewy body dementia, 1 –atypical AD – organic mental syndrome, 2 – vascular dementia) had positive flortaucipir uptake and the remaining one patient (mixed dementia, i.e., AD and another dementia, not specified in the report) negative PET – false negative case.

One patient with negative PET but positive Braak stage (B3) was a patient with frontal temporal dementia (FTD) clinical diagnosis and low amyloid pathology (low ADNC). In this case, microscopy demonstrated atypical tau, both glial and astrocytic in grey and white matter of frontal temporal and parietal lobes. This atypical tau could not further be differentiated from AD-type NFTs without further specialty stains that were not performed as part of the A16 study protocol. Specific non-AD pathologic diagnoses were not established.

One case (mixed dementia) showed positive flortaucipir uptake despite negative TS (W/o B3 and ADNC; false positive).

When regarding the cases per clinical diagnosis, both vascular dementia patients met histopathology criteria and had positive flortaucipir reads and neither of the two Parkinson's case had positive PET or met histopathology criteria. The remaining patients with mixed dementia (n=3), fronto-temporal dementia (n=2), atypical AD dementia (n=2), Lewy body dementia (n=5) showed mixed results with and without presence of positive PET, B3 and/or high ADNC.

The applicant calculated diagnostic performance of flortaucipir (18) PET in the subgroups of patients with dementia divided by presence or absence of atypical tau (in both hemispheres of the brain and in only one hemisphere) (Table 17). The subgroups are very small. Data in larger subgroups (atypical tau in both hemispheres – N=18 and patients without atypical tau – N=35) suggest that sensitivity was similar, while specificity was very poor in the population with atypical tau.

Table 17. Individual flortaucipir PET scan interpretation versus autopsy NFT score truth standard by atypical tau pathology – overview (efficacy analysis set #1)

Atypical Tau Status: Atypical Tau Pathology Present in Both Sides of the Brain							
	Number of Subjects (N = 18)				Sensitivity (%)	Specificity (%)	Accuracy (%)
Reader	True Positive	False Negative	False Positive	True Negative	(95% CI)	(95% CI)	(95% CI)
Reader 1	11	0	3	4	100.0 (74.1, 100.0)	57.1 (25.0, 84.2)	83.3 (60.8, 94.2)
Reader 2	10	1	2	5	90.9 (62.3, 98.4)	71.4 (35.9, 91.8)	83.3 (60.8, 94.2)
Reader 3	10	1	2	5	90.9 (62.3, 98.4)	71.4 (35.9, 91.8)	83.3 (60.8, 94.2)
Reader 4	10	1	2	5	90.9 (62.3, 98.4)	71.4 (35.9, 91.8)	83.3 (60.8, 94.2)
Reader 5	11	0	4	3	100.0 (74.1, 100.0)	42.9 (15.8, 75.0)	77.8 (54.8, 91.0)
Atypical Tau Status: Atypical Tau Pathology Present in Either Side of the Brain							
	Number of Subjects (N = 11)				Sensitivity (%)	Specificity (%)	Accuracy (%)
Reader	True Positive	False Negative	False Positive	True Negative	(95% CI)	(95% CI)	(95% CI)
Reader 1	7	0	2	2	100.0 (64.6, 100.0)	50.0 (15.0, 85.0)	81.8 (52.3, 94.9)
Reader 2	7	0	0	4	100.0 (64.6, 100.0)	100.0 (51.0, 100.0)	100.0 (74.1, 100.0)
Reader 3	7	0	0	4	100.0 (64.6, 100.0)	100.0 (51.0, 100.0)	100.0 (74.1, 100.0)
Reader 4	7	0	1	3	100.0 (64.6, 100.0)	75.0 (30.1, 95.4)	90.9 (62.3, 98.4)
Reader 5	7	0	2	2	100.0 (64.6, 100.0)	50.0 (15.0, 85.0)	81.8 (52.3, 94.9)
Atypical Tau Status: Atypical Tau Pathology Absent from Both Sides of the Brain							
	Number of Subjects (N = 35)				Sensitivity (%)	Specificity (%)	Accuracy (%)
Reader	True Positive	False Negative	False Positive	True Negative	(95% CI)	(95% CI)	(95% CI)
Reader 1	20	1	3	11	95.2 (77.3, 99.2)	78.6 (52.4, 92.4)	88.6 (74.0, 95.5)
Reader 2	19	2	0	14	90.5 (71.1, 97.3)	100.0 (78.5, 100.0)	94.3 (81.4, 98.4)
Reader 3	19	2	1	13	90.5 (71.1, 97.3)	92.9 (68.5, 98.7)	91.4 (77.6, 97.0)
Reader 4	19	2	3	11	90.5 (71.1, 97.3)	78.6 (52.4, 92.4)	85.7 (70.6, 93.7)
Reader 5	21	0	6	8	100.0 (84.5, 100.0)	57.1 (32.6, 78.6)	82.9 (67.3, 91.9)

Abbreviations: AD = Alzheimer's disease; CI = confidence interval; EFF1 = Efficacy Analysis Set #1; FN = false negative; FP = false positive; LR- = likelihood ratio negative; LR+ = likelihood ratio positive; NFT = neurofibrillary tangle; NPV = negative predictive value; PET = positron emission tomography; PPV = positive predictive value; τAD- = neocortical uptake not consistent with AD pattern; τAD+ = neocortical uptake consistent with AD pattern; τAD++ = neocortical uptake consistent with AD pattern with likelihood of disease progression; TN = true negative; TP = true positive; TS = truth standard.

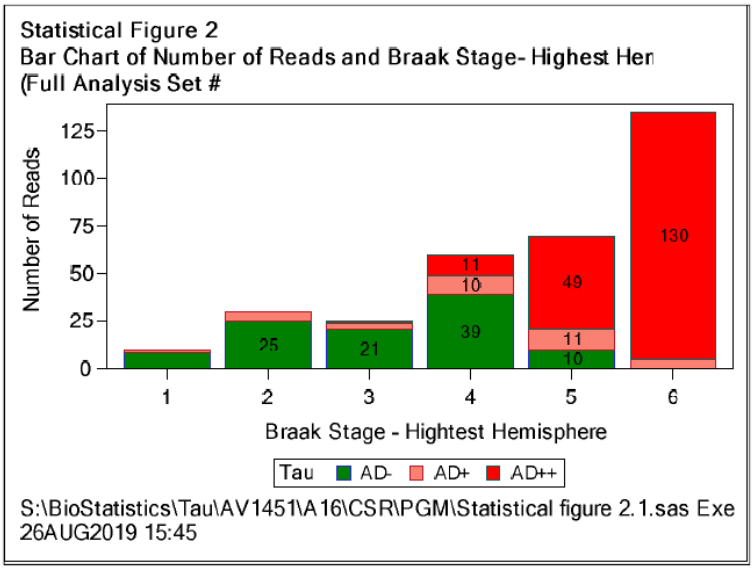
Note: The 95% CI was based on the Wilson score method.

Note: Subjects with mild cognitive impairment or dementia for whom death did not occur within 9 months of the first flortaucipir F 18 PET scan were given the opportunity to undergo a second flortaucipir F 18 PET scan. If the subject subsequently came to autopsy, the data obtained at the second scan (nearest to the time of death) has been used in the efficacy analysis. Cognitively normal subjects were not required to undergo repeats scans, even if death occurred more than nine months after a flortaucipir F 18 PET scan, as cortical tau deposition is expected to occur very slowly, if at all, in such subjects.

Note: Sensitivity = 100 × TP/(TP + FN); specificity= 100 × TN/(FP + TN); accuracy= 100 × (TP + TN)/(TP + FN + FP + TN); where FN = scan interpretation of τAD- and autopsy NFT score TS of B3; FP = scan interpretation of τAD+/τAD++ and autopsy NFT score TS of B0, B1, or B2; TN = scan interpretation of AD- and autopsy NFT score TS of B0, B1, or B2; and TP = scan interpretation of τAD+/τAD++ and autopsy NFT score TS of B3.

Primary efficacy analyses by different categories in visual read and for different thresholds for PET- (tAD++ vs. tAD+/tAD++) and TS-positivity (B2/B3 vs B3)

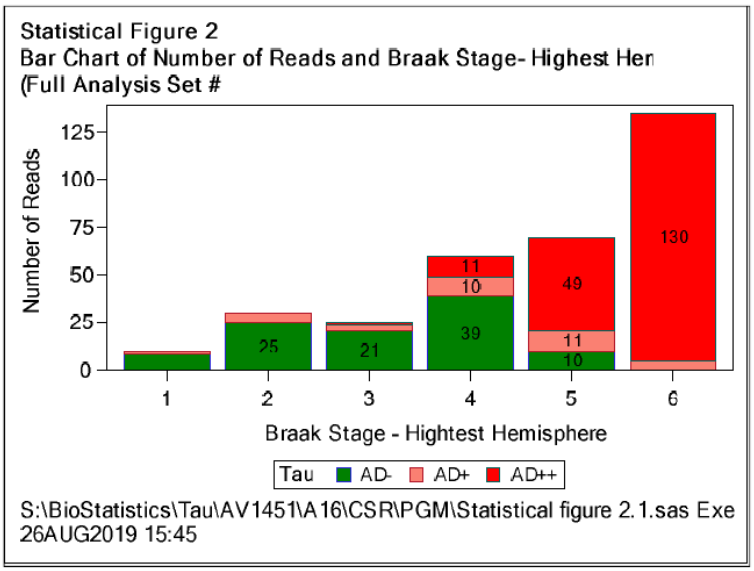
Analysis showed that majority of the visual reads with the outcome tAD++ corresponded to the Braak VI and Braak V. Overall number of tAD+ PET reads was low and these were spread across various



Abbreviations: τAD- = neocortical uptake not consistent with τAD pattern; τAD+ = neocortical uptake consistent with τAD pattern; τAD++ = neocortical uptake consistent with τAD pattern with likelihood of disease progression.

Braak categories (

Figure 6).



Abbreviations: τAD- = neocortical uptake not consistent with τAD pattern; τAD+ = neocortical uptake consistent with τAD pattern; τAD++ = neocortical uptake consistent with τAD pattern with likelihood of disease progression.

Figure 6 Comparison of the findings of visual read and Braak stages per category

To assess the correspondence of flortaucipir scan with B2/B3 (TS positive) vs B0/B1 (TS negative), an exploratory analysis was performed. This was the originally proposed approach for the primary efficacy analyses but was revised in a protocol amendment.

The results showed that the success criteria for this study would have been met for identifying B2/B3 pathology, though with lower sensitivity values than using B3 alone as the truth standard for positivity (Table 18). However, the B2/B3 analysis results were driven by B3 pathology detection. When isolating the 17 B2 cases in comparison to B0/B1, the visual reads did not identify these earlier pathological cases with acceptable sensitivity and specificity.

Table 18 Individual Flortaucipir PET Scan Interpretation versus Modified Autopsy NFT Score Truth Standard (B2 versus B0/B1) – Overview (τAD++/+ vs τAD-), Efficacy Analysis Set #1

	Number of Subjects(N=25)						
Reader	True Positive	False Negative	False Positive	True Negative	Sensitivity(%) (95% CI)	Specificity(%) (95% CI)	Accuracy(%) (95% CI)
Reader 1	7	10	1	7	41.2 (21.6, 64.0)	87.5 (52.9, 97.8)	56.0 (37.1, 73.3)
Reader 2	2	15	0	8	11.8 (3.3, 34.3)	100.0 (67.6, 100.0)	60.0 (40.7, 76.6)
Reader 3	2	15	1	7	11.8 (3.3, 34.3)	87.5 (52.9, 97.8)	64.0 (44.5, 79.8)
Reader 4	5	12	1	7	29.4 (13.3, 53.1)	87.5 (52.9, 97.8)	52.0 (33.5, 70.0)
Reader 5	9	8	3	5	52.9 (31.0, 73.8)	62.5 (30.6, 86.3)	56.0 (37.1, 73.3)
Majority Read	4	13	1	7	23.5 (9.6, 47.3)	87.5 (52.9, 97.8)	56.0 (37.1, 73.3)

For the purpose of assessing if a more stringent visual interpretation criteria had impact on reader sensitivity and specificity, further analyses were performed to compare pathological scores with a visual read score, when positive PET was defined as τAD++. For this τAD++ cases (n=34) were considered scan positive, compared to all other non-τAD++ (n=30) PET scan reads. For this analysis, the majority read sensitivity and specificity for reads versus NFT scores were 84.6% and 96.0%, respectively. For NIA-AA pathologic diagnosis the majority read sensitivity was 86.8% with specificity of 96.2% to detect high ADNC. Lower bounds of CIs were >50% for all readers for both TS (B3 and high ADNC). More stringent visual read criterion (τAD++ for definition of positive PET) produces an increase in specificity with a modest decrease in sensitivity compared to use of the AD+ visual criteria.

State of cognition vs. CERAD, Braak and PET

The study covered a full range of tau pathology from Braak NFT Stages I to VI. Forty one of 66 cases (64 plus 2 frontrunners with interpretable visual reads) were classified as Braak Stages V to VI (B3) NFT pathology; 40 of 41 (97.6%) of these cases had CERAD moderate or frequent amyloid neuritic plaques. Only one of these cases was asymptomatic (cognitively normal) prior to death. In the subjects with Braak Stage III/IV (B2) NFT pathology, 12 of 17 (69.4%) were amyloid positive (CERAD) and 10 of 17 (58.8%) were clinically symptomatic.

SACFAS – the analysis set including the supplemental autopsy cases (n=82)

Combining the supplemental cases (most of whom were low tau burden, non-AD dementia cases) with the primary efficacy cohort and frontrunners resulted in an increase in observed specificity with minimal reduction in sensitivity. All 5 readers met the criterion for lower confidence limits of sensitivity and specificity ≥50% for both the NFT and AD diagnosis (high ADNC) truth standards.

- Analysis vs. B3 NFT: The sensitivity (95% CI) ranged from 89.1% (77.0, 95.3) to 93.5% (82.5, 97.8) and for specificity from 66.7% (50.3, 79.8) to 94.4% (81.9, 98.5).
- Analysis vs. NIA-AA high ADNC: The sensitivity (95% CI) ranged from 95.1% (83.9, 98.7) to 97.6% (87.4, 99.6); the specificity (95% CI) ranged from 65.9% (50.5, 78.4) to 90.2% (77.5, 96.1).

Summary of discrepant read cases

Of the 66 cases with visual reads (including 2 frontrunners) sixteen cases had at least one discrepant read when compared to pathology primary endpoint results.

Further assessment of these 16 cases by Avid expert readers identified 4 key error categories, descriptive of the types of discrepancies that occurred between the visual reads and pathology results:

- **Atypical PET retention patterns**: These scans demonstrated relatively small, isolated foci of uptake in the parietal or occipital lobes with little to no temporal lobe uptake, leading to both positive and negative read errors.
- **Mild but visible tracer uptake associated with borderline pathology cases**: Cases in this category exhibited mild tracer uptake, often in the posterior lateral temporal lobe, with Braak IV or lower NFT stage and moderate to frequent amyloid pathology.
- **Accurate interpretation with unmatched Braak stage**: These scans were interpreted accurately based on the read methodology and training but did not correspond to the expected Braak stage.
- **Over interpretation by reader**: Image noise, particularly in temporal cortex, was mistaken for neocortical tau uptake in certain scans (8 cases) leading to false positive reads.

Study 18F-AV-1451-FR01 (FR01)

Methods FR01

This was a retrospective reader study. No new data were collected and new evaluation of PET images collected in the studies A16 and A05C was done to re-assess the diagnostic performance of flortaucipir (^{18}F) PET against histopathological findings in autopsy (gained in the study A16 and from the supplemental population) and to analyse inter and intra-reader variability utilising the images collected in the A16 and A05C studies (the latter including patients with MCI).

Study Participants FR01

No patients were enrolled. Data from the A16 and A05C studies were used.

Treatments FR01

None.

Objectives FR01

Key objectives of the study were re-assessment of diagnostic performance of flortaucipir (18F) PET against neuropathology and evaluation of inter and intra-reader agreement in the population similar to the target indication.

Outcomes/endpoints FR01

Primary Objective measures

Primary Objective 1, Analysis 1 - Accuracy of Flortaucipir PET Scan Interpreted as AD Pattern in Detecting NFT Tau Stage

- The diagnostic performance (sensitivity/specificity) of 5 independent readers' interpretations of ante-mortem flortaucipir PET images (τAD+/τAD++ or τAD-) for detection of a pattern of flortaucipir neocortical uptake that corresponds to NFT Score (truth standard) of B3 at autopsy.

Primary Objective 1, Analysis 2 - Accuracy of Flortaucipir PET Scan Interpreted as AD Pattern in Detecting AD Neuropathological Change

- Diagnostic performance (sensitivity/specificity) of 5 independent readers' interpretations of ante-mortem flortaucipir PET images (τAD+/τAD++ or τAD-) for detection of a pattern of flortaucipir neocortical uptake that corresponds to high levels of ADNC at autopsy. For individuals with cognitive impairment, high levels of AD neuropathologic change were considered adequate to explain cognitive impairment or dementia symptoms.

Primary Objective 2 Analysis - Assessment of Inter-reader Reliability

- The inter-reader reliability was assessed based on the 5 readers' visual interpretations of 242 scans comprised of 83 scans from Study A16 autopsy cases, and 159 cases from Study A05.

Four secondary objectives/analyses further evaluated the precision of flortaucipir PET scan results. The secondary efficacy analyses are described in the protocol appendix.

Secondary objective measures

Secondary Analysis 1

- The relationship between ante-mortem flortaucipir PET imaging of an AD pattern with uptake beyond the temporal/occipital regions (τAD++) and NFT and ADNC, as measured at autopsy.

Secondary Analysis 2

- Inter-reader reliability of flortaucipir PET scans with an AD pattern that is beyond the temporal/occipital regions (interpreted as τAD++) pattern across the 5 readers.

Secondary Analysis 3

- Inter-reader agreement of flortaucipir PET scan interpreted as AD pattern for the intended clinical use population (patients from the Study A05).

Secondary Analysis 4

- Intra-reader agreement of flortaucipir PET scan visual interpretation interpreted as AD pattern τAD (τAD+/ τAD++) vs. non AD pattern (τAD-)

Similar image reading methodology as in the A16 study was applied with 5 qualified and fully blinded readers analysing the flortaucipir (18F) images. In contrast to the A16 study, specific training was conducted in the readers prior to the reading procedure.

Sample size FR01

Assuming 80% sensitivity and 80% specificity, 14 TS positive or TS negative cases were needed to show the lower bounds of 95% confidence interval (CI; 2-sided) greater than 50%, for either sensitivity or specificity, with a Wilson score method to calculate the 95% CI. Out of the 83 Study A16 cases with autopsy results, 47 were neurofibrillary tangle (NFT) stage B3 (TS positive for primary analysis 1) and 41 were high ADNC (TS positive for primary analysis 2). This sample size was considered adequate to provide sufficient power to assess the accuracy of flortaucipir F 18 PET scan in detecting underlying pathological changes.

Fleiss' Kappa will be used to assess the reader-to-reader consistency in flortaucipir F 18 scan visual interpretation. Five independent qualified physicians will read 242 unique scans (83 autopsy scans from Study A16 and 159 scans from Study A05). Assuming the Fleiss' Kappa is expected to be 0.7, this sample size provides over 90% power in detecting a kappa value greater than 0.6, under a two-sided type I error rate of 0.05.

Randomisation and blinding (masking) FR01

No randomisation was done as no patients were recruited.

Reading procedure was fully blinded to the clinical and pathology information, as well as to the previous study results and to the planned analyses in the FR01 study.

Statistical methods FR01

Statistical methods applied for the primary efficacy analyses in this study were similar to those in the A16. Majority read method with the same finding in the same readers for sensitivity and specificity was applied with at least 50% threshold for lower bound of the 95%CI. Three hypotheses were tested for the two co-primary objectives, as detailed in Section 10.2 of the SAP in the Documentation of

Statistical Methods appendix. The trial was considered successful if all three primary efficacy analyses achieved their success criteria.

The precision of flortaucipir PET scan interpretation was assessed in 2 ways: inter- and intra-reader reliability.

Fleiss' Kappa was used to assess the reader-to-reader consistency in flortaucipir F 18 scan visual interpretation. Assuming the Fleiss' Kappa is expected to be 0.7, the sample size of 242 PET images provides over 90% power in detecting a kappa value greater than 0.6, under a two-sided type I error rate of 0.05.

Inter-Reader Agreement of Flortaucipir F 18 PET Scan Interpreted as AD Pattern:

The overall reader to reader reliability across 5 readers for interpreting the flortaucipir PET scan was assessed using Fleiss' Kappa. The primary hypothesis to be tested was that the lower bound of two-sided 95% confidence interval of Fleiss' Kappa will be greater than or equal to 0.6. The degree of agreement between two readers was assessed in a pair-wise manner using Cohen's Kappa statistics. The percent of agreement between the two readers was calculated for each reader pair. The overall percent of agreement across 5 readers was also calculated.

Inter-Reader Reliability of Flortaucipir F 18 PET Scan Interpreted as τAD++ Pattern:

This analysis assessed the reliability of flortaucipir F 18 scan interpreted as τAD++ pattern across 5 readers. Same analyses as described above was applied.

Inter-Reader Agreement of Flortaucipir F 18 PET Scan Interpreted as AD Pattern for Intended Clinical Use Population:

The analysis was done in the subjects with MCI enrolled in study A05. The same type of analysis as above was conducted.

Intra-Reader Agreement of Flortaucipir F 18 PET Scan Visual Interpretation Interpreted as AD Pattern:

Intra-reader reliability will be assessed using randomly selected 20 cases (EFFP3). These randomly selected cases were read twice by every reader. A Cohen's Kappa statistics was used to assess the agreement of the two reading results (scans that are interpreted as an AD pattern, i.e., τAD (τAD+/ τAD++) vs τAD-) by every reader.

Results FR01

Participant flow FR01

The study did not recruit any patients; data from the studies A16 and A5C were utilised. In total, data from 242 cases were included containing patients with normal (CN) and impaired cognition (mostly dementia) from the study A16 and the A16 supplement (N=83), as well as all MCI and AD cases from the A05C pivotal trial (N=159).

Recruitment FR01

The study did not recruit any patients.

Study start date: 27 March 2019

Data cutoff date: 03 May 2019

Conduct of the study FR01

The protocol was amended once on 21 March 2019. The following key changes were made in protocol amendment 1 appendix.

- Study design (Number of scans was updated from 163 to 242, "in a random sequence" was deleted)
- The number of scans in the analysis populations and analyses was updated

No protocol deviations were reported in this study.

Baseline data FR01

The mean age of the analysis population was 75.9 years (range of 50 to 100 years). The mean age of autopsy cases was 81.6 years, compared with 72.9 years for non-autopsy cases.

The percentage of male and female subjects was 52.5% and 47.5%, respectively. Overall, 97.1% of subjects were Caucasian, followed by Black or African American (1.7%) and Asian (1.2%).

The population was generally well-educated. 26.4% of subjects completed at least a high school education and 69.8% of the subjects completed college/university education.

Table 19 Demographics and baseline characteristics

Characteristic		**Autopsy Cases (Study A16) (N=83)**	**Non-autopsy Cases (Study A05) (N=159)**	**Total (N=242)**
Age (years)	N	83	159	242
	Mean	81.6	72.9	75.9
	Min, Max	55, 100	50, 97	50, 100
Gender				
Male	n (%)	42 (50.6%)	85 (53.5%)	127 (52.5%)
Female	n (%)	41 (49.4%)	74 (46.5%)	115 (47.5%)
Race				
Asian	n (%)	1 (1.2%)	2 (1.3%)	3 (1.2%)
Black or African American	n (%)	1 (1.2%)	3 (1.9%)	4 (1.7%)
Caucasian	n (%)	81 (97.6%)	154 (96.9%)	235 (97.1%)
Ethnicity				
Hispanic or Latino	n (%)	4 (4.8%)	9 (5.7%)	13 (5.4%)
Not Hispanic or Latino	n (%)	79 (95.2%)	150 (94.3%)	229 (94.6%)
Highest level of education completed				
Elementary school	n (%)	0	0	0

Middle school	n (%)	3 (3.6%)	2 (1.3%)	5 (2.1%)
High school	n (%)	32 (38.6%)	32 (20.1%)	64 (26.4%)
College/university	n (%)	48 (57.8%)	121 (76.1%)	169 (69.8%)
Other	n (%)	0	4 (2.5%)	4 (1.7%)

Abbreviations: n = number of subjects; N = number of subjects in the analysis set.

The information on what type of diagnoses was made for patients could not be located in the study report.

Numbers analysed FR01

One subject's scan was excluded from the primary efficacy analyses based on 4 of 5 readers determining it to be unevaluable for the same reason (noise or low count density).

Table 20 Summary of Quality of Scans for Visual Interpretation

	Autopsy Cases (Study A16) (N=83)	**Non-autopsy Cases (Study A05) (N=159)**	**Total (N=242)**
Scans Evaluable	82 (98.8%)	159 (100%)	241 (99.6%)
Scans Unevaluable	1 (1.2%)	0	1 (0.4%)
Reasons for Scans Unevaluable	1 (1.2%)	0	1 (0.4%)
Extreme noise or low count density that impairs evaluation	1 (1.2%)	0	1 (0.4%)

Abbreviation: N = number of subjects in the analysis set.

The following analysis sets were applied:

- Efficacy Analysis Population for Accuracy Evaluation (EFFA) – 82 patients with autopsy
- Efficacy Analysis Population for Precision Evaluation (EFFP):
    - EFFP1 - (n=241) included reading results for all valid and interpretable scans from all cases from studies A16 and A05. The analysis set was used to conduct primary objective 2 efficacy analysis, and secondary objective 2 efficacy analysis
    - EFFP2 – (n=159 from the A05 study) The analysis set was used to conduct analysis for secondary objective 3
    - EFFP3 – (n=20 cases randomly selected to evaluate intra-reader reliability). The analysis set was used to conduct efficacy analysis for secondary objective 4.

Outcomes and estimation FR01

The study met the first primary objective: flortaucipir demonstrated statistically significant sensitivity and specificity for detecting tau neurofibrillary pathology, as determined by the interpretations of at least 3 of the 5 readers (Table below). The sensitivity (95% CI) ranged from 87.2% (74.8, 94.0) to

93.6% (82.8, 97.8). For specificity, the lower bound of the 95% CI was ≥50% for 4 of the 5 readers; for these 4 readers, the specificity (95% CI) ranged from 77.1% (60.9, 87.9) to 91.4% (77.6, 97.0).

The study also met the second primary objective: flortaucipir demonstrated statistically significant sensitivity and specificity for detecting high levels of AD neuropathic changes, as determined by the interpretations of at least 3 of the 5 readers (Table 21 below). The sensitivity (95% CI) ranged from 95.1% (83.8, 98.6) to 97.6% (87.4, 99.5). For specificity, the lower bound of the 95% CI was ≥50% for 4 of the 5 readers; for these 4 readers, the specificity (95% CI) ranged from 70.7% (55.5, 82.3) to 85.4% (71.5, 93.1).

The lower confidence limits of the majority read score also exceeded 50% for both endpoints.

Table 21 Individual flortaucipir PET scan interpretation (τAD+/τAD++ or τAD-) versus autopsy NFT score (upper half of the table) and ADNC (lower half of the table) truth standard (N = 82)

Reader Number	**Sensitivity (%) (95% CI)**	**Specificity (%) (95% CI)**	**PPV (%) (95% CI)**	**NPV (%) (95% CI)**
1	91.5 (80.07, 96.64)	77.1 (60.98, 87.93)	84.3 (71.99, 91.83)	87.1 (71.15, 94.87)
2	89.4 (77.41, 95.37)	91.4 (77.62, 97.04)	93.3 (82.14, 97.71)	86.5 (72.02, 94.09)
3	87.2 (74.83, 94.02)	85.7 (70.62, 93.74)	89.1 (76.96, 95.27)	83.3 (68.11, 92.13)
4	93.6 (82.84, 97.81)	62.9 (46.34, 76.83)	77.2 (64.79, 86.16)	88.0 (70.04, 95.83)
5	89.4 (77.41, 95.37)	77.1 (60.98, 87.93)	84.0 (71.49, 91.66)	84.4 (68.25, 93.14)
Majority	89.4 (77.41, 95.37)	80.0 (64.11, 89.96)	85.7 (73.33, 92.90)	84.8 (69.08, 93.35)
Reader Number	**Sensitivity (%) (95% CI)**	**Specificity (%) (95% CI)**	**PPV (%) (95% CI)**	**NPV (%) (95% CI)**
1	95.1 (83.86, 98.65)	70.7 (55.52, 82.39)	76.5 (63.24, 86.00)	93.5 (79.28, 98.21)
2	95.1 (83.86, 98.65)	85.4 (71.56, 93.12)	86.7 (73.82, 93.74)	94.6 (82.30, 98.50)
3	95.1 (83.86, 98.65)	82.9 (68.74, 91.47)	84.8 (71.78, 92.43)	94.4 (81.86, 98.46)
4	97.6 (87.40, 99.57)	58.5 (43.37, 72.24)	70.2 (57.34, 80.47)	96.0 (80.46, 99.29)
5	95.1 (83.86, 98.65)	73.2 (58.07, 84.31)	78.0 (64.76, 87.25)	93.8 (79.85, 98.27)
Majority	95.1 (83.86, 98.65)	75.6 (60.66, 86.17)	79.6 (66.36, 88.52)	93.9 (80.39, 98.32)

Abbreviations: ADNC = AD neuropathological change; CI = confidence interval; N = number of subjects in the analysis set; NFT = neurofibrillary tangle; NIA-AA = National Institute on Aging–Alzheimer's Association; NPV = negative predictive value; PET = positron emission tomography; PPV = positive predictive value.

LRs in the analysis against NFT ranged between 2.5 and 10.4 for LR+ and were 0.1 for all readers and for LR-. Accuracy (95% CI) ranged from 80.5 (70.63, 87.62) to 90.2 (81.91, 94.97).

For the analysis against ADNC LR+ 2.4 to 3.9 and LR- 0.1. Accuracy (95% CI) ranged from 78.0 (67.95, 85.64) to 90.2 (81.91, 94.97).

Sensitivity analyses which tested tAD defined in "whole brain" showed similar diagnostic performance as the primary analyses.

Diagnostic performance based on the more stringent definition of PET image "positivity (Secondary objective analyses 1 and 2)

Use of more stringent definition of the positive tau PET, i.e., tAD++ (instead of tAD+/tAD++) against the TS of NFT and ADNC (highest scores – B3 and "high" for ADNC, respectively), led to worsening of sensitivity and improvement of specificity compared to the definition of tAD+/tAD++ for positive PET image.

*Table 22 Secondary endpoint: individual flortaucipir PET scan interpretation **(τAD++ vs τAD-/τAD+)** versus autopsy NFT score (upper half of the table) and ADNC (lower half of the table) truth standard (N = 82)*

Reader Number	**Sensitivity (%) (95% CI)**	**Specificity (%) (95% CI)**	**PPV (%) (95% CI)**	**NPV (%) (95% CI)**
1	87.2 (74.83, 94.02)	88.6 (74.05, 95.46)	91.1 (79.27, 96.49)	83.8 (68.86, 92.35)
2	80.9 (67.46, 89.58)	100.0 (90.11, 100.00)	100.0 (90.82, 100.00)	79.5 (65.50, 88.85)
3	76.6 (62.78, 86.40)	94.3 (81.39, 98.42)	94.7 (82.71, 98.54)	75.0 (60.56, 85.43)
4	89.4 (77.41, 95.37)	82.9 (67.32, 91.90)	87.5 (75.30, 94.14)	85.3 (69.87, 93.55)
5	83.0 (69.86, 91.11)	94.3 (81.39, 98.42)	95.1 (83.86, 98.65)	80.5 (65.99, 89.77)
Majority	83.0 (69.86, 91.11)	94.3 (81.39, 98.42)	95.1 (83.86, 98.65)	80.5 (65.99, 89.77)
Reader Number	**Sensitivity (%) (95% CI)**	**Specificity (%) (95% CI)**	**PPV (%) (95% CI)**	**NPV (%) (95% CI)**
1	92.7 (80.57, 97.48)	82.9 (68.74, 91.47)	84.4 (71.22, 92.25)	91.9 (78.70, 97.20)
2	85.4 (71.56, 93.12)	92.7 (80.57, 97.48)	92.1 (79.20, 97.28)	86.4 (73.29, 93.60)
3	85.4 (71.56, 93.12)	92.7 (80.57, 97.48)	92.1 (79.20, 97.28)	86.4 (73.29, 93.60)
4	92.7 (80.57, 97.48)	75.6 (60.66, 86.17)	79.2 (65.74, 88.27)	91.2 (77.04, 96.95)
5	87.8 (74.46, 94.68)	87.8 (74.46, 94.68)	87.8 (74.46, 94.68)	87.8 (74.46, 94.68)
Majority	87.8 (74.46, 94.68)	87.8 (74.46, 94.68)	87.8 (74.46, 94.68)	87.8 (74.46, 94.68)

Abbreviations: ADNC = AD neuropathological change; CI = confidence interval; N = number of subjects in the analysis set; NFT = neurofibrillary tangle; NIA-AA = National Institute on Aging–Alzheimer's Association; NPV = negative predictive value; PET = positron emission tomography; PPV = positive predictive value.

Inter-reader and intra-reader agreement

For inter-reader analysis in total population (primary objective 2) there was a high degree of agreement, both in the overall percentage agreement (93.8%) and when comparing 2 readers (range 91.3 to 97.1%). The Fleiss' kappa was also high: 0.87 for overall agreement (95% CI: 0.83, 0.91); p<0.0001 for all comparisons. The lower bound of two-sided 95% CI was > 0.6 for all comparisons.

Inter-Reader Reliability of Flortaucipir PET Scan Interpreted as τAD++ Pattern agreement (secondary objective 2) remained high with overall percentage agreement of 92% and Fleiss' kappa of 0.84 for overall agreement (95% CI: 0.80, 0.88) p<0.0001 for all comparisons.

Inter-reader agreement evaluated in the target population of the patients with MCI and dementia (secondary objective 3; n=159 from the study A05C) also showed high degree of agreement, both in the overall percentage agreement (95.1%) and when comparing reader pairs (range 93.1% to 97.5%). The Fleiss' kappa was also high: 0.90 for overall agreement; p<0.0001 for all comparisons.

Intra-reader agreement of Flortaucipir PET Scan Visual Interpretation Interpreted as AD Pattern (Secondary objective 4; n=20) showed high degree of intra-reader agreement of the 2 reading results by every reader (range 80% to 95%); p<0.0077 for all comparisons. However, Kappa was very variable (range: 0.6 to 0.89) and lower bound of the 95% CI was very low (0.24 to 0.52) in 3 of 5 readers.

Table 23 Intra-reader agreement – individual data per reader (n=20)

Reader	**% Intra-Reader Agreements[a]**	**Kappa[b]**	**95% CI[b]**	**P-value[b]**
1	95.0	0.89	(0.69, 1.00)	<.0001
2	85.0	0.71	(0.41, 1.00)	.0010
3	80.0	0.60	(0.24, 0.95)	.0077
4	95.0	0.89	(0.67, 1.00)	<.0001
5	90.0	0.79	(0.52, 1.00)	0.0004

Abbreviations: CI = confidence interval; N = number of subjects in the analysis set; PET = positron emission tomography.

[a] Percentage of agreement within a reader has been computed as follows: number of images for which reader had the same interpretation at initial and second read divided by the total number of images evaluated twice by a reader, multiplied by 100%.

[b] Cohen's kappa Statistics.

Several further sensitivity analyses with inter and intra-reader agreement were done with intra-reader analysis based on stringent definition of tAD-positive PET (tAD++) being most noteworthy. This analysis showed improvement in intra-reader agreement compared to the original definition of positive PET.

Diagnostic performance was re-calculated utilizing less stringent definition for positive NFT pathology (i.e., B2+3 vs B0+1) in the autopsy population and showed worsening in sensitivity with improved specificity.

2.6.5.3. Main studies – Indication II

Flortaucipir (18F) can be used to help establish a diagnosis of AD and to inform risk for progression of cognitive and functional impairment due to AD. This indication is no longer claimed.

Study 18F-AV-1451-A05 (Study A05)

Methods A05

For consistency and convenient reading both phases of study A05, Exploratory (A05E) and Confirmatory (A05C), will be presented in this section. Only the confirmatory phase of this study A05C is considered to be pivotal.

Title: An Open-Label, Multicentre Study, Evaluating the Safety and Imaging Characteristics of 18F-AV-1451 in Cognitively Healthy Volunteers, Subjects with Mild Cognitive Impairment, and Subjects with Alzheimer's Disease.

EudraCT or clinicaltrials.gov numbers are not available.

Study 18F-AV-1451-A05 is a Phase 2/3 cross-sectional and longitudinal open-label multicentre observational study evaluating the safety and imaging characteristics of flortaucipir PET in cognitively healthy volunteers, subjects with mild cognitive impairment and subjects with AD. The study was conducted in 2 phases, an Exploratory/hypothesis generating phase 2 (A05E), and a Confirmatory/validation phase 3 (A5C), which had separate subjects and analyses.

The study synopsis states that follow-up PET scans were not conducted in the confirmatory cohort subjects, but the study report mentions that follow-up PET scans for subjects in the confirmatory cohort were conducted under a separate protocol, 18F-AV-1451-A18 (A18). This was not mentioned in the Summary of Clinical Efficacy. The follow-up study A18 is shortly described in the section for the supportive studies.

Study Participants A05

For both the Exploratory and Confirmatory Phases (

Table 24), screening assessments included demographic information, cognitive testing, safety assessment, and magnetic resonance imaging (MRI), including both volumetric and standard clinical sequences. Raters administering the cognitive testing were blinded to the flortaucipir scans for subjects in the Confirmatory Cohort. Subjects who qualified for the study had both a florbetapir F 18 PET imaging session and a flortaucipir PET imaging session at baseline but not at screening. The option to participate in cerebrospinal fluid (CSF) collection by lumbar puncture (LP) was also offered to subjects >50 years of age at some centres in the Exploratory Phase.

In both the Exploratory and Confirmatory Phases, subjects >50 years of age who completed the baseline flortaucipir PET scans were asked to return for follow-up visits at 9 (±2) months and 18 (±2) months following the initial flortaucipir scan; longitudinal follow-up visits were not conducted for the young cognitively healthy control group, as these subjects were not expected to show any change in flortaucipir binding or cognitive performance over an 18-month time span. Cognitive assessments and updates to concomitant medications and medical history were collected at each follow-up visit. Follow-up flortaucipir PET scans were also performed at 9 and 18 months in the Exploratory cohort, but not the Confirmatory cohort subjects. Subjects or their designated decision maker were contacted by phone at 5 and 14 months following the initial flortaucipir scan to collect updated concomitant medications and medical history.

Table 24 Number of subjects included in the exploratory and confirmatory phase of study A05

Number of subjects	**Exploratory Cohort**	**Confirmatory Cohort**
Planned	230	150
Enrolled	223	160
Treated (at least 1 dose):	222	160
Completed	167	111

Table 25 Study populations of studies A05C and PX01

Study Design Element	**Study A05C**	**Study PX01**
Age	>50 years of age.	55 to 85 years old
Cognitive status	**MCI or dementia with suspected neurodegenerative cause.**	MCI due to AD or probable AD by NIA-AA criteria
Cognitive testing	MMSE between 20 and 27, inclusive.	MMSE 20 to 30 inclusive CDR Global of 0.5 or 1 with a memory box score of >0.5 RBANS delayed memory score<85
Amyloid status	Unknown at study entry; baseline amyloid PET scans obtained after enrolment.	Amyloid positive by florbetapir PET or CSF measurement
Key exclusions	Clinically significant psychiatric disease, structural abnormalities on MRI (stroke, brain mass) likely to interfere with PET, history of TBI, clinically significant cardiovascular disease or risk factors for QT prolongation, alcohol or substance abuse/dependence, and other clinically significant medical conditions (for example, infectious disease, cancer).	Significant neurological disease, history of clinically evident stroke, clinically important carotid or vertebrobasilar stenosis or plaque, history of multiple concussions, current mental disorder or psychiatric diagnosis, history of alcohol or substance abuse

Abbreviations: AD = Alzheimer's disease; CDR = Clinical Dementia Rating scale; CSF = cerebrospinal fluid; MCI = mild cognitive impairment; MMSE = Mini-Mental State Examination; MRI = magnetic resonance imaging; NIA-AA = National Institute on Aging and the Alzheimer's Association; PET = positron emission tomography; RBANS = Repeatable Battery for the Assessment of Neuropsychological Status; TBI =traumatic brain injury

As the applicant points out, the exploratory cohort of study A05 included healthy volunteers (subdivided in young and older cognitively normal) and patients with ranging cognitive status. These were clinically normal with MMSE >29 (n=58), clinically diagnosed MCI (NIA-AA criteria; Albert et al. 2011) with MMSE >24 (n=98) and clinically diagnosed possible or probable AD (NIA-AA criteria; McKhann et al. 2011) divided in two groups:

- mild AD group, with MMSE>20 (n=39)
- moderate AD group with MMSE 10-19 inclusive (n=12).

The diagnosis MCI or AD in the exploratory cohort (A05E) was made according to the US National Institute on Aging and the Alzheimer's Association (NIA-AA) criteria (Albert et al. 2011 and McKhann et al. 2011) which were the diagnostic research criteria at the time the study was initiated in 2014. This is in principle acceptable. However, it should be noted that diagnostic criteria have been refined since then moving the research framework from a clinical-biological diagnosis to a purely biological definition of AD, centred around a biomarker definition relying on amyloid β, tau and neurodegeneration (ATN) status (Jack et al. 2018). According to these criteria the diagnosis of AD requires both, amyloid β marker (CSF or PET) and tau marker (CSF or PET).

It should further be noted that this purely biological definition has been critically discussed recently by Dubois et al 2021 in a paper by the European Internal Working Group (IWG) on Alzheimer's disease. This resulted in a recommendation that the clinical diagnosis of AD still requires both, the presence of a specific clinical phenotype and biomarker evidence of AD pathology (amyloid-positive and tau-positive).

In addition, it should be noted that the confirmatory phase (A05C) obviously included an independent subject population restricted to subjects clinically diagnosed as MCI or dementia with a suspected neurodegenerative cause with an MMSE score between 20 and 27, inclusive. It is unclear whether this independent population is comparable to the exploratory population with respect to the diagnostic criteria, since neurodegenerative cause may not be restricted to AD. Further taking into consideration the differences in the inclusion criteria between the exploratory and the confirmatory phase of the trial, the applicant was requested to clarify whether the participants in the Exploratory Cohort of study A05 could be considered a different population from the one in the Confirmatory cohort. The applicant clarified in the responses that study A05E was only hypothesis generating and this hypothesis needed to be confirmed with adequate enough evidence. Despite using enriched criteria mentioned above, study A05C did not yield robust results, since the primary analysis did not meet the pre-specified success criteria (lower 95% confidence limits >1.0).

Treatments A05

All subjects received single IV bolus administrations of florbetapir with target dose of 370 MBq (10 mCi) and flortaucipir F 18 injection with a target dose of 370 MBq (10 mCi) at baseline. Flortaucipir PET scans were used for the consistency with an AD pattern (τAD+, τAD++) or the absence an AD pattern (τAD-). The florbetapir PET scans were used to identify patients with the presence of beta amyloid as (Aβ+) positive or the absence as (Aβ-) negative. No comparator was administered in this study.

Objectives A05

Exploratory Phase: The primary objective of the cross-sectional component was to compare flortaucipir imaging results among subjects with Alzheimer's disease (AD), mild cognitive impairment (MCI) and cognitively healthy older individuals. The primary objective of the longitudinal component was to assess the rate of change of tau deposition as measured by flortaucipir F 18 uptake over time.

The secondary objective of the cross-sectional component was to establish a database of cognitively healthy individuals to show the spectrum of flortaucipir imaging results in cognitively healthy individuals across a range of age strata.

Confirmatory Cohort: The primary objective of the confirmatory cohort, which was comprised of subjects with AD and MCI, was to provide independent validation of the relationships observed in the exploratory analyses of the first phase. In particular, the goal of the second phase is to confirm the relationship between flortaucipir uptake in the brain as measured by positron emission tomography (PET) and the subsequent rate of cognitive decline observed over longitudinal follow up. As defined in the statistical analysis plan, this required assessing whether or not a baseline flortaucipir PET scan is visually interpreted as τAD++, predicts a higher risk of subjects' clinically meaningful cognitive and functional deterioration within 18 months of scan, as measured by the Clinical Dementia Rating Scale Sum of Boxes (CDR-SB) change from baseline.

The secondary objective of the Confirmatory (second) Phase longitudinal component, as defined in the statistical analysis plan was to assess the diagnostic performance of baseline tau positivity according to

a flortaucipir scan visual interpretation, for predicting subjects' clinically meaningful cognitive and functional deterioration within 18 months of scan, as measured by the CDR-SB scales.

In both the Exploratory and Confirmatory Phases, scans were evaluated qualitatively (visual interpretation) and quantitatively (standardised uptake value ratio [SUVr]).

According to the applicant, study A05C also included an assessment of sensitivity/specificity for flortaucipir PET positivity to pre-specific thresholds for cognitive and functional decline. The hypothesis tested was that, of the 5 independent imaging physicians, at least 3 will have the lower bounds of 2-sided 95% CIs ≥50%, for both sensitivity and specificity. This analysis was included in the studies to satisfy a request from FDA for sensitivity and specificity to predict progression. Since this endpoint is not met, it was not further discussed by the applicant. This is not supported since this is important information that a claim regarding the diagnostic performance of flortaucipir for prediction of progression for individual patients cannot be made. See below.

For the secondary objective of the study A05 confirmatory phase, please see the section for Study A05C - Ancillary analyses.

Outcomes/endpoints A05C

In this section the outcomes of the confirmatory cohort A05C will be described. Five imaging physicians independently visually interpreted the flortaucipir PET scans as 1 of 3 read patterns as follows (further described in the Summary of Clinical Efficacy):

- not consistent with an AD pattern (τAD-)
- consistent with a Moderate AD pattern (τAD+), or
- consistent with an Advanced AD pattern (τAD++).

Read Outcome		**Objective Image Features**
Not consistent with AD pattern (τAD-)		No increased neocortical activity, or increased neocortical activity isolated to the mesial temporal, anterolateral temporal, and/or frontal regions.
AD pattern (τAD)	τAD+	In either hemisphere, increased neocortical activity in the posterolateral temporal (PLT) or occipital region(s).
	τAD++	In either hemisphere, increased neocortical activity in the parietal/precuneus region(s), or frontal region(s) with increased uptake in the PLT, parietal, or occipital region(s).

For analysis, visual reads were dichotomised as follows for the prognosis performance studies:

- a τAD++ pattern (Advanced tau AD pattern predicted to be associated with near-term/18-month notable progression)
- a non-τAD++ pattern, that is either
    - a moderate tau AD pattern, predicted to have lower likelihood of progression [τAD+], or
    - a non-AD pattern [τAD-].

Readers were blinded to all clinical information. All readers were trained/retrained in person on the flortaucipir PET scan read methodology using the proposed clinically applicable visual read method that was used across all 4 pivotal efficacy studies.

The majority interpretations of 5 independent readers was derived using individual readers' interpretation results to classify scans for analysis. For binary evaluations of majority read, interpretations were classified as either τAD++ or non-τAD++ (τAD+ or τAD-) according to the median choice of 3 of the 5 readers.

A single external expert nuclear medicine physician interpreted the florbetapir PET scans as either positive (Aβ+) or negative (Aβ-) in accordance with the current approved interpretation methods. Both florbetapir and flortaucipir scans were also evaluated quantitatively according to published methods (Joshi et al. 2015, Devous et al. 2018, Southekal et al. 2018, Pontecorvo et al. 2017, 2019).

Primary Cognitive/Functional Outcome Assessment – CDR-SB

The applicant selected CDR-SB as the primary cognitive and functional efficacy assessment for the AO5C study, as well as PX01. The CDR scale (Berg 1988) examines 6 categories of cognitive functioning domains. Each domain is scored on a scale ranging from 0 to 3 (including 0.5), with 0.5 indicating mild impairment and higher scores indicating greater levels of impairment. A global rating of dementia can be generated from the domain scores. A CDR-SB can be generated from the total across domains. The CDR-SB sum scores range from 0 to 18, with higher scores indicating greater cognitive and functional impairment.

CDRs and other cognitive/functional assessments were obtained at baseline, 6 months, 12 months, and 18 months in both studies. Subjects in Study PX01 also had a 24-month assessment, but due to the parent treatment study discontinuation for futility, relatively few subjects had 24-month assessments. Therefore, the 18-month assessment was chosen for the primary endpoint analysis in Study PX01 which also matched Study A05C.

The CDR-SB worsening 1 point or more was nominally chosen as the threshold for the flortaucipir prognosis study design to determine when patients were considered to have deteriorated cognitively and functionally. Mean change from baseline was also evaluated.

Image Interpretation and Analysis: Five nuclear medicine or radiology physicians independently interpreted the flortaucipir PET scans as either not consistent with an AD pattern (τAD-); or Consistent with an AD pattern (τAD+, τAD++) according to the criteria described above.

Secondary/Exploratory Cognitive and Functional Outcome Assessments

Nominal thresholds for clinically meaningful cognitive and functional decline used in the prognosis studies are outlined as follows. These assessments were considered exploratory in Study A05C and secondary analyses in Study PX01.

- Clinical Dementia Rating (CDR) Scale Global (any stage increase)
- Mini-Mental State Examination (MMSE) (decrease of 3 points or more)
- Pfeffer Functional Activities Questionnaire (FAQ) (increase of 3 points or more)
- Alzheimer's Disease Assessment Scale – Cognitive Subscale (ADAS-Cog) (increase by 4 points or more)

The applicant included other key efficacy variables, flortaucipir SUVr, flortaucipir visual interpretation, florbetapir SUVr, florbetapir visual interpretation.

The primary evaluation was binary and interpretations were classified as either τAD++ or non-τAD++ (which comprised of τAD+ and τAD-) according to 3 of 5 raters. It was a "yes" or "no" evaluation. More "granulation" included for τAD+ and for τAD- may have been more helpful and may have provided more details for the risk of progression in a staggered manner. However, the design of the confirmatory cohort was based on the evaluation of the primary hypothesis that the hazard of progressing to the clinically meaningful event as determined by CDR-SB value change (1 point or more increase) within 18 months would be significantly greater for subjects with flortaucipir scans rated (by the majority of the five readers) with a τAD++ pattern, as compared to those with scans rated as non-τAD++ (τAD- and τAD+ but not τAD++).

Sample size A05C

In Study A05 exploratory data, 75% (15/20) of τAD++ subjects had an event while only 27% (7/26) of non-τAD++ subjects experienced an event. Assuming a similar τAD++ rate and that a similar event rate would be observed in this Confirmatory Cohort, a sample of 120 subjects in total should have provided approximately 90% power to detect a risk ratio of at least 2, under a 2-sided type I error rate of 5%.

Preliminary analysis of the Exploratory Phase data showed that clinical assessments may have a relatively large variance. It was important to ensure that enough subjects will be in each group for sensitivity and specificity calculation, such as the subjects with or without a significant disease progression. To that end, the applicant assumed at least 20% of the subjects were to be in each group at the 18-month follow-up. With this assumption, a 5% 2-sided type I error rate, and that 95% CI will be calculated using Wilson score method; a sample of 115 subjects will provide >90% probability to observe a lower bound of 95% CI above 50%, for both sensitivity and specificity. Also, assuming an approximate 20% drop-out rate over the 18-month follow-up, this study aimed to enrol 150 subjects in total for the Confirmatory (second) Phase.

A subgroup of subjects from Study A05 Confirmatory Cohort was selected for their flortaucipir scans to be read twice in a random sequence to explore intra-reader reliability of scan interpretation. In Study A05 Exploratory Cohort data, the τAD++ rate was 47% among AD/MCI subjects with baseline MMSE 20 to 27. Assuming that this would be the rate of τAD++ in the Confirmatory Cohort, a simulation showed that a sample size with 20 random selected cases would have an approximately 95% probability to have the τAD++ rate in the range from 30% to 70% for the selected cases. This relatively balanced τAD++ rate should ensure a meaningful kappa statistic calculation for assessing intra-reader reliability.

The primary objective of the Confirmatory (second) Phase was to assess the hazard ratio of subjects with flortaucipir scan rated as τAD++ versus subjects with scans rated as non-τAD++ in progressing to a clinically meaningful event (CDR-SB with 1 point or more increase) within 18 months of scan. In Study A05 exploratory data, 75% (15/20) of τAD++ subjects had an event while only 27% (7/26) of non-τAD++ subjects experienced an event. Assuming a similar τAD++ rate and that a similar event rate would be observed in this Confirmatory Cohort, a sample of 120 subjects in total should have provided approximately 90% power to detect a risk ratio of at least 2, under a 2-sided type I error rate of 5%.

The sample size estimation is considered acceptable.

Randomisation and blinding (masking) A05C

Randomisation

There was no randomisation. Study A05 was a Phase 2/3 cross-sectional and longitudinal observational study that evaluated imaging characteristics of flortaucipir in control subjects and patients with clinically defined MCI and AD dementia (AD).

Blinding (masking)

The investigational products were used open label.

To avoid bias, raters administering the cognitive testing were blinded to the flortaucipir F 18 scans for subjects in the Confirmatory Cohort.

Avid personnel were blinded to the Confirmatory Phase longitudinal cognitive data. Avid personnel also remained blinded to the PET scans (both florbetapir F 18 and flortaucipir F 18) performed on subjects in the Confirmatory Phase, with the exception of periodic QA assessments of a random subset of image data (not to exceed 20% unless issues were identified that warranted additional review). These QA reviews served to ensure image quality and data integrity for the scans collected and managed by the imaging core laboratory. Avid staff performing the PET scan QA assessment were not aware of diagnostic group or cognitive test scores for the subjects/images being reviewed.

The independent visual readers (Confirmatory Cohort) were blind to patient clinical status for the cases being reviewed.

Statistical methods A05C

General Considerations

Frequency distributions, including counts and percentages were included for all categorical outcomes. Summary statistics, including mean, standard deviation (SD), median, minimum and maximum values, are presented for all continuous outcomes. Unless otherwise specified, hypothesis testing was 2-sided with a type I error rate of 0.05. All statistical analyses were performed using SAS® version 9.3 or higher.

Mixed model with repeated measures (MMRM) analyses also modelled mean change in each cognitive/functional variable as related to majority flortaucipir visual interpretation. Finally the inter-rater and intra-rater reliability of the PET interpretation was assessed with kappa statistics.

The majority interpretation of 5 independent readers was derived using individual readers' interpretation results to classify scans for analysis. For binary evaluations of majority read, interpretations were classified as either τAD++ or non-τAD++ (τAD+ or τAD-) according to the median choice of 3 of the 5 readers.

The primary analysis of Study A05C evaluated the HR of progressing to a CDR-SB value change (1 point or more increase) within 18 months, for subjects with flortaucipir scans rated as τAD++, as compared to those with scans rated as non-τAD++ (τAD- and τAD+). Time to first occurrence of the event was modelled using a Cox proportional hazard model by baseline tau status as determined by majority flortaucipir scan visual reading results from the 5 independent imaging physicians, and the HR of τAD++ rated subjects progressing to the event over non-τAD++ rated subjects along with the 95% CI were provided. The Cox proportional hazard model was adjusted for baseline age, ANART, and baseline CDR-SB score.

The confirmatory phase was informed by the exploratory phase of study A05 which identified a pattern of flortaucipir retention that according to the applicant is unique to amyloid-positive subjects.

Statistical analysis of the confirmatory part appears acceptable in general. However, the primary objective to investigate hazard ratios instead of risk ratios or likelihood ratios might be questioned in principle. It is, however, understood that the time-to-event analysis evaluating a hazard ratio has been selected to include censored subjects that discontinue prematurely.

Risk ratios would prevent from interpretational issues in case of non-proportional hazard. They may also be derived from the time-to-event analysis (milestone survival probabilities). Nevertheless, none of the subjects were censored at 18 months facilitating the evaluation without time-to-event model.

However, risk ratios depend on the prevalence, i.e. included population. Likelihood ratios are considered to be somewhat more informative, since they would be independent of the prevalence (hence severity of the disease) under the assumption that sensitivity and specificity can be transferred. E.g. likelihood ratios for a positive outcome indicate the factor by which the odds of worsening is increased after obtaining a positive test result. Obviously, a likelihood ratio of 1 would indicate that no information at all is given by the test. Nevertheless, the applicant did not plan for a confirmatory analysis based on likelihood ratios.

Results A05E and A05C

A05E

The overarching goal of the Exploratory Phase was to further investigate the pattern of flortaucipir PET imaging across the range of disease, in cognitively healthy subjects through patients with cognitive decline. Additionally, the Exploratory Phase investigated relationships between flortaucipir PET and cognitive decline over the 18-month study period and served to generate hypotheses tested in the Confirmatory Phase.

For the Exploratory phase, of the 223 enrolled subjects, 1 did not receive flortaucipir, leaving 222 in the primary analysis (flortaucipir as a function of diagnosis).

According to the applicant, the results of the Exploratory Phase cross-sectional component suggest that there may be a pattern of flortaucipir retention that is unique to patients with AD pathology, which is not found in Aβ- subjects, and increases in intensity and extent with disease severity in Aβ+ but not Aβ- subjects. On quantitative analyses cognitively impaired (clinically defined AD or MCI) Aβ+ subjects had significantly higher flortaucipir SUVr than the corresponding Aβ- subjects in all neocortical areas, including the MUBADA composite, as well as in amygdala, parahippocampus, and anterior but not posterior hippocampus, caudate, or putamen.

Consistent with autopsy-based histopathological findings suggesting that accumulation of abnormal tau in AD begins in the temporal lobe, the largest and most consistently observed differences between Aβ+ and Aβ- subjects were in temporal lobe regions (temporal lobe and fusiform cortex). Within the Aβ+ subjects, flortaucipir global (MUBADA) SUVr was higher on average in subjects with clinically more advanced disease.

The primary objective of the Exploratory Phase longitudinal component as specified in the protocol was to assess the rate of change of tau deposition as measured by flortaucipir uptake over time. An additional objective was to explore associations between baseline flortaucipir retention with clinical and functional measures, in order to generate hypotheses regarding the predictive relationships between flortaucipir and cognitive decline that could be tested in the Confirmatory Phase of the study. According to the applicant the results of the Exploratory Phase longitudinal component demonstrate that

flortaucipir PET signal increased from baseline to 18 months in Aβ+ but not Aβ- subjects, and that the magnitude of the change was predicted most strongly by the baseline flortaucipir PET SUVr value.

Additionally, there was a consistent significant relationship between flortaucipir PET signal and cognitive outcome across multiple test measures and different methods of statistical analysis regardless of whether flortaucipir PET signal was determined quantitatively or by visual interpretation. Thus, a stepwise regression analysis identified baseline MUBADA SUVr as a significant predictor of 18-month change in MMSE, ADAS, FAQ, and ADAS+FAQ, even though the analysis population was limited to amyloid-positive subjects.

Table 26: MMRM analysis of cognition change over 18 months in AD/MCI subjects with a τAD++ PET pattern versus subjects lacking a τAD++ pattern – exploratory phase

	LS Mean Change (± SEM) at 18 Months		**p-value**
Test	**τAD++ N = 31**	**Non-τAD++ N = 68**	**τAD++ vs Non-τAD++**
MMSE	-5.0 ± 0.63	-1.4 ± 0.43	<.0001
ADAS	8.3 ± 1.03	2.1 ± 0.69	<.0001
FAQ	6.1 ± 0.84	1.5 ± 0.60	<.0001
ADAS + FAQ	14.2 ± 1.53	4.0 ± 1.06	<.0001

Abbreviations: SEM = standard error of mean.

The results of the Exploratory Phase cross-sectional component suggest that there may be a pattern of flortaucipir retention that is unique to patients with AD pathology, which is not observed in Aβ- subjects. The analysis was used to compare the mean SUVr values between diagnostic groups (AD, MCI, and OCN) within amyloid beta status (Aβ+, Aβ-), which was performed on the MUBADA SUVr as well as the SUVr for each brain region. It should be noted that the YCN group served as a negative control to facilitate the understanding of a true negative flortaucipir scan and was not included in the analysis due to the obvious age difference between the groups. The older Aβ- cognitively normal (OCN) individuals who, as expected, did not accumulate tau pathology over 18 months were an adequately robust control group for the Study A05E longitudinal analysis objectives.

It also appears that the group classified as τAD++ had statistically worse performance on the ADAS, MMSE or FAQ. A comparison with the AD- in the OCN group could have been performed and discussed.

Flortaucipir PET signal increased from baseline to 18 months in Aβ+ but not Aβ- subjects. Although the longitudinal part of this study according to the applicant seems to suggest that the group of patients with an τAD++ pattern seems to have an increased cognitive decline over 18 months, it is noted that this analysis was only undertaken in Aβ+ subjects.

Overall the exploratory phase was hypothesis generating and the hypotheses that will be tested in the confirmatory phase of the study is whether flortaucipir uptake in the brain as measured by PET will be able to predict cognitive decline.

The two Lilly experts who also interpreted flortaucipir PET images for the integrated efficacy analysis were the same as the two Avid imaging experts (AKA and MDD).

A05C

In contrast to the Exploratory Phase, the Confirmatory Phase was designed to determine whether a flortaucipir PET scan visual read of "τAD++ pattern" predicts a higher risk of subjects' clinically meaningful cognitive and functional deterioration within 18 months of scan.

Participant flow

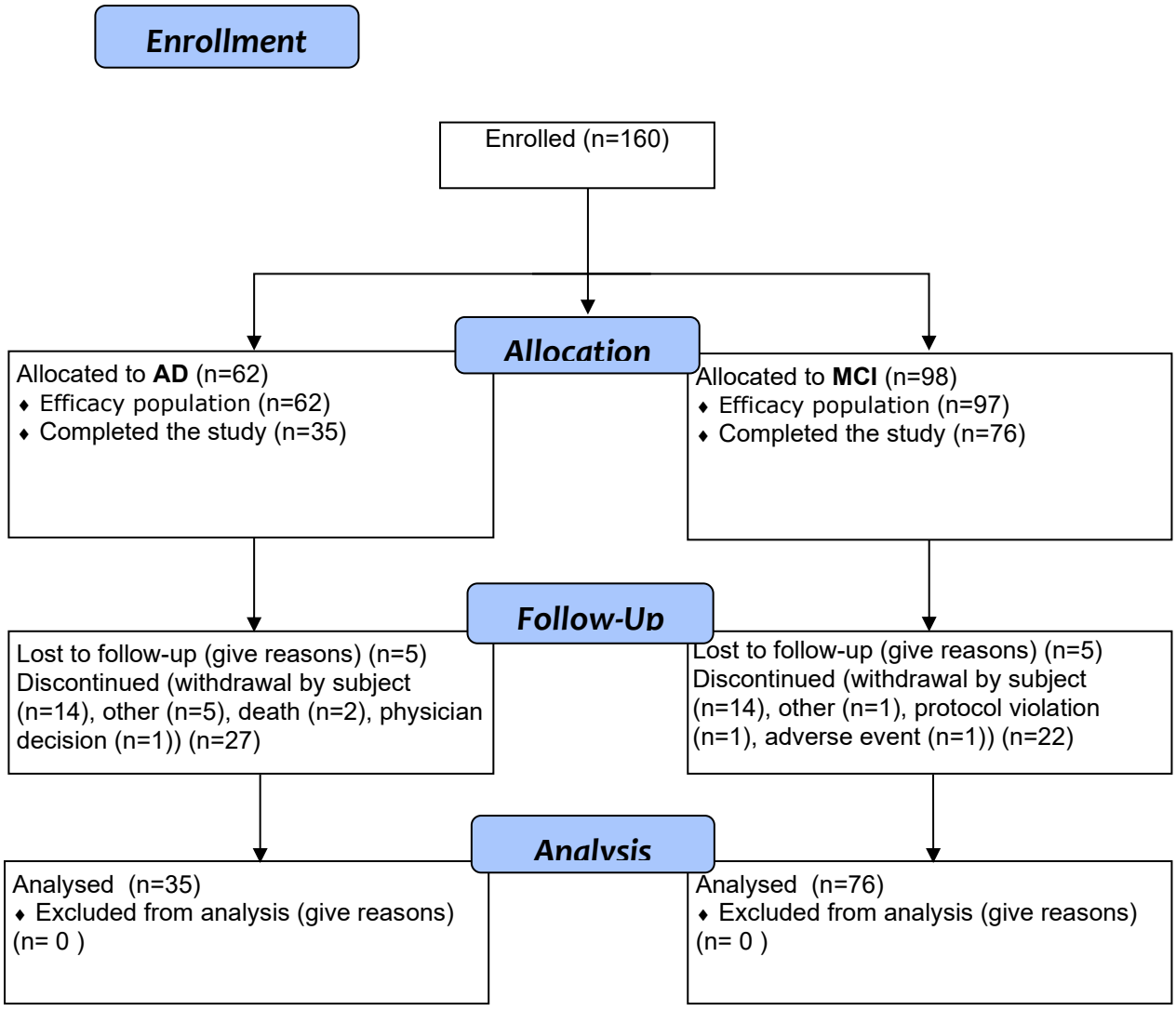


Figure 7 Subject disposition – Confirmatory Cohort

Conduct of the study A05C

The protocol was amended twice during the conduct of the study. In Protocol Amendment 1 (24 October 2014), the study was divided into 2 phases, Exploratory (first) Phase and Confirmatory (second) Phase. The Confirmatory Phase was to focus on the relationship of baseline PET tau images to change in longitudinal clinical measure. The second phase of the study was designed to provide independent validation of the relationships observed in the exploratory analyses of the first phase.

Despite the fact that the important protocol deviations are not presented in a grouped manner, there are no deviations that could have had an impact on the outcome of the study.

Baseline data A05C

Baseline flortaucipir PET images from 159 subjects (AD, n = 62; MCI, n = 97) were included in the primary analysis (Table 27). A total of 111 subjects (AD, n = 35; MCI, n = 76) completed the 18-month follow-up efficacy assessments for the study.

At baseline, 78 patients (49.1%) were τAD++, 11 (6.9%) were τAD+, and 70 (44.0%) were τAD- by majority read. Table in the Summary of Clinical Efficacy and below shows flortaucipir PET status broken down by baseline amyloid status.

Table 27: Study A05C, flortaucipir baseline PET status as a function of amyloid status

	Predicted to Progress	**Not Predicted to Progress in 18 Months**	
	τAD++	**τAD+**	**τAD-**
Aβ+ (N = 101)	76	9	16
Aβ- (N = 58)	2	2	54
Total	78	11	70

Abbreviations: Aβ = beta amyloid; PET = positron emission tomography.

Notably, 85/89 subjects (95.5%) with a τAD majority read flortaucipir PET scan result were positive for amyloid on a florbetapir PET scan. While not specifically designed to assess diagnostic performance, Study A05C may provide an even stronger test of the relationship between the τAD flortaucipir pattern and elevated amyloid, due to the higher proportion of Aβ- subjects (58/159), in contrast to Study A16, where only 9 of the 64 subjects in the primary analysis were amyloid negative by CERAD criteria.

The mean age of the safety population was 72.9 years (range of 68 to 80 years). The mean age of AD subjects was 73.6 years, compared with 72.5 years for MCI subjects.

The percentage of male and female subjects was 53.8% and 46.3%, respectively. Overall, 96.9% of subjects were Caucasian, followed by African Americans (1.9%) and Asians (1.3%). Most subjects (94.4%) were of non-Hispanic ethnicity.

Mean weight in the overall safety population was 75.63 kg, with a range of 63.7 to 86.1 kg, and a mean height of 168.36 cm (range of 160.0 to 176.3 cm).

Greater than half (53.1 %) of the subjects in the overall safety population had completed a college or university education and 23.1% had completed graduate school.

The proportion of subjects who were ApoE E4 positive was numerically higher in the AD group (48.4%) than in the MCI group (33.7%) group.

The mean MMSE score was 23.5 ± 2.23 for the AD group and 25.8 ± 1.64 for the MCI group. The mean ADAS-Cog 11 score was 17.7 ± 6.76 for the AD group and 12.1 ± 4.61 for the MCI group.

There are differences in the mean baseline values of the cognitive assessments between the AD and the MCI group, which is reasonable and anticipated. The lack of a comparator group creates difficulties in the interpretation of any potential prognostic ability of flortaucipir 18.

Numbers analysed A05C

A total of 160 subjects (AD, n = 62 and MCI, n = 98) were included in the enrolled population of the confirmatory cohort of the study (Table 28). The enrolled population consisted of all subjects that had signed informed consent and had data in the EDC system. The safety population (n = 160) consisted of all subjects who received at least 1 dose injection of either flortaucipir or florbetapir F 18. The efficacy

population (n = 159 [AD, n = 62 and MCI, n = 97]) included all subjects who received an injection of flortaucipir plus had valid flortaucipir imaging data available (either visual reads or SUVr). A total of 111 subjects (AD, n = 35 and MCI, n = 76) completed the Confirmatory Phase of the study.

Forty-nine subjects (AD, n = 27 [43.5%] and MCI, n = 22 [22.4%]) discontinued from the study (Table 29). Reasons for discontinuation are also presented in Table 29. One subject, 1147-935, discontinued from the study due to a diagnosis of severe breast cancer 500 days after last dose of study drug. This case was listed by the investigator as a discontinuation due to an AE, although the investigator determined this event was not related to study drug.

Study 18F-AV-1451-A05 Part II: Confirmatory Cohort

Table 28: Summary of subject disposition - all enrolled population

	AD (N=62)	**MCI (N=98)**	**Total (N=160)**
Total Number of Subjects			
All Enrolled Population [a]	62 (100.0%)	98 (100.0%)	160 (100.0%)
Safety Population [b]	62 (100.0%)	98 (100.0%)	160 (100.0%)
Florbetapir Safety Population [c]	62 (100.0%)	98 (100.0%)	160 (100.0%)
Baseline Scan	62 (100.0%)	98 (100.0%)	160 (100.0%)
Flortaucipir Safety Population [d]	62 (100.0%)	98 (100.0%)	160 (100.0%)
Baseline Scan	62 (100.0%)	98 (100.0%)	160 (100.0%)
Efficacy Population [e]	62 (100.0%)	97 (99.0%)	159 (99.4%)
Completed	35 (56.5%)	76 (77.6%)	111 (69.4%)
Terminated	27 (43.5%)	22 (22.4%)	49 (30.6%)

Note: AD = Alzheimer's Disease, MCI = Mild Cognitive Impairment. Percentages are based on the number of enrolled subjects in each diagnosis group.
[a] The All Enrolled Population consists of all subjects that have signed informed consent and have data in the EDC system.
[b] The Safety Population consists of all subjects that have received at least one dose injection of either Flortaucipir or Florbetapir F 18.
[c] The Florbetapir Safety Population consists of all subjects who received at least one dose injection of Florbetapir F 18.
[d] The Flortaucipir Safety Population consists of all subjects who received at least one dose injection of Flortaucipir.
[e] The Efficacy Population includes all subjects with a valid interpretable PET images and at least one clinical/cognitive assessment.
[f] Reason for Termination percentages use the Safety Population as the denominator.

Table 29: Summary of subject disposition - all enrolled population – reason for termination

	AD (N=62)	**MCI (N=98)**	**Total (N=160)**
Reason for Termination [f]			
Withdrawal by Subject	14 (22.6%)	14 (14.3%)	28 (17.5%)
Lost to follow-up	5 (8.1%)	5 (5.1%)	10 (6.3%)
Other	5 (8.1%)	1 (1.0%)	6 (3.8%)
Death	2 (3.2%)	0	2 (1.3%)
Protocol Violation	0	1 (0.6%)	1 (0.6%)
Adverse Event	0	1 (0.6%)	1 (0.6%)
Physician decision	1 (1.6%)	0	1 (0.6%)
Pregnancy	0	0	0
Study terminated by Sponsor	0	0	0
Technical Problems	0	0	0

The efficacy population included all subjects with a valid, interpretable PET image and at least 1 clinical/cognitive assessment, either from baseline or post baseline visits. The primary/secondary efficacy population include all subjects with a valid, interpretable PET image, and who have completed the relevant clinical/cognitive assessment from baseline, and any of post flortaucipir scan visits, 9month and/or 18-month visits.

A large number of patients were terminated from the study in the AD group (27/62) and in the MCI group (22/98) resulting in a percentage of 44% and 22% for the AD and MCI group, respectively and with the total number being as high as 49/111 (44%). Taking into consideration the relatively small number of patients included in the confirmatory phase compared to the patients with AD and MCI and the large number of discontinuations, the size of the study is of concern for the support of an indication to inform risk for progression of cognitive and functional impairment due to AD since this indication is no longer claim this major objection was no longer pursued.

Outcomes and estimation A05C

For each imaging session, subjects were to receive a single IV bolus injection of approximately 370 MBq (10 mCi) of flortaucipir F 18 followed by a saline flush. At approximately 80 minutes post dose, a continuous 20-minute brain scan (4 acquisitions of 5 minutes duration each) was obtained.

The primary objective of the Confirmatory (second) Phase was to assess the hazard ratio of subjects with flortaucipir scan rated as τAD++ versus subjects with scans rated as non-τAD++ in progressing to a clinically meaningful event (CDR-SB with 1 point or more increase) within 18 months of scan. In Study A05 exploratory data, 75% (15/20) of τAD++ subjects had an event while only 27% (7/26) of non-τAD++ subjects experienced an event. According to the applicant's assumptions that a similar τAD++ rate and that a similar event rate would be observed in the Confirmatory Cohort, a sample of 120 subjects in total should have provided approximately 90% power to detect a risk ratio of at least 2, under a 2-sided type I error rate of 5%.

In contrast to the Exploratory Phase, the Confirmatory Phase was designed to test a diagnostically oriented statistical hypothesis, specifically to determine whether a flortaucipir PET scan visual read of "τAD++" predicts a higher risk of subjects' clinically meaningful cognitive and functional deterioration within 18 months of scan. For the purpose of the primary analysis, clinically significant deterioration was defined as at least a 1 point worsening on the CDR-SB. A CDR-SB change of 1 point has been considered a potential measure of response in therapeutic trials. A CDR is an interview-based assessment of function in the real world. The average 18-month change in CDR-SB in clinical trials of early AD patients can be 2 or more points. An 18-month change of less than 1 CDR-SB point would mean the individual had little or no change in real world function, which would be clinically important in the context of a therapeutic trial. An alternative criterion for clinically significant change would be a change in the global CDR.

In the systematic review by Andrews et al (2019), across the overall cohort of National Alzheimer's Coordinating Center (NACC) participants (n = 19,566) with cognitive abilities ranging from normal to moderate-severe AD dementia, on average, a 1-2 point increase in CDR-SB, a 1-3 point decrease in MMSE, and 3–5 point increase in FAQ were indicative of a meaningful decline in the clinician's assessment. Hence, the chosen CDR-SB change of 1 point can be considered a potential measure of response in therapeutic trials.

The hazard ratio for the primary analysis was 1.581 with a lower bound of the 95% confidence limit of 0.968. Thus, the primary analysis did not meet the pre-specified success criteria (lower 95% confidence limits >1.0). However, as shown in the table below, analyses of other outcome parameters showed that nominally significant hazard ratios were obtained, indicating that a greater proportion of

τAD++ subjects passed the Prespecified threshold for clinically significant worsening, on other indices including CDR-SB change > 2.5, CDR Global stage change (change greater than 0), MMSE (≥3 point decrease), and Functional Activities Questionnaire (FAQ, ≥3 point increase), but not ADAS-Cog (≥4 point increase).

Table 30: Confirmatory cohort study A05C, hazard ratios for flortaucipir PET classification of τAD++ vs. non- τAD++

Test	HR	Confidence Limit	p-value
CDR-SB change >1	**1.581**	**0.968**, 2.581	**0.067**
CDR-SB change ≥2.5	2.595	1.261, 5.339	0.0096
CDR Global change >0	2.371	1.320, 4.259	0.0039
MMSE change, ≥3 point decrease]	2.506	1.285, 4.888	0.007
FAQ change, ≥3 point increase	2.785	1.671, 4.640	0.0001
ADAS-Cog change, ≥4 point increase	1.359	0.767, 2.408	**0.2934**

Abbreviations: ADAS-Cog = Alzheimer's Disease Assessment Scale Cognitive subscale 11; CDR-SB = Clinical Dementia Rating Scale Sum of Boxes; CL = confidence limit; FAQ = Functional Activities Questionnaire; MMSE = Mini-Mental State Examination; HR = hazard ratio; PET = positron emission tomography.

Interestingly a nominally statistically significant hazard ratio (lower bound, 95% confidence limit >1.0) was obtained for the pre-specified threshold for clinically significant change in CDR Global (any worsening, such as ≥0.5-point increase), as well as for a modified CDR-SB threshold (≥2.5-point change). Pre-specified criteria for clinically significant change were also met for MMSE (≥3 point change), FAQ (change ≥3), albeit not for the ADAS.

According to the applicant, the statistical result for the primary analysis may also have been influenced by the method for calculating the hazard ratio. The hazard ratio provides an estimate of relative risk, based on time to event, adjusted based on certain factors. Hazard ratios are appropriate for freely occurring event-based outcomes (for example, death and heart attack), but can be more difficult to interpret when applied to criterion test scores, since subjects may meet the criteria on 1 occasion, but not at a later occasion (a particular problem when the criterion threshold is set low, as in CDR-SB change of 1). In the current study, 4 subjects in the non-AD++ cohort, and 1 subject in the AD++ cohort met the criterion of CDR-SB change at 9 months but not at 18 months. If the outcome criteria are adjusted such that cases that show a 1-point increase at 9 months, but not 18 months, are considered as non-events, then the observed hazard ratio is 1.937 (lower bound, 95% confidence limit 1.157, p = 0.0118). Consistent with this analysis, unadjusted risk ratios, which utilise only the final, 18-month outcome for calculation, also showed a significant elevation for cases classified as τAD++ vs. non-τAD++ for 4 of the readers. Thus, both the majority read (adjusted for patients that showed change at 9, but not at 18 months) and the read of 4 of the 5 individual raters showed a significant increase in risk of a clinically significant 1-point worsening of CDR-SB being evident at 18-months follow-up.

Additionally, the applicant is of the view that it is important to consider that the primary statistical comparison (τAD++ vs. non-τAD++) is not a comparison of subjects with advanced NFT stage (e.g., τAD++) versus subjects devoid of pathologic tau aggregates. The non-τAD++ group includes 11 patients with a moderate AD tau pattern (τAD+). Of these, 9 were Aβ+ by amyloid scan, and thus, likely meet pathologic criteria for at least Intermediate, if not High AD neuropathic change (albeit with less widespread tau distribution than the τAD++ group) with some degree of associated cognitive progression. Moreover, 16 of 70 τAD- subjects were also amyloid positive, raising the possibility that these subjects had early Alzheimer's pathologic change (Jack et al, 2018). Thus, these subjects might also be expected to show some cognitive deterioration. The τAD++ scans therefore should not be

thought of as identifying the only patients at risk for cognitive deterioration. The applicant believes that τAD++ scans rather potentially identify a subgroup of patients at greater risk for near term (18 months) deterioration above and beyond their clinical diagnosis and amyloid status.

In summary, the applicant noted that although the primary efficacy outcome was not met (CDR-SB hazard ratio lower 95% confidence limit = 0.968, p = 0.0670), a consistent pattern was observed across multiple outcome measures indicating that a τAD++ pattern on flortaucipir PET was associated with an increased hazard and/or risk ratio for clinically significant deterioration, and increased mean deterioration, by comparison to a τAD+ or a τAD- pattern. Importantly, these findings probably underestimate the relationship between flortaucipir PET signal and cognitive deterioration in an unselected population, since these results were obtained in a population with MMSE 20 to 27, that was intentionally chosen (based on results from the Exploratory Phase) to eliminate subjects with high likelihood of high tau and rapid deterioration (MMSE <20) or low tau and slow deterioration (MMSE >27). Taken together, these results could have suggested that flortaucipir PET signal may have utility in identifying patients with increased risk for cognitive decline. However, this primary endpoint was not met (CDR-SB hazard ratio lower 95% confidence limit = 0.968, p = 0.0670).

Ancillary analyses A05C (

***Table 31*)**

According to the applicant, in order to better understand the degree to which flortaucipir PET signal provides information regarding near term cognitive decline that is beyond the background information easily obtained in clinical/demographic information, sensitivity analyses were performed looking at LS mean as a function of tau PET signal corrected for baseline age, ANART score, and baseline score for the respective outcome variable. Because there is a high degree of overlap between τAD pattern and Aβ status, the first sensitivity analysis evaluated the impact of τAD pattern in Aβ+ subjects only; thus, also providing information above that which could be determined from amyloid status alone. In each case there was a significant difference (or in the case of CDR-SB a trend toward difference, p = 0.0962) in LS mean change from baseline between the Aβ+, τAD++, (excluding the 2 Aβ- subjects) and the Aβ+, non-τAD++ groups, suggesting that flortaucipir PET signal provides information beyond that which is available from amyloid status alone.

A second sensitivity analysis looked separately at the effects of flortaucipir PET status in subjects with amnestic versus non-amnestic MCI. There was overlap between amnestic and flortaucipir PET status such that 34 of 39 τAD++ subjects were amnestic, but not all amnestic subjects were τAD++ (32 of 57 amnestic subjects were non-τAD++). Among amnestic subjects, significantly greater decline in CDR-SB, MMSE, and FAQ was seen for τAD++ as opposed to non-τAD++ subjects. Among non-τAD++ subjects, there were no differences between amnestic and non-amnestic subjects. Both amnestic and non-amnestic non-τAD++ subjects declined slowly by comparison to τAD++ subjects. Thus, combined flortaucipir PET visual read classification and amnestic status yielded better risk stratification than amnestic status alone.

Table 31 below shows the mean change for τAD++ and non-τAD++ for CDR-SB, MMSE, FAQ, and ADAS-Cog11. For each variable, mean worsening from baseline was significantly greater for τAD++ than non-τAD++ subjects. For example, Aβ+ τAD++ showed statistically greater mean cognitive and functional decline than either Aβ+ non-τAD++ (except on CDR-SB) or Aβ- subjects on each of the cognitive and functional test scales. Similarly, τAD++ amnestic MCI showed significantly more rapid decline (except on ADAS-Cog11) than non-τAD++ amnestic MCI.

Table 31: Study A05C, LS mean cognitive and functional test score change from baseline to 18 months τAD++ vs Non-τAD++

Test	τAD++ LS Mean (± SEM)	Non-τAD++ LS Mean (± SEM)	p-value
CDR-SB Change	2.09 ± 0.35	0.66 ± 0.32	.0030
MMSE change	-3.29 ± 0.67	-0.33 ± 0.59	.0013
FAQ change	6.93 ± 0.75	1.78 ± 0.69	<.0001
ADAS	6.41 ± 1.00	2.50 ± 0.91	.0056

Abbreviations: ADAS = Alzheimer's Disease Assessment Scale; CDR-SB = Clinical Dementia Rating Scale Sum of Boxes; FAQ = Functional Activities Questionnaire; LS = least squares; MMSE = Mini-Mental State Examination; SEM = standard error of the mean.

Another exploratory analysis used a multivariate regression model to explore the contribution made by flortaucipir PET imaging to predicting cognitive change against a background of clinical/demographic information including diagnosis stage (MCI/AD), baseline MMSE, age, years of education, and ApoE4 status. Although the model assigned more of the variance to the MMSE factor (excluding flortaucipir Partial r^2=0.1163, p=0.0004 and including flortaucipir Partial r^2=0.1163, p=0.0003), flortaucipir PET tau status remained a significant factor (Partial r^2=0.0392, p=0.0334).

It can be anticipated that patients with Aβ+ τAD++ would show a greater cognitive and functional decline compared to either Aβ+ non-τAD++ or Aβ-. However, these analyses cannot confirm the prognostic value of flortaucipir in patients with AD. It is important to note also that due to the negative result in the primary analysis, no confirmatory claim can be based on these data and the sensitivity analyses.

Secondary Analysis of Efficacy – Assessment of sensitivity/specificity

Study A05C and PX01 also included an assessment of sensitivity/specificity for flortaucipir PET positivity to pre-specific thresholds for cognitive and functional decline. Similar to the analyses of diagnostic performance in Studies A16 and FR01, the hypothesis tested was that, of the 5 independent imaging physicians, at least 3 will have the lower bounds of 2-sided 95% CIs ≥50%, for both sensitivity and specificity. This analysis was included in the studies to satisfy a request from FDA for sensitivity and specificity to predict progression. This endpoint was not met, and since Lilly does not intend to make claims regarding the diagnostic performance of flortaucipir for prediction of progression for individual patients, this endpoint was not further discussed in the applicant´s summary.

In the case of study A05 Confirmatory Phase (Table 32*)* the secondary analysis of efficacy regarding an Assessment of sensitivity/specificity is presented below.

The secondary objective of the Confirmatory Phase was to assess the diagnostic performance of baseline tau positivity according to a flortaucipir scan visual read (τAD++, vs. non-τAD++), for predicting subjects' clinically meaningful cognitive and functional deterioration within 18 months of scan, as measured by the CDR-SB scales. This analysis used dichotomised CDR-SB change (1 point or more increase vs. otherwise) as a TS to assess the diagnostic performance of baseline tau status as determined by flortaucipir scan interpretation. The assessments were conducted for each of the 5 independent imaging readers. Sensitivity, specificity, NPV, PPV, LR+, and LR- were calculated. The hypothesis tested was that, of the 5 independent imaging physicians, at least 3 will have the lower bounds of 2-sided 95% CIs ≥50%, for both sensitivity and specificity. As shown in the following Table 37 for all 5 readers the lower confidence limits of specificity exceeded the target of 50%, but the lower confidence limits for sensitivity did not exceed 50% for any of the readers. However, the lower confidence limit for overall accuracy exceeded 50% and the lower confidence limit for relative risk ratio (PPV/(1-NPV)) exceeded 1.0 for 4 of the 5 readers.

Table 32 Reader diagnostic statistics for baseline Tau status predicting CDR-SB clinically meaningful change (≥1) – confirmatory phase

Reader	Sensitivity % (CL)	Specificity % (CL)	PPV % (CL)	NPV % (CL)	Relative Risk Ratio (CL)
1	59.6% (**46.1%**, 71.8%)	65.5% (52.7%, 76.4%)	60.8% (47.1%, 73.0%)	64.4% (51.7%, 75.4%)	1.708 (1.136, **2.568**)
2	53.8% (**40.5%**, 66.7%)	65.5% (52.7%, 76.4%)	58.3% (44.3%, 71.2%)	61.3% (48.8%, 72.4%)	1.507 (1.016, **2.235**)
3	55.8% (**42.3%**, 68.4%)	65.5% (52.7%, 76.4%)	59.2% (45.2%, 71.8%)	62.3% (49.7%, 73.4%)	1.570 (1.055, **2.336**)
4	59.6% (**46.1%**, 71.8%)	65.5% (52.7%, 76.4%)	60.8% (47.1%, 73.0%)	64.4% (51.7%, 75.4%)	1.708 (1.136, **2.568**)
5	50.0% (**36.9%**, 63.1%)	65.5% (52.7%, 76.4%)	56.5% (42.2%, 69.8%)	59.4% (47.1%, 70.5%)	1.391 (0.942, **2.055**)

Abbreviations: CDR-SB = Clinical Dementia Rating Scale Sum of Box; CL = confidence limit; NPV = negative predictive value; PPV = positive predictive value.

The secondary objective of study A05C was not met, since the lower limit of the confidence interval limits for sensitivity did not exceed 50% for any of the readers. In addition, the relative risk ratios ranged from 1.391 to 1.708, and the upper limits of the 95% CI of the relative risk ratios were all above 2.0 ranging from 2.055 to 2.568. The applicant did not discuss these results, as they did not intend in the end to make claims regarding the diagnostic performance of flortaucipir for prediction of progression for individual patients. However, this is exactly what would have been expected if such a prognosis claim had been included in the wording of the indication.

The mean value for sensitivity for every reader was between 50 and 60%, but not greater. The sensitivity is lower than specificity in Study A05C, with specificity having a mean value of only 65.5% for every reader. Please see also comments in the section for the intra and inter-reader variability (below).

Instead of the risk ratios the likelihood ratios can be considered to be more informative, since they would be independent of the prevalence (hence severity of the disease) under the assumption that sensitivity and specificity can be transferred. For example, likelihood ratios for a positive outcome indicate the factor by which the odds of worsening are increased after obtaining a positive test result. Obviously, a likelihood ratio of 1 would indicate that no information at all is given by the test.

Whereas the applicant did provide likelihood ratios confidence intervals for likelihood ratios were neither pre-planned nor presented, which is unfortunate due to a number of possible options. In any case, difference between likelihood ratios (LR+) and risk ratios is small. In both cases, the amount of information given by the test appears rather small. Considering the lower limit of the classical (asymptotic) 95% confidence interval, the values for all readers are close to 1, for reader 5 even lower than 1. Hence, it cannot be shown that the test is very informative. Even considering the point estimate, the amount of obtained information may not be larger than simple clinical assessment. It should further be noted that in case the argument of the applicant was that some of the readers show a LR+ larger than 1 (but not all), a multiplicity issue would arise. Adjusted confidence intervals for the likelihood ratio would all include 1.

In any case, due to the negative result in the primary analysis no confirmatory claim can be made.

Intra and inter-reader variability (study A05C)

Similar to Studies A16 and FR01, inter-reader reliability in the prognosis studies (A05C and PX01) was assessed primarily using Fleiss' kappa with its 95% CI for scan reads dichotomised to tAD++

(Advanced AD pattern, expected to progress clinically) vs. non-tAD++ (Moderate AD pattern or negative for AD, not expected to progress clinically). There were no pre-specified hypotheses tested for inter-reader reliability.

Intra-reader agreement was evaluated for prognosis reads in A05C only by randomly choosing 20 scans to be read twice by each reader. A Cohen's kappa with 95% CI was calculated for each of the 5 readers to show the intra-reader scan reading consistency.

Inter-reader reliability was assessed for prognosis reads (Advanced [τAD++] AD flortaucipir PET pattern vs. non-Advanced [non-τAD++] AD flortaucipir PET pattern). Agreement among reader pairs was >92% and the calculated kappa was >0.83 for all pairs. The overall Fleiss Kappa was 0.8893 (CI: 0.8401, 0.9385), indicating excellent agreement across all readers. Intra-reader agreement on the 20 cases interpreted twice by each reader was at least 90% and read/re-read kappa was >0.79 in all cases, indicating excellent reliability of the interpretation.

In quantitative analyses, AD-signature-derived neocortical target VOI was significantly correlated with change from baseline for each cognitive and functional test score (Table 33).

Table 33: Study A05C, correlation (Pearson r) between baseline AD-signature VOI SUVr and change from baseline cognitive and functional test scores

	MMSE	**FAQ**	**ADAS-Cog 11**	**CDR-SB**
N	107	108	108	108
r	-0.4147	0.2690	0.3950	0.3243
p-value	<.0001	.0047	<.0001	.0006

Abbreviations: AD = Alzheimer's disease; ADAS-Cog 11 = Alzheimer's Disease Assessment Scale-Cognitive subscale 11; CDR-SB = Clinical Dementia Rating Scale Sum of Box; FAQ = Functional Activities Questionnaire; MMSE = Mini-Mental State Examination; N = number of patients; SUVr = standardised uptake value ratio; VOI = volume of interest.

Table in the CSR and Table 34 below, summarise read/re-read agreement on the 20 cases that were interpreted twice by each reader. Observed agreement was at least 90% and read/re-read kappa was greater than 0.79 in all cases, indicating excellent reliability of the interpretation.

Table 34: Intra-reader agreement of visual qualitative rating of flortaucipir PET images – confirmatory phase

	Observed Agreement (%)	**Kappa (95% CI)**
Reader 1	90.00	0.7917 (**0.5180**, 1.0000)
Reader 2	95.00	0.9000 (0.7099, 1.0000)
Reader 3	90.00	0.8000 (**0.5424**, 1.0000)
Reader 4	95.00	0.894 (0.6915, 1.0000)
Reader 5	100.00	1.000 (1.0000, 1.0000)

Abbreviation: PET = positron emission tomography.

According to the applicant, intra- and inter-reader reliability analyses suggested agreement on 90% or more of cases and kappa was greater than 70% for all intra- or inter-reader pairings. However, for two of the readers the lower limit of the 95% CI for Fleiss' kappa lower than 0.6, suggesting that the level of agreement was not that reassuring. In the sensitivity and specificity analysis, the lower confidence limits for sensitivity did not exceed 50% for any of the readers. These raised concerns with respect to the reliability and sensitivity of the ability of flortaucipir F 18 in predicting a change in patient's functional outcome. The applicant did not discuss these results, as they did not intend in the end to make claims regarding the diagnostic performance of flortaucipir for prediction of progression for

individual patients. However, this is exactly what would have been expected if such a prognosis claim had been included in the wording of the indication.

Study 18F-AV-1451-PX01 (Study PX01)

Evaluation of the Relationship between Baseline Flortaucipir PET Signal and Cognitive Change in Subjects with Early Alzheimer's Disease Participating in the I8D-MC-AZES Protocol Addendum D5010C00009 (2.1) (Tau Imaging)

Methods PX01

The second pivotal study for the claim of informing on the risk of progression, Study PX01, was designed to take advantage of a therapeutic trial cohort that had flortaucipir PET scans and up to 2 years of cognitive follow-up to further evaluate the relationship between flortaucipir PET signal and cognitive and functional decline. Design details of Study PX01 are discussed in Summary of Clinical Efficacy.

It is worth mentioning that this was a retrospective analysis study with the exclusion of drop-outs, which may have consequently created a selection bias since the results may not be generalizable across a broader population. This is a basic limitation of this study. In addition, the applicant was asked to clarify on what basis and when it was decided that the PX01 study is a confirmatory clinical trial (i.e. pivotal study) as opposite to other studies with very similar design which were considered to be supportive studies only. It was unclear on how the GCP could be assured as the data for PX01 study were collected as part of another terminated prematurely study. The impact of the early termination of the AZES study was discussed in this context. Taking into consideration the methodological limitations, it was considered that study PX01 does not fulfil robustness conditions for a confirmatory clinical trial and therefore it can only be considered as a supporting trial. It was unclear on how the GCP could be assured as the data for PX01 study were collected as part of another prematurely terminated study. The impact of the early termination of the AZES study was discussed in this context (see Exploratory Analyses and Table 46 below).

Study Participants PX01

In contrast to Study A05C, no new subjects were enrolled in this study PX01, since this study used flortaucipir PET scans from placebo subjects of the AZES trial. The study population included placebo patients from parent Study AZES, which consisted of subjects who:

- were amyloid positive by florbetapir PET or lumbar puncture,
- aged 55 to 85 years with MCI due to AD or probable AD by NIA-AA criteria,
- had MMSE of 20 to 30, inclusive,
- had a CDR Global score of 0.5 (MCI), or 0.5 or 1 (AD) with a memory box score≥0.5, and
- had a score of ≤85 on the Delayed Memory Index of the RBANS.

Subjects that completed a baseline flortaucipir scan and at least 18 months of cognitive follow-up were included in the Study PX01 analysis.

All subjects were amyloid positive by florbetapir PET or lumbar puncture. Subjects whose scans were read in Study PX01 had a baseline flortaucipir scan, an 18-month CDR assessment, and met Study AZES inclusion criteria.

Images from 205 subjects who had a valid baseline flortaucipir scan, and a CFB value of CDR-SB at 18 months were interpreted.

In comparison to the A05C study this is a retrospective analysis in a population that had a higher baseline risk for AD. One inclusion criterion in study PX01 was amyloid positivity in contrast to study A05 which did not have that subject inclusion criterion reflecting a different study population.

Treatments PX01

No study drug was administered in PX01. At each flortaucipir imaging visit in parent Study AZES, all subjects received a single IV bolus administration of approximately 240 MBq (6.5 mCi) of flortaucipir F 18 Injection followed by a saline flush. The applicant was asked to present how many patients received an active treatment (lanabecestat 20mg and 50 mg), while participating in the parent study. The applicant clarified that a total of 68 subjects received 20 mg of lanabecestat (20 subjects with MCI and 48 subjects with AD), 64 subjects received 50 mg of lanabecestat (22 subjects with MCI and 42 subjects with AD), and 73 subjects received placebo.

Flortaucipir F 18 was administered as a 240 MBq, which is at a different concentration from the recommended single intravenous dose for flortaucipir (18F) injection of 370 MBq (10 mCi) for the intended use of Tauvid.

The applicant has discussed the use of two different doses i.e. 370 MBq in study A05C and 240 MBq in study PX01, their relation to each other and their contribution for the initially proposed indication to inform on the risk for progression of cognitive and functional impairment due to AD.

Objectives PX01

Study PX01 evaluated the relationship between baseline flortaucipir PET Signal and cognitive and functional change in subjects with early AD participating in the I8D-MC-AZES Tau Imaging Protocol Addendum.

<u>Primary Objective</u>

- To assess whether a visual interpretation (τAD++ vs. non-τAD++ pattern) of the baseline flortaucipir PET scan can predict the risk of subjects' clinically meaningful cognitive and functional deterioration within 18 months of scan, as measured by the CDR-SB CFB.

<u>Secondary Objectives</u>

- To assess whether a visual interpretation (τAD++ vs. non-τAD++ pattern) of the baseline flortaucipir PET scan can predict the risk of subjects' clinically meaningful cognitive and functional deterioration within 18 months of scan, as measured by MMSE, ADASCog11, FAQ, and CDR global change from baseline.
- To assess the relationship between visual interpretation (τAD++ vs. non-τAD++ pattern) of the baseline flortaucipir PET scan and magnitude of cognitive and functional deterioration within 18 months of scan, as measured by the mean change from baseline of CDR-SB, MMSE, ADAS-Cog11, and FAQ.
- To assess inter-reader reliability of the flortaucipir F 18 PET scan visual interpretation by 5 independent, blinded readers. In Study PX01, 4 of the 5 readers had previously been trained on the read method as part of Study A05C.

Safety endpoints were not evaluated in this study.

One of the secondary objectives of this study (as in the case of A05C) was to assess the inter-reader reliability of flortaucipir (τAD++ vs. non-τAD++) using Fleiss' Kappa statistics (please see below).

Exploratory Objectives

- To assess whether a visual interpretation (τAD++ vs. non-τAD++ pattern) of the baseline flortaucipir PET scan can predict the risk of subjects' clinically meaningful cognitive and functional deterioration within 24 months of scan, as measured by the CDR, MMSE, ADAS-Cog11, and FAQ change from baseline.
- To assess the diagnostic performance of baseline flortaucipir F 18 PET scan in predicting clinically meaningful deterioration at 18 months

Outcomes/endpoints PX01

According to the protocol, the following were the primary and secondary objective analyses.

Primary analysis

The primary analysis for this study evaluated whether the baseline tau status as determined by flortaucipir F 18 scans will predict the risk of subjects' clinically meaningful cognitive and functional deterioration within 18 months of scan. Clinically meaningful deterioration (CMD) for the primary objective analysis was defined as CDR-SB change from baseline with an increase of 1 point or more.

Tau Status Based on Flortaucipir

Flortaucipir images were visually assessed by 5 independent readers blinded to all demographic and clinical data for the subjects.

Clinical Dementia Rating Scale

The CDR (Berg 1988) examines 6 categories of cognitive functioning domains. Each domain is scored on a scale ranging from 0 to 3 (including 0.5). A CDR-SB was generated as the sum of the values in each of the 6 domains. The CDR-SB sum scores range from 0 to 18, with higher scores indicating greater cognitive impairment. The CDR was scheduled to be performed at screening, and at 6 months, 12 months, 18 months, and 24 months postbaseline. For the primary analysis, the CDR-SB score CFB was dichotomised in 2 groups: CDR-SB score change of 1 point or more at 18 months, or otherwise. The CDR global score (a global rating of dementia) was derived based on CDR item/domain values. The CDR global was dichotomised in 2 groups for exploratory analysis: any change in CDR global, or otherwise.

Secondary analyses

Clinically Meaningful Deterioration at 18 Months by Tau Status assessed by:

Mini-Mental State Examination decreased by 3 points or more

The MMSE (Folstein et al. 1975) is a brief instrument used to assess cognitive function in elderly patients. The range for the total MMSE score is 0 to 30, the sum of each correct answer, with higher scores indicating better cognition. The score on a continuous scale was used for analysis, when assessing the average change from baseline by tau status.

Alzheimer's Disease Assessment Scale—Cognitive subscale increased by 4 points or more

The Alzheimer's Disease Assessment Scale (ADAS; Rosen et al. 1984) was designed to assess the severity of the dysfunction in the cognitive and non-cognitive behaviours characteristic of persons with AD. The cognitive subscale of the ADAS, the ADAS-Cog11, consists of 11 items assessing areas of

function most typically impaired in AD: orientation, verbal memory, language, and praxis. The overall score for ADAS-Cog11 ranges from 0 to 70, with higher scores indicating greater disease severity; the ADAS-Cog11 is calculated as the sum of all 11 individual component scores. The ADAS-Cog11 total score was used as a continuous variable for the exploratory analysis to assess the average change from baseline by tau status.

Pfeffer Functional Activities Questionnaire, (Pfeffer et al. 1982) increased by 3 points or more

Functional status is conceptualised as the "ability to perform self-care, self-maintenance, and physical activities." The FAQ was developed to assess instrumental activities of daily living involving higher level functional skills such as shopping alone, writing checks, and remembering appointments. The FAQ asks the informant to rate the patient's ability using the following scoring system: Dependent = 3; Requires assistance = 2; Has difficulty but does by self = 1; Normal = 0; Never did [the activity] but could do now = 0; Never did and would have difficulty now = 1. The sum scores range from 0-30, where higher scores indicate greater functional impairment. The FAQ sum score was used as a continuous variable in analysis when assessing the average change from baseline by tau status.

CDR global with any increase

The CDR global change from baseline was categorised into 3 groups: >0, = 0 and <0. In addition, CDR global change from baseline was categorised as any increase (change >0) vs. no change or decrease (change ≤0).

Mean Change of Cognitive/Functional Assessments at 18 Months by Flortaucipir PET Visual Interpretation (τAD++ vs. non-τAD++)

Mean change at 18 months by tau status of:

- CDR-SB,
- MMSE,
- ADAS-Cog11, and
- FAQ.

Inter-reader Reliability

Inter-reader reliability of flortaucipir scan interpretation (τAD++ vs. non-τAD++) across the 5 independent readers

Exploratory Analyses

Cognitive/Functional Assessments Deterioration at 24 Months, by Tau Status

- Subjects who completed both baseline and 24-month clinical assessments.
- Assessment of whether a visual interpretation (τAD++ vs. non-τAD++ pattern) of the baseline flortaucipir PET scan can predict the risk of subjects' clinically meaningful cognitive and functional deterioration within 24 months of scan, as measured by:
    - CDR (CDR-SB and CDR Global),
    - MMSE,
    - ADAS-Cog11, and
    - FAQ change from baseline.

Clinically Meaningful Deterioration by τAD vs τAD- Tau Status

- Contrast scan visual interpretations of τAD (τAD+/ τAD ++) vs. τAD- using both 18 months and 24 months follow-up data.

Diagnostic Performance of Flortaucipir Scan in Predicting Clinical Deterioration

- Diagnostic performance (sensitivity/specificity) assessed by using CMD as truth standard (TS).

Safety Measurements

Safety measurements were not collected in this study.

It should be noted that this "pivotal" study PX01 is a retrospective evaluation of scans from 205 subjects who received a lower dose (240 MBq) compared to study A05C (370 MBq).

Sample size PX01

This study was planned to include up to 205 subjects who had a valid baseline flortaucipir scan, and a change from baseline value of CDR-SB at 18 months. In Study TZAX, which had a similar design and entry criteria, approximately 75% of subjects were rated as τAD++ and 25% were rated as non-τAD++. Among non-τAD++ subjects (reference group), approximately 40% had an increase of 1 point or more on CDR-SB at 18 months follow up visit. Assuming a distribution of visual reads and CDR-SB change in this study is similar to that for TZAX, a sample of 205 subjects was estimated to provide 90% power to detect a risk ratio of 1.65 or larger, under a two-sided type I error rate of 5%.

As already mentioned, the study population included patients from parent Study AZES. Images from 205 subjects who had a valid baseline flortaucipir scan, and a CFB value of CDR-SB at 18 months were interpreted. This number of images is considered reasonable for the intended evaluation. However, no control group was used.

Randomisation and blinding (masking) PX01

There was no randomisation. No new subjects were recruited, and no drug was administered in this study PX01. This report describes the results of testing an in-person reader training program using images collected in a substudy of parent study I8DMC-AZES (AZES). This study was conducted in the United States.

However, images were randomised before interpretation. A description of the methods for image randomisation and coding is provided in Section 6.2 of the IRC located in the Documentation of Inter-Laboratory Standardization Methods and Quality Assurance Procedures appendix.

Blinding (masking)

A blinded design was not used in PX01 for flortaucipir exposure because all subjects received flortaucipir F 18. Subjects and investigators were blind to lanabecestat treatment throughout the trial.

Physician readers were blinded to subject diagnosis and all demographic and clinical data from Study AZES. The readers were not to have participated in the on-site evaluation of subjects at the study site, or to have been an employee of the sponsor or the Imaging Core Laboratory (ICL). Only the ICL read monitors were to be present with the readers while the blinded reads were being performed. The readers were to receive no assistance from the ICL staff regarding image interpretation. Additional details on the blinded read and a description of the methods for image randomisation and coding is provided in the IRC and in the Blinded Read Manual located in the Documentation of Inter-Laboratory Standardization Methods and Quality Assurance Procedures appendix.

Statistical methods PX01

Definitions and Conventions

All analysis was performed using SAS version 9.2 or higher.

Data were summarised using descriptive statistics (number of subjects [n], mean, standard deviation [SD], median, minimum [min], and maximum [max]) for continuous variables and frequency counts and percentages for discrete variables. Percentages were rounded to one decimal place, except 100% will be displayed without any decimal places and percentages will not be displayed for zero counts.

The tables and listings were numbered using a decimal system to reflect main levels of unique tables and listings and sub-levels of replicate tables and the listings with maximum two digits per level (e.g., Table XX.YY.ZZ…). Tables will be numbered as 14.YY.ZZ. Baseline analysis was reported in table series 14.1, and efficacy analysis in series 14.2. Listings will be numbered as 16.YY.ZZ. Unless otherwise specified, hypothesis testing will be two-sided with type I error rate of 0.05.

Adjustments for Covariates

Multivariate models using cognitive or function assessments as the dependent variables were adjusted for treatment arm [lanabecestat (20mg, 50 mg) or placebo], baseline cognitive score, years of education (categorical), and age.

Handling of Dropouts or Missing Data

Dropout subjects were not replaced in this study. For situations with no rules for handling missing data the default will be no imputation.

Likelihood-based mixed effects models for repeated measures were used to handle missing data for the cognitive/functional assessment mean CFB analyses. The model parameters were simultaneously estimated using restricted likelihood estimation incorporating all of the observed data. Estimates have been shown to be unbiased when the missing data are missing at random and when there is ignorable non-random missing data.

Repeated measures analyses only used data from visits where the data was scheduled to be collected. When subjects discontinue from the study early, there may be efficacy data measurements at visits where the variables were not scheduled to be collected. These data appeared in listings only.

In addition to design differences from study A05C, the statistical test for the primary analysis in Study PX01 evaluated a risk ratio, rather than a hazard ratio as used in Study A05C. The primary rationale for use of a hazard ratio in Study A05C was that the hazard ratio allows one to use all subjects, including those that discontinue prematurely, in comparing event rates between 2 groups (that is, тAD++ versus non-тAD++). Since subjects in Study PX01 were selected on the basis of having completed the required 18-month follow-up period, the hazard ratio was not necessary. In addition, risk ratios have the advantage of requiring no assumption about proportionality of hazard function between groups or time to occurrence of an event, but simply compare the status of patients in different groups at a fixed time point (that is, the 18-month CDR-SB score). According to the applicant risk ratios are arguably more appropriate for evaluating cognitive change in this particular trial since all subjects completed 18 months follow-up and there were no concerns about the bias introduced by missing data (for example, early drop-outs from the study).

Since this was a retrospective analysis with the exclusion of drop-outs this may create a selection bias because the results may not be generalizable across a broader population.

Instead of the considered risk ratios, likelihood ratios are considered to be more appropriate, since they would be independent of the prevalence (hence severity of the disease) under the assumption

that sensitivity and specificity can be transferred. E.g. likelihood ratios for a positive outcome indicate the factor by which the odds of worsening is increased after obtaining a positive test result. Obviously, a likelihood ratio of 1 would indicate that no information at all is given by the test.

Results PX01

Participant flow

A participant's flow is not available since no new subjects were enrolled in this study. The scans of 205 subjects who had a valid baseline tau scan and CDR assessments from baseline and 18 months in the AZES PET substudy were evaluated.

Recruitment PX01

There were no new subjects, only those in study AZES.

Conduct of the study PX01

Based on the nature of the study, it is expected that no protocol deviations could have occurred.

Baseline data PX01

Baseline demographic characteristics are presented and summarised in the following Tables . A listing of patient demographic characteristics is provided in the Demographic Data.

The mean age of the analysis population was 71.0 years (range 55 to 85 years). The mean age of subjects with AD was 70.4 years compared with 72.3 years for subjects with MCI.

The percentage of male and female subjects was 51.2% and 48.8%, respectively. Overall, 83.9% of subjects were White or Caucasian, followed by Asian (13.2%), Black or African American (2.4%), and Other (0.5%).

Greater than half (64.9%) of the subjects had completed 13+ years of education, 27.8% had completed 10 to 12 years of education, and 7.3% of subjects had completed 6 to 9 years of education.

There was no comparator group in this study, as in the case of study A05C. The two groups AD and MCI due to AD had differences between them in the cognitive assessments, since the level of impairment was different.

Numbers analysed PX01

A summary of the scans evaluated by тAD and disease status is provided in Table 35.

Table 35: Summary of scan reads by clinical diagnosis

	Mild AD (N=141)	MCI Due to AD (N=64)	Total (N=205)
Available scans	141	64	205
Scans read	141	64	205
τAD++, n (%)	120 (85.1)	42 (65.6)	162 (79.0)
τAD+, n (%)	9 (6.4)	6 (9.4)	15 (7.3)
τAD-, n (%)	12 (8.5)	16 (25.0)	28 (13.7)

Note: Percentages are based on the number of scans read.
Abbreviations: AD = Alzheimer's disease; MCI = mild cognitive impairment; N = number of subjects in the population; n = number of subjects within the category.

Five readers independently interpreted the 205 flortaucipir PET scans collected from the AZES PET substudy. No new subjects were enrolled in this study.

For each efficacy variable, the analysis included subjects who had a valid baseline flortaucipir scan in the Study AZES PET substudy and a change from baseline value of CDR-SB at 18 months.

All subjects had valid and interpretable scans and none were excluded from the analyses.

Outcomes and estimation PX01

The observed risk ratio for a CMD in CDR-SB in Study PX01 was 1.36, which was statistically significant (p=0.0313). The primary result is supported by the secondary analyses, including a statistically significant risk ratio in ADAS-Cog11 18-month change from baseline measure (p=0.0141), as well as trends toward increased risk for τAD++ vs non- τAD++ subjects on the MMSE and FAQ scales. Further, the mean differences between the τAD++ and non-τAD++ groups continued to increase and remained statistically significant for all measures, except FAQ, out to 24 months.

The LS mean changes from the MMRM models also revealed significantly greater deterioration in τAD++ than non- τAD++ subjects at both 18 and 24 months for each of the cognitive and functional outcome variables except FAQ at 24 months (p=.0969).

Primary Efficacy Evaluation

Risk Ratio of τAD++ vs. non-τAD++ Subjects in Progression to Clinically Meaningful Deterioration Evaluated by CDR-SB at 18 Months

The primary objective analysis was to evaluate whether the baseline tau status as determined by flortaucipir F 18 scans would predict the risk of subjects' clinically meaningful cognitive and functional deterioration within 18 months of scan. Clinically meaningful deterioration for the primary objective analysis was defined as CDR-SB CFB with an increase of 1 point or more.

The hypothesis tested was that the risk of progressing to a clinically meaningful event (≥1 point change) in CDR-SB at 18 months would be significantly greater for subjects in the τAD++ group compared to the non-τAD++ group (τAD- and τAD+).

The risk ratio of τAD++ vs. non-τAD++ subjects in progression to CMD is shown in Table 36. The analysis showed a risk ratio of 1.36 for subjects in the τAD++ group as compared to those in the non-τAD++ group (τAD- and τAD+) in terms of progression to CMD within 18 months. The predefined hypothesis was confirmed with an associated p=.0313.

Table 36: Risk ratio of τAD++ vs. Non-τAD++ CDR-SB clinically meaningful deterioration (≥1) at 18 months

Tau Read	CMD (⩾1 CFB) n (%)	No CMD (<1 CFB) n (%)	Total N
τAD++	119 (73.5)	43 (26.5)	162
τAD+	10 (66.7)	5 (33.3)	15
τAD-	14 (50.0)	14 (50.0)	28
Risk Ratio (95% CI)			
τAD++ vs. non-τAD++	**1.36 (1.028, 1.785)**	–	–
p-Value	**0.0313**	–	–

Abbreviations: AD = Alzheimer's disease; CDR-SB = Clinical Dementia Rating Scale Sum of Boxes; CFB = change from baseline; CI = confidence interval; CMD = clinically meaningful deterioration; N = number of subjects in the population; n = number of subjects within the category.

Secondary Efficacy Evaluations

Risk Ratio of τAD++ vs. Non-τAD++ Subjects in Progression to Clinically Meaningful Deterioration Evaluated by MMSE, ADAS-Cog11, FAQ, and CDR Global at 18 Months (Table 37)

The same analysis used for the primary variable was used to test of the risk of progressing to a clinically meaningful event in MMSE, ADAS-Cog11, FAQ, and CDR global at 18 months.

The risk of progressing to CMD at 18 months was numerically greater for subjects in the τAD++ group as compared to those in the non-τAD++ group (τAD- and τAD+) for all the assessed measurements. The pre-specified success criteria were met in Study PX01. An Advanced AD flortaucipir PET pattern was associated with an increased risk of at least a 1-point worsening on CDR-SB (risk ratio = 1.36, p=.0313). There was also a statistically significant increase in risk of at least a 4-point worsening in ADAS (risk ratio = 1.771, p =.0141).

Table 37: Risk ratios of τAD++ vs. Non-τAD++ subjects in progression to clinically meaningful deterioration by MMSE, ADAS-Cog11, FAQ, and CDR global at 18 months

Score	Risk Ratio (95% CI)	p-Value
CDR-SB (⩾1 point change)	1.36 (1.028, 1.785)	0.0313
MMSE ⩽-3 point change	1.35 (0.962, 1.886)	**.0833**
ADAS-Cog$_{11}$ ⩾4 point change	1.77 (1.122, 2.796)	.0141
FAQ ⩾3 point change	1.32 (0.984, 1.776)	**.0639**
CDR global >0 point change	1.28 (0.815, 2.020)	**.2814**

The primary analysis showed a risk ratio of 1.36 for subjects in the τAD++ group as compared to those in the non-τAD++ group (τAD- and τAD+) in terms of progression to clinically meaningful deterioration (CMD) within 18 months. However, the secondary analysis and for the risk of progression to CMD at 18 months, only ADAS-Cog11 change ≥4 reached statistical significance. For the other endpoints/indices MMSE ≤-3 point change, FAQ ≥3 point change and CDR global >0-point change the values did not achieve statistical significance. Thus, the results do not show consistency across all endpoints.

Mean Change of Cognitive/Functional Assessments at 18 Months by Flortaucipir PET Visual Interpretation (τAD++ vs. non-τAD++) (Table 38)

The relationship between visual interpretation (τAD++ vs. non-τAD++ pattern) of the baseline flortaucipir PET scan and the magnitude of cognitive and functional deterioration within 18 months of scan, as measured by the mean CFB of CDR-SB, MMSE, ADAS-Cog11, and FAQ was assessed.

Table 38: Mean Change in Cognitive/Functional Assessments at 18 Months τAD++ vs. Non-τAD++ MMRM Models

	CDR-SB		**MMSE**		**ADAS-Cog$_{11}$**		**FAQ**	
τAD++ vs Non-τAD++								
	τAD++	non-τAD++	τAD++	non-τAD++	τAD++	non-τAD++	τAD++	non-τAD++
LSM ± SE	2.22 (0.215)	1.31 (0.379)	-4.89 (0.377)	-2.12 (0.647)	6.53 (0.660)	1.97 (1.181)	5.22 (0.537)	2.67 (0.895)
p-Value	.0305		.0001		.0006		.0097	

Abbreviations: AD = Alzheimer's disease; ADAS-Cog11 = 11-item version of the Alzheimer's Disease Assessment Scale—Cognitive subscale; CDR-SB = Clinical Dementia Rating Scale Sum of Boxes; FAQ = Pfeffer Functional Activities Questionnaire; LSM = least-squares mean; MMRM = mixed-model repeated measures; MMSE= Mini-Mental State Examination; SE = standard error.

At 18 months, all variables were significantly worse in the τAD++ group compared to the non-τAD++ group. The mean differences between the τAD++ and non-τAD++ groups continued to increase and remained statistically significant for all measures, except FAQ.

Inter-reader Reliability

Inter-reader reliability for prognosis interpretation (τAD++ vs. non-τAD++) was a secondary endpoint in Study PX01. The inter-reader reliability of flortaucipir scan interpretation (τAD++ vs. non-τAD++) across the 5 independent readers was assessed using Fleiss Kappa statistics. Fleiss' kappa was comparable to that for other studies in the development program. Percent agreement among reader pairs was greater than 91% for all pairs. The overall Fleiss Kappa was 0.754 (95% CI: 0.711, 0.797), indicating substantial agreement across all readers. The overall percent of agreement, Fleiss Kappa, and 95% CI around the kappa was summarised, and pairwise comparisons between readers are presented with simple kappa statistics evaluating agreement in the PX01 Clinical Study Report (please see below).

The following Table 39 summarises reliability statistics among each pair of readers and overall.

Reliability among reader pairs was greater than 91% for all pairs. The overall Fleiss Kappa was 0.754 (95% CI: 0.711, 0.797), indicating substantial agreement across all readers.

Table 39 Inter-reader Reliability

	Percent Agreement
τAD++ vs. Non-τAD++	91.4% (1874/2050)
	Fleiss Kappa (95% CI)
τAD++ vs. Non-τAD++	0.754 (0.711, 0.797)

Abbreviation: AD = Alzheimer's disease, CI = confidence interval.

Table 40: Inter-reader Consistency in study PX01

	Kappa (95% CI: tAD++ vs. Non-tAD++)	**Kappa (95% CI: tAD++ vs. Non-tAD++)**	**Kappa (95% CI: tAD++ vs. Non-tAD++)**	**Kappa (95% CI: tAD++ vs. Non-tAD++)**
	For Reader 1	For Reader 2	For Reader 3	For Reader 4
Reader 2	0.820 (0.729, 0.910)			
Reader 3	0.746 (0.634, 0.858)	0.698 (**0.582**, 0.815)		
Reader 4	0.806 (0.711, 0.900)	0.709 (**0.598**, 0.820)	0.681 (**0.561**, 0.801)	
Reader 5	0.789 (0.687, 0.890)	0.768 (0.665, 0.872)	0.868 (0.780, 0.957)	0.670 (**0.550**, 0.791)

For the overall percent of agreement, Fleiss Kappa was just above 0.75 i.e. at the minimum acceptable level for excellent agreement. However, upon looking at the lower limit of 95% CI, it appears that not all the values were higher than 0.60 for all the readers. This does not provide reassurance for the robustness of the results. It is further important to highlight that 4 of the 5 readers in this study had previously participated in A05C study and therefore they were trained on the read method as part of Study A05C. In this context, it is unclear if these readers could be classified as independent readers for this study.

Exploratory Analyses

Risk Ratio of τAD++ vs. Non-τAD++ Subjects in Progression to Clinically Meaningful Deterioration Evaluated by CDR-SB, MMSE, ADAS-Cog11, FAQ, and CDR Global at 24 Months

This analysis tested the hypothesis that the risk of progressing to a clinically meaningful event as determined by value change at 24 months would be significantly greater for subjects in the τAD++ group as compared to those in the non-τAD++ group (τAD- and τAD+). Assessments of whether a visual interpretation (τAD++ vs. non-τAD++ pattern) of the baseline flortaucipir PET scan can predict the risk of subjects' clinically meaningful cognitive and functional deterioration within 24 months of scan, as measured by the CDR, MMSE, ADAS-Cog11, and FAQ CFB are shown in the Table below.

The 24-month results reflect the continued deterioration of the subjects in the study, as well as the effects of reduced sample size due to early termination of the trial. At 24 months postscan, the risk ratio for the CDR-SB was no longer significant due to an increased proportion of subjects meeting the threshold level of deterioration in both the τAD++ and non-τAD++ groups.

Table 41: Study PX01, risk ratios by flortaucipir PET classification τAD++ vs non-τAD++, results at 18 months

	18 Months		**24 Months**	
Score	**RR (95% CI)**	**p-Value**	**RR (95% CI)**	**p-Value**
CDR-SB ⩾1 point change	1.36 (1.028, 1.785)	.0313	1.26 (0.955, 1.653)	**.1029**
MMSE ⩽-3 point change	1.35 (0.962, 1.886)	.0833	1.83 (1.138, 2.942)	.0127
ADAS-Cog$_{11}$ ⩾4 point change	1.77 (1.122, 2.796)	.0141	1.64 (1.075, 2.510)	.0219
FAQ ⩾3 point change	1.32 (0.984, 1.776)	.0639	1.30 (0.938, 1.792	**.1161**
CDR global >0 point change	1.28 (0.815, 2.020)	.2814	1.24 (0.749, 2.042)	**.4071**

Abbreviations: ADAS = Alzheimer's Disease Assessment Scale; CDR-SB = Clinical Dementia Rating Scale Sum of Boxes; FAQ = Functional Activities Questionnaire; MMSE = Mini-Mental State Examination; PET = positron emission tomography; RR = risk ratio.

Risk Ratio of τAD (τAD+ and τAD++) vs. τAD- Subjects in Progression to Clinically Meaningful Deterioration Evaluated by CDR-SB, MMSE, ADAS-Cog11, FAQ, and CDR Global at 18 Months and 24 Months

These analyses tested the hypothesis that the risk of progressing to a clinically meaningful event as determined by value change in by CDR-SB, MMSE, ADAS-Cog11, FAQ, and CDR global at 18 months and 24 months would be significantly greater for subjects in the τAD (τAD+ and τAD++) group as compared to the τAD- groups

Table 42: Risk ratio of subject progression to clinically meaningful deterioration at 18 months and 24 months: τAD++ vs. non-τAD++ (τAD- and τAD+) and τAD (τAD+ and τAD++) vs. τAD-

	τAD++ vs. Non-τAD++ (τAD- and τAD+)		**τAD (τAD+ and τAD++) vs. τAD-**	
Score	**RR (95% CI)**	**p-Value**	**RR (95% CI)**	**p-Value**
18 Months				
CDR-SB ⩾1 point change	1.36 (1.028, 1.785)	**.0313**	1.52 (1.050, 2.194)	**.0266**
MMSE ⩽-3 point change	1.35 (0.962, 1.886)	.0833	1.18 (0.800, 1.743)	.4031
ADAS-Cog$_{11}$ ⩾4 point change	1.77 (1.122, 2.796)	**.0141**	2.02 (1.108, 3.679)	**.0218**
FAQ ⩾3 point change	1.32 (0.984, 1.776)	.0639	1.43 (0.975, 2.091)	.0674
CDR global >0 point change	1.28 (0.815, 2.020)	.2814	1.50 (0.835, 2.691)	.1748
24 Months				
CDR-SB ⩾1point change	1.26 (0.955, 1.653)	.1029	1.27 (0.912, 1.778)	.1552
MMSE ⩽-3 point change	1.83 (1.138, 2.942)	**.0127**	1.58 (0.886, 2.812)	.1211
ADAS-Cog$_{11}$ ⩾4 point change	1.64 (1.075, 2.510)	**.0219**	1.41 (0.882, 2.250)	.1516
FAQ ⩾3 point change	1.30 (0.938, 1.792)	.1161	1.17 (0.823, 1.660)	.3839
CDR Global >0 point change	1.24 (0.749, 2.042)	.4071	1.35 (0.704, 2.585)	.3669

Abbreviations: AD = Alzheimer's disease; ADAS-Cog11 = 11-item version of the Alzheimer's Disease Assessment Scale—Cognitive subscale; CDR = Clinical Dementia Rating Scale; CDR-SB = Clinical Dementia Rating Scale Sum of Boxes; CI = confidence interval; FAQ = Pfeffer Functional Activities Questionnaire; MMSE= Mini-Mental State Examination; RR = risk ratio.

According to the applicant, consistent with the hypothesis, risk ratios for the τAD vs τAD- comparison at 18 months were generally numerically greater than 1, and similar to those for the primary analysis (τAD++ vs non-τAD++), with similar levels of significance. The applicant's review of the cognitive deterioration across the three levels (τAD++, τAD+, τAD-) of flortaucipir majority interpretation showed that for most of the cognitive/functional outcome variables, the percentage of τAD+ subjects meeting the threshold for clinically meaningful deterioration was intermediate between the τAD++ and τAD- groups. However, the number of subjects classified as τAD+ was too small (14 or 15 depending on the analysis) to allow any substantive conclusions about progression in these subjects. In contrast to results of the 24-month follow-up analyses for the primary comparison (τAD++ vs non-τAD++), the 24-month follow-up risk ratios for the τAD vs τAD- comparison were somewhat smaller and failed to achieve significance. Perhaps this result is a consequence of continued accumulation of tau and cognitive deterioration in this purely amyloid positive population over this longer period, and perhaps in part because of the further reduction in sample size.

It should be kept in mind that the population of study PX01 was a selected enriched population. Furthermore, although the parent Study AZES was designed to follow patients for 24 months, it was terminated early due to lack of efficacy in the treatment group. Thus, the number of patients available for evaluation at 24 months was approximately half that available at 18 months.

During the exploratory analyses for the Risk Ratio of τAD++ vs. Non-τAD++ subjects in progression to Clinically Meaningful Deterioration evaluation by CDR-SB, MMSE, ADAS-Cog11, FAQ, and CDR Global at 18 and 24 Months, only CDR-SB and ADAS were statistically significant at 18 months. However, the lower limit of the confidence intervals were close to 1, which does not provide reassurance for the robustness of the results. Furthermore, MMSE ≤-3 point change, FAQ ≥3 point change and CDR global >0 point change did not achieve statistical significance for subjects in the τAD++ group as compared to those in the non-τAD++ group (τAD- and τAD+), as in the case of the above-mentioned secondary analysis. According to the applicant, comparable trends were observed at 24 months, suggesting an increased risk of deterioration in subjects with Advanced (τAD++) flortaucipir PET patterns. However, this cannot be agreed upon, since at 24 months the values of CDR-SB, FAQ and CDRE Global the endpoints did not reach statistical significance and the lower limit of the 95% CI was below 1. In the case of MMSE and ADAS-Cog11 the lower limit was marginally above 1.

In the exploratory analysis of the Risk Ratio of τAD (τAD+ and τAD++) vs. τAD- subjects in progression to Clinically Meaningful Deterioration which was evaluated by CDR-SB, MMSE, ADAS-Cog11, FAQ, and CDR Global at 18 Months and 24 Months the results for the τAD vs τAD- comparison at 18 months were generally numerically greater than 1. However, no statistical significance was achieved (only for CDR-SB and ADAS-Cog11) and the lower limit of the 95% CI ranged from 0.8 to 1.1 depending on the endpoint/index chosen. As the applicant stated, the 24-month follow-up risk ratios for the τAD vs τAD- comparison were somewhat smaller and failed to achieve significance. It appears that there is no continuity in the results from 18 to 24 months, which does not provide reassurance for the "predictive abilities" of flortaucipir.

Mixed Model Repeated Measures for Comparison of LSM Changes at 18 Months for:

- τAD (τAD+ and τAD++) vs. τAD-
- τAD++ vs. τAD

The mean change for τAD++ and non-τAD++ for CDR-SB, MMSE, FAQ and ADAS-Cog11. As was true for Study A05C, for each variable in Study PX01, mean worsening from baseline was significantly greater for τAD++ than non-τAD++ subjects.

Diagnostic Performance of Flortaucipir Scan in Predicting Clinical Deterioration – Sensitivity and Specificity

The diagnostic performance of baseline tau status in predicting Clinically Meaningful Deterioration (CMD) at 18 months was evaluated by CDR-SB, MMSE, ADAS-Cog11, FAQ, and CDR Global.

This analysis used dichotomised change at 18 months as a truth standard to assess the diagnostic performance of baseline tau status as determined by flortaucipir scan (Table 43, below). Sensitivity, specificity, accuracy, PPV, and NPV, along with their respective 95% Wilson score confidence intervals are presented in Table below. Diagnostic performance (sensitivity/specificity) was assessed by using CMD as truth standard (TS). There was a higher proportion of CMD among τAD++ than non-τAD++ in all endpoint measurements. Sensitivity was high, with τAD++ detected more than 80% of subjects with clinically meaningful deterioration. However, specificity was very low.

Table 43: Diagnostic performance of baseline Tau status in predicting clinical meaningful change at 18 months

CDR-SB			
Tau read	CMD (⩾1 CFB)	No CMD (<1 CFB)	Total
τAD++	119	43	162
Non-τAD++	24	19	43
Total	143	62	205
Sensitivity (%)	83.2 (76.24, 88.45)		
Specificity (%)	**30.6** (20.58, 42.97)		
MMSE			
Tau read	CMD (⩽-3 CFB)	No CMD (>-3 CFB)	Total
τAD++	112	49	161
Non-τAD++	21	21	42
Total	133	70	203
Sensitivity (%)	84.2 (77.07, 89.43)		
Specificity (%)	**30.0** (20.54, 41.54)		
ADAS-Cog$_{11}$			
Tau read	CMD (⩾4 CFB)	No CMD (<4 CFB)	Total
τAD++	98	59	157
Non-τAD++	14	28	42
Total	112	87	199
Sensitivity (%)	87.5 (80.11, 92.41)		
Specificity (%)	**32.2** (23.30, 42.57)		
FAQ			
Tau read	CMD (⩾3 CFB)	No CMD (<3 CFB)	Total
τAD++	111	47	158
non-τAD++	23	19	42
Total	134	66	200
Sensitivity (%)	82.8 (75.56, 88.28)		
Specificity (%)	**28.8** (19.27, 40.64)		
CDR Global			
Tau Read	CMD (>0 CFB)	No CMD (0 CFB)	Total
τAD++	71	91	162
non-τAD++	16	27	43
Total	87	118	205
Sensitivity (%)	81.6 (72.19, 88.35)		
Specificity (%)	**22.9** (16.23, 31.24)		

Diagnostic Performance of Baseline Tau Status in Predicting Clinical Meaningful Change at 18 Months Abbreviations: AD = Alzheimer's disease; ADAS-Cog11 = 11-item version of the Alzheimer's Disease Assessment Scale—Cognitive subscale; CDR = Clinical Dementia Rating Scale; CDR-SB = Clinical Dementia Rating Scale Sum of Boxes; CFB = change from baseline; CMD = clinically meaningful deterioration; FAQ = Pfeffer Functional Activities Questionnaire; MMSE = Mini-Mental State Examination.

With respect to the diagnostic performance of flortaucipir scan in predicting clinical deterioration Sensitivity was moderately high (81.6% to 87.5%), with τAD++ detected more than 80% of subjects with clinically meaningful deterioration. However, specificity was very low ranging from 22.9% to 32.2%. Similarly, to the case of study A05C, the assessment of sensitivity/specificity for flortaucipir PET positivity to pre-specific thresholds for cognitive and functional decline cannot be considered successful.

It is questionable whether a comparison performed post-hoc using a dichotomisation approach without a control group and showing inconsistent results across endpoints using various analyses with very low

specificity could provide useful robust information. However, the applicant did not intend in the end to make claims regarding the diagnostic performance of flortaucipir for prediction of progression for individual patients.

Study PX01 – Subgroup analyses

The primary and secondary analyses, with the exception of inter-reader consistency, were repeated by age group (≤75 years; > 75 years), sex (female; male), and baseline CDR global score (≥1; < 1). The analysis in the SAP was not run by the CDR global subgroups, as this model was already adjusted for baseline CDR global score.

In subgroup analyses few significant differences in risk ratio at 18 months for τAD++ vs. non-τAD++ were seen for age, gender, and CDR global score. However, it can be agreed with the applicant that the number of subjects is relatively small in the subgroups and the range of subjects is restricted by cognitive and amyloid status, so interpretation of subgroup analyses should be made with caution. The effect of tau status on these subgroups would be better understood in larger, less homogenous populations.

Ancillary analyses PX01

To avoid misinterpretation of potentially biased estimates caused by missing data, the MMRM analyses were run on the set of 90 subjects who completed the 24-month visit (Table 44 below). The results of the sensitivity analyses were similar to those reported in the PX01 Clinical Study Report from the combined 18- and 24-month MMRM analysis though numerically different likely due to differences in sample and sample size.

Table 44 Sensitivity analysis of mixed model repeated measures at 18 months for comparison of τAD++ vs. Non-τAD++ change in CDR-SB, MMSE, ADAS-Cog11, and FAQ in subjects who completed 24 months

	CDR-SB		**MMSE**		**ADAS-Cog$_{11}$**		**FAQ**	
	τAD++	non-τAD++	τAD++	non-τAD++	τAD++	non-τAD++	τAD++	non-τAD++
LSM ± SE	2.10 (0.285)	1.28 (0.490)	-4.88 (0.578)	-2.17 (0.958)	5.95 (0.969)	1.56 (1.709)	4.93 (0.764)	2.36 (1.205)
p-Value	.1403		**.0131**		**.0242**		.0601	

Abbreviations: AD = Alzheimer's disease; ADAS-Cog11 = 11-item version of the Alzheimer's Disease Assessment Scale—Cognitive subscale; CDR-SB = Clinical Dementia Rating Scale Sum of Boxes; FAQ = Pfeffer Functional Activities Questionnaire; LSM = least-squares mean; MMRM = mixed-model repeated measures; MMSE = Mini-Mental State Examination; SE = standard error.

According to the applicant, in subgroup analyses few significant differences in risk ratio at 18 months for τAD++ vs. non-τAD++ were seen for age, gender, and CDR global score. However, the number of subjects is relatively small in the subgroups and the range of subjects is restricted by cognitive and amyloid status, so interpretation of subgroup analyses should be made with caution. The effect of tau status on these subgroups would be better understood in larger, less homogenous populations.

With this sensitivity analysis, statistically significant results could not be obtained with CDR-SB and FAQ, which did not provide reassurance for the robustness of the outcome. As in the secondary and exploratory analyses the results did not show consistency across the endpoints selected, which is of concern.

- **Summary of main efficacy results**

The following tables summarise the efficacy results from the main studies supporting the present

application. These summaries should be read in conjunction with the discussion on clinical efficacy as well as the benefit risk assessment (see later sections).

Indication I: Neuropathologic Correspondence for diagnostic performance and reader agreement Studies

Table 45: Summary of efficacy for pivotal study 18F-AV-1451-A16 (Study A16)

Title: A Clinico-Pathological Study of the Correspondence between ^{18}F-AV-1451 PET Imaging and Post-Mortem Assessment of Tau Pathology			
Study identifier	Protocol No.: 18F-AV-1451-A16		
Design	Phase 3, open-label study examining the correspondence between ante-mortem imaging with flortaucipir and post-mortem tau pathology in terminally ill subjects with AD or mild cognitive impairment and terminally ill subjects who were cognitively normal. Images were assessed as having an AD tau pattern of flortaucipir retention (moderate/τAD+ [within the posterior lateral temporal/occipital regions] or advanced/τAD++ [beyond the posterior lateral temporal/occipital regions]) or no flortaucipir retention/retention in a non-AD pattern (τAD-)		
	Duration of main phase:		Single PET scan with follow-up to autopsy; cognitively impaired subjects remaining alive at 9 months could optionally receive a second scan with an additional 9-month follow-up phase; cognitively normal subjects were followed up to autopsy regardless of duration
Hypothesis	For both Primary Analysis 1 and Primary Analysis 2, for the same 3 out of 5 readers, the lower bound of the 95% CIs for both sensitivity and the specificity will be ≥50%.		
Treatment group	Flortaucipir		370 MBq (10 mCi) as an IV bolus administration
Endpoints and definitions	Co-primary endpoint 1	Diagnostic Performance of Individual Readers (NFT Score)	Sensitivity and specificity of 5 independent readers' interpretations of ante-mortem flortaucipir PET imaging of an AD tau pattern of flortaucipir neocortical uptake that corresponds to NFT Score of B3 (Braak V/VI) as measured at autopsy
	Co-primary endpoint 2	Diagnostic Performance of Individual Readers (High Levels of ADNC)	Sensitivity and specificity of 5 independent readers' interpretations of ante-mortem flortaucipir imaging of an AD tau pattern of flortaucipir neocortical uptake that corresponds to high levels of ADNC as defined by NIA-AA criteria as measured at autopsy
	Secondary endpoint	Diagnostic Performance Majority Read (NFT Score)	Sensitivity and specificity of majority of 5 readers' interpretation of AD pattern tau PET scan corresponding to NFT Score of B3 (Braak V/VI) as measured at autopsy
	Secondary endpoint	Diagnostic Performance Majority Read (High Levels of ADNC)	Sensitivity and specificity of majority of 5 readers' interpretation of AD pattern tau PET scan corresponding to high levels of ADNC as defined by NIA-AA criteria as measured at autopsy
	Secondary endpoint	Inter-reader Agreement	Fleiss' Kappa statistics were used to assess inter-reader agreement for the diagnostic decisions associated with primary efficacy analysis 1 and 2
Database lock	20 August 2018		

Results and Analysis		
Analysis description	**Co-primary Analysis 1:** Diagnostic Performance of Individual Readers (NFT Score)	
Analysis population, time point description, and statistical model	Efficacy Analysis Set 1 At autopsy within 9 months of baseline/last scan The sensitivity and specificity along with their 2-sided 95% CIs (based on the Wilson score method) were calculated for each of the 5 readers. The first primary efficacy analysis was considered to be met if for at least the same 3 of 5 readers, the lower bounds of the 95% CI for both sensitivity and specificity were ≥50% (that is, statistically significant at 2-sided significance level of 0.05).	
Descriptive statistics and estimate variability	Treatment group	Flortaucipir
	Number of patients	64
	Sensitivity: % median (range) of 5 readers Range % (95% CI) across 5 individual readers	92.3 (92.3-100.0) 5 of 5 readers met success criteria 92.3 (79.7, 97.3) to 100 (91.0, 100.0)
	Specificity: % median (range) of 5 readers Range % (95% CI) across 5 individual readers	76.0 (52.0-92.0) 3 of 5 readers met success criteria 52.0 (33.5, 70.0) to 92.0 (75.0, 97.8)
Analysis description	**Co-primary Analysis 2:** Diagnostic Performance of Individual Readers (High Levels of ADNC)	
Analysis population, time point description, and statistical model	Efficacy Analysis Set 2 At autopsy within 9 months of baseline/last scan The sensitivity and specificity along with their 2-sided 95% CIs (based on the Wilson score method) were calculated for each of the 5 readers. The second primary efficacy analysis was considered to be met if for at least the same 3 out of 5 readers, the lower bounds of the 95% CI for both sensitivity and specificity were ≥50% (that is, statistically significant at 2-sided significance level of 0.05).	
Descriptive statistics and estimate variability	Treatment group	Flortaucipir
	Number of patients	64
	Sensitivity: % median (range) of 5 readers Range % (95% CI) across 5 individual readers	94.7 (94.7-100.0) 5 of 5 readers met success criteria 94.7 (82.7, 98.5) to 100 (90.8, 100.0)
	Specificity: % median (range) of 5 readers Range % (95% CI) across 5 individual readers	76.9 (50.0-92.3) 3 of 5 readers met success criteria 50 (32.1, 67.9) to 92.3 (75.9, 97.9)
Analysis description	**Secondary Analysis:** Diagnostic Performance Majority Read (NFT Score)	
Analysis population, time point description, and statistical model	Efficacy Analysis Set 1 At autopsy within 9 months of baseline/last scan The sensitivity and specificity along with their 2-sided 95% CIs (based on the Wilson score method) were calculated for the majority interpretation of 5 independent readers relative to the appropriate truth standard.	
Descriptive statistics and estimate variability	Treatment group	Flortaucipir
	Number of patients	64
	Sensitivity (95% CI)	92.3% (79.7, 97.3)
	Specificity (95% CI)	80.0% (60.9, 91.1)

Analysis description	**Secondary Analysis:** Diagnostic Performance Majority Read (high levels of ADNC)	
Analysis population, time point description, and statistical model	Efficacy Analysis Set 2 At autopsy within 9 months of baseline/last scan The sensitivity and specificity along with their 2-sided 95% CIs (based on the Wilson score method) were calculated for the majority interpretation of 5 independent readers relative to the appropriate truth standard.	
Descriptive statistics and estimate variability	Treatment group	Flortaucipir
	Number of patients	64
	Sensitivity (95% CI)	94.7% (82.7, 98.5)
	Specificity (95% CI)	80.8% (62.1, 91.5)
Analysis description	**Secondary Analysis:** Inter-reader Agreement	
Analysis population, time point description, and statistical model	Efficacy Analysis Set 1 and Efficacy Analysis Set 2 All participants from the study for which a visual read flortaucipir PET result was obtained, regardless of whether the subject had a valid autopsy. Fleiss' kappa statistics were used to assess inter-reader agreement for the diagnostic decision associated with Primary Analysis 1 and 2. For evaluation of agreement in the context of Primary Analysis 1, Fleiss' kappa statistic was calculated with its 2-sided 95% CI and associated p-value.	
Descriptive statistics and estimate variability	Treatment group	Flortaucipir
	Number of patients	105
	Overall percentage agreement	89.9%
	Fleiss' kappa (95% CI)	0.80 (0.74, 0.86)
	p-value	<0.001

Abbreviations: AD = Alzheimer's disease; ADNC = Alzheimer's disease neuropathologic change; CI = confidence interval; IV = intravenous; NFT = neurofibrillary tangle; NIA-AA = National Institute on Aging-Alzheimer's Association; PET = positron emission tomography.

Table 46: Summary of efficacy for pivotal study 18F-AV-1451-FR01 (Study FR01)

Title: A reader study to assess accuracy and reliability of flortaucipir F 18 PET scan interpretation		
Study identifier	Protocol No.: 18F-AV-1451-FR01	
Design	Study to evaluate the performance of physician readers trained to read flortaucipir-PET scans. No new subjects were recruited and no drug was administered in this study. Images used were collected in Studies 18F-AV-1451-A16 and 18F-AV-1451-A05. Images were assessed for AD tau pattern as described in Study A16.	
	Duration of main phase:	Not applicable
Hypothesis	For both Primary Analysis 1 and Primary Analysis 2, for at least the same 3 out of 5 readers, the lower bound of the 95% CIs for both sensitivity and the specificity will be ≥50%.	
Treatment groups	Flortaucipir (administered in parent studies)	370 MBq (10 mCi) as an IV bolus administration (administered in parent studies)

Endpoints and definitions	Co-primary endpoint 1: Analysis 1	Diagnostic Performance of Individual Readers (NFT Score)	Sensitivity and specificity of 5 independent readers' interpretations of ante-mortem flortaucipir PET imaging for detection of a τAD (τAD+/ τAD++) pattern of flortaucipir neocortical uptake that corresponds to NFT Score of B3 (Braak V/VI) as measured at autopsy
	Co-primary endpoint 1: Analysis 2	Diagnostic Performance of Individual Readers (High Levels of ADNC)	Sensitivity and specificity of 5 independent readers' interpretations of ante-mortem flortaucipir imaging for detection of a τAD (τAD+/ τAD++) pattern of flortaucipir neocortical uptake that corresponds to high levels of ADNC as defined by NIA-AA criteria as measured at autopsy
	Co-primary endpoint 2	Inter-reader Reliability/Agreement	Fleiss' Kappa statistics were used to assess inter-reader agreement across all readers of flortaucipir PET scan interpretation as τAD (τAD+/ τAD++) vs. non-AD pattern (τAD-)
	Secondary endpoint	Diagnostic Performance of τAD++ Flortaucipir PET Images to Detect B3 (Braak V/VI) NFTs and high ADNC	Relationship between ante-mortem flortaucipir PET imaging of τAD++ and NFT Score of B3 (Braak V/VI), as measured at autopsy
	Secondary endpoint	Inter-reader Reliability of Reader Interpretation of τAD++ Flortaucipir PET Images	Inter-reader reliability of flortaucipir PET scans with a τAD++ pattern across the 5 readers
	Secondary endpoint	Inter-reader Reliability of Flortaucipir PET Scan Interpretation in the Population of Intended Use	Inter-reader agreement of flortaucipir PET scan interpreted as τAD (τAD+/ τAD++) vs. non-AD pattern (τAD-) for the non-autopsy intended clinical use population (scans from Study A05)
	Secondary endpoint	Intra-reader Reliability of Reader Interpretation of Flortaucipir-PET Imaging	Intra-reader agreement of flortaucipir PET scan visual interpretation, which was interpreted as τAD (τAD+/ τAD++) vs. non-AD pattern (τAD-) on scans randomly read twice by each reader
Database lock	03 May 2019		
Results and Analysis			
Analysis description	**Co-primary Endpoint 1 Analysis 1:** Diagnostic Performance of Individual Readers (B3 [Braak V/VI] **NFT Score**)		
Analysis population, time point description, and statistical model	Efficacy Analysis Population for Accuracy Evaluation Valid scan and autopsy in Study A16 (plus supplemental autopsy dataset and frontrunners) The sensitivity and specificity along with their 2-sided 95% CIs (based on the Wilson score method) were calculated for each of the 5 readers. The Primary Endpoint 1 Analysis 1 was considered to be met if for at least the same 3 of 5 readers, the lower bounds of the 95% CI for both sensitivity and specificity were ≥50% (that is, statistically significant at 2-sided significance level of 0.05).		
Descriptive statistics and	Treatment group		Flortaucipir
	Number of scans		82

estimate variability	Sensitivity: % median (range) of 5 readers Range % (95% CI) across 5 individual readers	89.4 (87.2-93.6) 5 of 5 readers met success criteria 87.2 (74.8, 94.0) to 93.6 (82.8, 97.8)
	Specificity: % median (range) of 5 readers Range % (95% CI) across 5 individual readers	77.1 (62.9-91.4) 4 of 5 readers met success criteria 62.9 (46.3, 76.8) to 91.4 (77.6, 97.0)
Analysis description	**Co-primary Endpoint 1 Analysis 2:** Diagnostic Performance of Individual Readers (High Levels of ADNC)	
Analysis population, time point description, and statistical model	Efficacy Analysis Population for Accuracy Evaluation Valid scan and autopsy in Study A16 (plus supplemental autopsy dataset and frontrunners) The sensitivity and specificity along with their 2-sided 95% CIs (based on the Wilson score method) were calculated for each of the 5 readers. The Primary Endpoint 1 Analysis 2 was considered to be met if for at least the same 3 of 5 readers, the lower bounds of the 95% CI for both sensitivity and specificity were ≥50% (that is, statistically significant at 2-sided significance level of 0.05).	
Descriptive statistics and estimate variability	Treatment group	Flortaucipir
	Number of scans	82
	Sensitivity: % median (range) of 5 readers Range % (95% CI) across 5 individual readers	95.1 (95.1-97.6) 5 of 5 readers met success criteria 95.1 [83.9, 99.0] to 97.6 [87.4, 100.0]
	Specificity: % median (range) of 5 readers Range % (95% CI) across 5 individual readers	73.2 (58.5-85.4) 4 of 5 readers met success criteria 58.5 [43.4, 72.2] to 85.4 [71.6, 93.1]
Analysis description	**Co-Primary Endpoint 2:** Inter-reader Reliability/Agreement	
Analysis population, time point description, and statistical model	Efficacy Analysis Population for Precision Evaluation 1 Valid scan and autopsy in Study A16 or valid baseline scan in Study A05 The inter-reader reliability of flortaucipir PET scan visual interpretation was assessed using Fleiss' kappa. p-values were calculated using the normal approximation method. The lower bound of the 95% CI for Fleiss' kappa was to be greater than or equal to 0.6 to meet the inter-reader reliability criterion.	
Descriptive statistics and estimate variability	Treatment group	Flortaucipir
	Number of scans	241
	Overall percentage agreement	93.8%
	Fleiss' kappa (95% CI)	0.87 (0.83, 0.91)
	p-value	<0.0001
Analysis description	**Secondary Endpoint:** Diagnostic Performance of tAD++ Flortaucipir PET Images to Detect B3 (Braak V/VI) NFTs	
Analysis population, time point description, and statistical model	Efficacy Analysis Population for Accuracy Evaluation Valid scan and autopsy in Study A16 (plus supplemental autopsy dataset and frontrunners) The sensitivity and specificity along with their 2-sided 95% CIs (based on the Wilson score method) were calculated for each of the 5 readers. The secondary endpoint was considered to be met if for at least the same 3 of 5 readers, the lower bounds of the 95% CI for both sensitivity and specificity were ≥50% (that is, statistically significant at 2-sided significance level of 0.05).	

Descriptive statistics and estimate variability	Treatment group	Flortaucipir
	Number of scans	82
	Sensitivity: % median (range) of 5 readers	83.0 (76.6-89.4) 5 of 5 readers met success criteria
	Specificity: % median (range) of 5 readers	94.3 (82.9-100.0) 5 of 5 readers met success criteria
Analysis description	**Secondary Endpoint:** Diagnostic Performance of τAD++ Flortaucipir PET Images to Detect high ADNC	
Analysis population, time point description, and statistical model	Efficacy Analysis Population for Accuracy Evaluation Valid scan and autopsy in Study A16 (plus supplemental autopsy dataset and frontrunners) The sensitivity and specificity along with their 2-sided 95% CIs (based on the Wilson score method) were calculated for each of the 5 readers. The secondary endpoint was considered to be met if for at least the same 3 of 5 readers, the lower bounds of the 95% CI for both sensitivity and specificity were ≥50% (that is, statistically significant at 2-sided significance level of 0.05).	
Descriptive statistics and estimate variability	Treatment group	Flortaucipir
	Number of scans	82
	Sensitivity: % median (range)	87.8 (85.4-92.7)
	Specificity: % median (range)	87.8 (75.6-92.7)
Analysis description	**Secondary Endpoint:** Inter-reader Reliability of Reader Interpretation of τAD++ Flortaucipir PET Images	
Analysis population, time point description, and statistical model	Efficacy Analysis Population for Precision Evaluation 1 Valid scan and autopsy in Study A16 or valid baseline scan in Study A05 The inter-reader reliability of flortaucipir PET scan visual interpretation was assessed using Fleiss' kappa. p-values were calculated using the normal approximation method. The lower bound of the 95% CI for Fleiss' kappa was to be greater than or equal to 0.6 to meet the inter-reader reliability criterion.	
Descriptive statistics and estimate variability	Treatment group	Flortaucipir
	Number of scans	241
	Overall percentage agreement	92.0%
	Fleiss' kappa (95% CI)	0.84 (0.80, 0.88)
	p-value	<0.0001
Analysis description	**Secondary Endpoint:** Inter-reader Reliability of Flortaucipir PET Scan Interpretation in the Population of Intended Use	
Analysis population, time point description, and statistical model	Efficacy Analysis Population for Precision Evaluation 2 (Study A05 scans) Valid baseline scan in Study A05. The inter-reader reliability of flortaucipir PET scan visual interpretation was assessed using Fleiss' kappa. p-values were calculated using the normal approximation method. The lower bound of the 95% CI for Fleiss' kappa was to be greater than or equal to 0.6 to meet the inter-reader reliability criterion.	
Descriptive statistics and estimate variability	Treatment group	Flortaucipir
	Number of scans	159
	Overall percentage agreement	95.1%
	Fleiss' kappa (95% CI)	0.90 (0.85, 0.95)

	p-value			<0.0001	
Analysis description	**Secondary Endpoint:** Intra-reader Reliability of Reader Interpretation of Flortaucipir-PET Imaging				
Analysis population, time point description, and statistical model	Efficacy Analysis Population for Precision Evaluation 3 20 randomly selected cases to be read twice The degree of agreement between the first and second read of the same case in a pairwise manner using Cohen's kappa statistics.				
Descriptive statistics and estimate variability	Treatment group			Flortaucipir	
	Number of scans			20	
		% Intra-reader Agreements	**Cohen's kappa**	**95% CI**	**p-value**
	Reader 1	95.0	0.89	0.69, 1.00	<0.0001
	Reader 2	85.0	0.71	0.41, 1.00	0.0010
	Reader 3	80.0	0.60	0.24, 0.95	0.0077
	Reader 4	95.0	0.89	0.67, 1.00	<0.0001
	Reader 5	90.0	0.79	0.52, 1.00	0.0004

Abbreviations: AD = Alzheimer's disease; ADNC = Alzheimer's disease neuropathologic change; CI = confidence interval; IV = intravenous; NFT = neurofibrillary tangle; NIA-AA = National Institute on Aging-Alzheimer's Association; PET = positron emission tomography.

Indication II: Pivotal trials concerning prognosis performance (this claim is not included in the final wording of the indication)

Table 47: Summary of efficacy for pivotal study 18F-AV-1451-A05C (Study A05C)

Title: An Open Label, Multicenter Study, Evaluating the Safety and Imaging Characteristics of 18F-AV-1451 in Cognitively Healthy Volunteers, Subjects with Mild Cognitive Impairment, and Subjects with Alzheimer's Disease			
Study identifier	Protocol No.: 18F-AV-1451-A05		
Design	Phase 3, open-label, confirmatory phase to confirm the relationship between flortaucipir uptake in the brain as measured by PET in subjects with AD or MCI and the subsequent rate of cognitive decline observed over longitudinal follow up that was observed in the Exploratory Phase (A05E) of the study.		
	Duration of main phase:		18 months
Hypothesis	The primary hypothesis of the confirmatory phase was that the hazard of progressing to a 1 point or more worsening in CDR-SB within 18 months would be significantly greater for subjects with flortaucipir scans rated (by the majority of the 5 readers) as a τAD++ pattern, as compared to those with scans rated as non-τAD++ (τAD- and τAD+).		
Treatment groups	Flortaucipir		370 MBq (10 mCi) as an IV bolus administration
Endpoints and definitions	Primary endpoint	Relationship between Neocortical Flortaucipir Uptake and the Subsequent Rate of Cognitive Decline	To assess whether or not the baseline τAD++ tau pattern, according to a flortaucipir scan visual interpretation, predicts a higher risk of subjects' cognitive and functional deterioration within 18 months of scan, as measured by the CDR-SB change from baseline (defined as a 1 point or greater worsening on CDR-SB)

	Secondary endpoint	Diagnostic Performance of Flortaucipir Visual Read	To assess the diagnostic performance of baseline τAD++ tau pattern according to a flortaucipir scan visual interpretation, for predicting subjects' cognitive and functional deterioration within 18 months of scan, as measured by the CDR-SB scales
Database lock	27 February 2018		
Results and Analysis			
Analysis description	**Primary Analysis:** Relationship between Neocortical Flortaucipir Uptake and the Subsequent Rate of Cognitive Decline		
Analysis population, time point description, and statistical model	All subjects with a valid, interpretable PET image, and who have completed the relevant clinical/cognitive assessment from baseline, and any of post-flortaucipir scan visits, 9-month and/or 18-month visits. Flortaucipir baseline scan visual reads were dichotomised as having a τAD++ (hypothesised to be associated with near term/18 months progression) versus a non-τAD++ (τAD+ or τAD-). The final value for each scan was the majority classification across the 5 independent imaging physicians. For the purposes of the primary analysis, flortaucipir status was compared to cognitive and functional decline, defined as a 1 point worsening in CDR-SB. Time to first occurrence of the event was modelled using a Cox proportional hazard model adjusted for baseline age, ANART, and baseline CDR-SB score.		
Descriptive statistics and estimate variability	Treatment group	Flortaucipir	
	Number of patients	159	
	Cox proportional hazard ratio (confidence limit)	1.581 (0.968, 2.581)	
	p-value	0.067	
Analysis description	**Secondary Analysis:** Diagnostic Performance of Flortaucipir Visual Read		
Analysis population, time point description, and statistical model	All subjects with a valid, interpretable PET image, and who have completed the relevant clinical/cognitive assessment from baseline, and any of post-flortaucipir scan visits, 9-month and/or 18-month visits. Dichotomised CDR-SB change (1 point or more increase vs. otherwise) as a truth standard to assess the diagnostic performance of baseline tau status as determined by flortaucipir scan interpretation. Sensitivity and specificity were calculated for each of the 5 independent imaging readers. The hypothesis tested was that of the 5 independent imaging physicians, at least 3 will have the lower bounds of 2-sided 95% CIs ≥50% for both sensitivity and specificity.		
Descriptive statistics and estimate variability	Treatment group	Flortaucipir	
	Number of patients	110	
	Sensitivity: % median (range)	55.8 (50.0-59.6)	
	Specificity: % median (range)	65.5 (65.5-65.5)	

Abbreviations: AD = Alzheimer's disease; ANART = American National Adult Reading Test; CDR-SB = clinical dementia rating - sum of boxes; IV = intravenous; MCI = mild cognitive impairment; PET = positron emission tomography.

Table 48: Summary of efficacy for pivotal study 18F-AV-1451-PX01 (Study PX01)

Title: Evaluation of the Relationship between Baseline Flortaucipir PET Signal and Cognitive Change in Subjects with Early Alzheimer's Disease Participating in the I8D-MC-AZES Protocol Addendum D5010C00009 (2.1) (Tau Imaging)			
Study identifier	Protocol No.: 18F-AV-1451-PX01		
Design	Study to evaluate the relationship between baseline flortaucipir PET signal and cognitive change in subjects with early AD. No new subjects were recruited and no drug was administered in this study. Images used were collected in a substudy of parent Study I8D-MC-AZES (AZES).		
	Duration of main phase:		Minimum of 18 months post-baseline follow-up (in parent study)
Hypothesis	The primary hypothesis tested was that the risk of progressing to a ≥1-point change in CDR-SB at 18 months would be significantly greater for subjects rated (by the majority of the 5 readers) in the тAD++ group compared to the non-тAD++ group (тAD- and тAD+).		
Treatment groups	Flortaucipir (administered in the parent study)		240 MBq (6.5 mCi) as an IV bolus administration (administered in the parent study)
Endpoints and definitions	Primary endpoint	Risk Ratio for cognitive and functional deterioration on CDR-SB	To assess whether a visual interpretation (тAD++ vs. non-тAD++ pattern) of the baseline flortaucipir PET scan can predict the risk of subjects' cognitive and functional deterioration within 18 months of scan, as measured by the CDR-SB change from baseline
	Secondary endpoint	Risk Ratio for cognitive and functional deterioration on Various Clinical Measures	To assess whether a visual interpretation (тAD++ vs. non-тAD++ pattern) of the baseline flortaucipir PET scan can predict the risk of subjects' cognitive and functional deterioration within 18 months of scan, as measured by MMSE, ADAS-Cog11, FAQ, and CDR global change from baseline
	Secondary endpoint	Mean Change in Cognitive/Functional Assessments	To assess the relationship between visual interpretation (тAD++ vs. non-тAD++ pattern) of the baseline flortaucipir PET scan and magnitude of cognitive and functional deterioration within 18 months of scan, as measured by the mean change from baseline of CDR-SB, MMSE, ADAS-Cog11, and FAQ
	Secondary endpoint	Inter-reader Reliability of Reader Interpretation of Flortaucipir PET Imaging	To assess inter-reader reliability of the flortaucipir PET scan visual interpretation by 5 independent, blinded readers
Database lock	06 May 2019		

Results and Analysis			
Analysis description	**Primary Analysis:** Risk Ratio for Cognitive and Functional Deterioration on CDR-SB		
Analysis population, time point description, and statistical model	All subjects who had a valid baseline flortaucipir scan in the Study AZES PET substudy and a change from baseline value of CDR-SB at 18 months. The primary efficacy variable was the dichotomised CDR-SB score change from baseline (1 point or more increase vs. otherwise). A Poisson regression model was used to calculate the ratio of risk for τAD++ subjects over non-τAD++ (τAD+/τAD-) subjects. Cognitive and functional deterioration defined as ≥1-point change in CDR-SB was used as the dependent variable, and the model was adjusted for baseline age, years of education (categorical), CDR-SB score, and therapeutic treatment assignment from the AZES study (lanabecestat 20 mg or 50 mg, or placebo).		
Descriptive statistics and estimate variability	Treatment group	Flortaucipir	
	Number of patients	205	
	Risk ratio of τAD++ vs. non-τAD++ (95% CI)	1.36 (1.028, 1.785)	
	p-value	0.0313	
Analysis description	**Secondary Analysis:** Risk Ratio for Cognitive and Functional Deterioration on Various Clinical Measures		
Analysis population, time point description, and statistical model	Same as primary analysis with the following dependent variables 1 through 4 for cognitive and functional deterioration, and the adjustment of corresponding baseline scores: 1) MMSE decreased by 3 points or more 2) ADAS-Cog11 increased by 4 points or more 3) FAQ increased by 3 points or more 4) CDR global with any increase.		
Descriptive statistics and estimate variability	Treatment group	Flortaucipir	
	Number of patients	205	
		Risk Ratio (95% CI)	**p-value**
	MMSE ≤ -3-point change	1.35 (0.962, 1.886)	0.0833
	ADAS-Cog11 ≥4-point change	1.77 (1.122, 2.796)	0.0141
	FAQ ≥3-point change	1.32 (0.984, 1.776)	0.0639
	CDR global >0-point change	1.28 (0.815, 2.020)	0.2814
Analysis description	**Secondary Analysis:** Mean Change in Cognitive/Functional Assessments		
Analysis population, time point description, and statistical model	Same analysis population as primary analysis. MMRM was used to assess the mean change of MMSE, ADAS-Cog11, FAQ, and CDR-SB at 18 months by tau status. For each analysis, the change from baseline value from relative measurement was the dependent variable, and the model included the fixed effects of tau status (τAD++ or non-τAD++), visit (categorical covariate), tau status-by-visit interaction, therapeutic treatment assignment from the AZES study (lanabecestat 20 mg or 50 mg, or placebo), and years of education (categorical), as well as corresponding baseline measurement score and age as continuous covariates.		

Descriptive statistics and estimate variability	Treatment group		Flortaucipir	
	Number of patients		205 (162 τAD++; 43 non-τAD++)	
		τAD++ LSM (SE)	**Non-τAD++ LSM (SE)**	**p-value**
	CDR-SB	2.22 (0.215)	1.31 (0.379)	0.0305
	MMSE	-4.89 (0.377)	-2.12 (0.647)	0.0001
	ADAS-Cog11	6.53 (0.660)	1.97 (1.181)	0.0006
	FAQ	5.22 (0.537)	2.67 (0.895)	0.0097
Analysis description	**Secondary Analysis:** Inter-reader Reliability of Reader Interpretation of Flortaucipir PET Imaging			
Analysis population, time point description, and statistical model	Same analysis population as primary analysis. The inter-reader reliability of flortaucipir scan interpretation (τAD++ vs. non-τAD++) across the 5 independent readers was assessed using Fleiss Kappa statistics.			
Descriptive statistics and estimate variability	Treatment group		Flortaucipir	
	Number of patients		205	
	Overall percentage agreement		91.4% (1874/2050)	
	Fleiss' kappa (95% CI)		0.754 (0.711, 0.797)	

Abbreviations: AD = Alzheimer's disease; ADAS-Cog11 = 11-item version of the Alzheimer's Disease Assessment Scale–Cognitive subscale; CDR-SB = clinical dementia rating - sum of boxes; FAQ = Pfeffer Functional Activities Questionnaire; LSM = least-squares mean; MMRM = mixed model with repeated measures; MMSE = Mini-Mental State Examination; PET = positron emission tomography; SE = standard error.

2.6.5.4. Clinical studies in special populations

The target population is represented by elderly patients predominantly. The product is not expected to be used in children. Subgroup analysis conducted in various age groups and gender subgroups in the study A16 is described in the respective section of this document.

2.6.5.5. In vitro biomarker test for patient selection for efficacy

N/A

2.6.5.6. Analysis performed across trials (pooled analyses and meta-analysis)

Indication I:

A pooled analysis in 958 subjects enrolled in four studies A05E, A05C, TZAX, and PX01 showed, that

- 487 (51%) were classified as having an Advanced AD flortaucipir PET pattern (τAD++),
- 69 (7%) were classified as having a Moderate AD flortaucipir PET pattern (τAD+), and
- 401 (42%) were classified as having a non-τAD pattern (τAD-).
- Classification was not available for 1 subject.

τAD+ cohort was significantly older than either the τAD- or τAD++ subjects and τAD++ subjects tended to be more likely female, less educated, and had lower body mass index (consistent with

increased frailty). Patients with tAD++ pattern had highest proportion of population with positive family history of AD and the related genetic mutation.

The advanced flortaucipir PET pattern (τAD++) was, in descending order, more common in subjects clinically diagnosed as AD>MCI>subjective memory complaint (SMC)>other dementing disorder (ODD)> older cognitively normal (OCN), the probability of a non-AD (τAD-) pattern was exactly opposite in order. The Moderate flortaucipir PET pattern (τAD+) was observed least frequently than τAD++ or non-τAD patterns, but was higher in AD, MCI, and SMC subjects than in ODD or OCN subjects.

Within the Aβ+ subjects, the probability of an advanced AD flortaucipir PET pattern (τAD++) increased and the probability of a non-τAD pattern decreased from OCN/ODD through SMC, MCI, and AD. The probability of a Moderate τAD pattern (τAD+) was about 10% across the diagnostic categories. In contrast, in Aβ- subjects the probability of a Moderate or Advanced AD flortaucipir PET pattern (τAD+/τAD++) was low across all diagnostic groups. Overall flortaucipir PET scans for 272 of the 288 Aβ- subjects were interpreted as having a non-τAD pattern, resulting in a specificity of 94.4% for a τAD read to be associated with positive amyloid PET status.

Within the Aβ+ population, 81.3% (387/476) of participants with AD and 48.9% (136/278) with MCI were of Moderate or Advanced AD flortaucipir PET pattern (τAD+/τAD++).

Table 49 Distribution of flortaucipir PET visual interpretation by clinical diagnosis and amyloid status at baseline

	AD n=476	ODD n=27	MCI n=278	SMC n=45	OCN n=112	YCN n=20	Total n=958
Aβ+							
tAD++	349 (73.6%)	0	115 (41.4%)	9 (20%)	4 (3.6%)	0	477 (49.8%)
tAD+	38 (8.0%)	0	21 (7.6%)	2 (4.4%)	2 (1.8%)	0	63 (6.6%)
tAD-	54 (11.4%)	0	42 (15.1%)	9 (20.0%)	19 (17.0%)	0	124 (13.0%)
Aβ-							
tAD++	3 (0.6%)	2 (7.4%)	1 (0.4%)	3 (6.7%)	1 (0.9%)	0	10 (1.0%)
tAD+	1 (0.2%)	1 (3.7%)	1 (0.4%)	3 (6.7%)	0	0	6 (0.6%)
tAD-	29 (6.1%)	24 (88.9%)	98 (35.3%)	19 (42.2%)	86 (76.8%)	16 (100.0%)	272 (28.6%)
p-value 1:	<.0001	NE	<0.0001	0.0514	0.0007	NE	<0.0001
p-value 2:	<.0001						
p-value 3:	.0002						

Abbreviations: AD = Alzheimer's disease; MCI = mild cognitive impairment; n = number of patients; NE = not evaluated; OCN = old cognitively normal; ODD = other dementing disorder; PET = positron emission tomography; SMC = subjective memory complaint; YCN = young cognitively normal.

Note: Percentages are based on the total number of subjects in each cognitive status.

[1] from Fisher's exact test comparing frequency distribution of tAD++/tAD+/tAD- by amyloid status, within diagnosis.

[2] from Fisher's exact test comparing frequency distribution of tAD++/tAD+/tAD- by AD, MCI, SMC and OCN, for AB+ subjects.

[3] from Fisher's exact test comparing frequency distribution of tAD++/tAD+/tAD- by AD, MCI, SMC and OCN, for AB- subjects.

Indication II:

According to the applicant, the results of Studies A05C and PX01 were supported by analysis of the exploratory cohort of Study A05 (A05E) and Study TZAX (tau addendum of LY2062430 Study H8A-MC-LZAX). Although there were many differences among the 4 trials (see Summary of Clinical Efficacy), the most important difference between the supportive Studies A05E and TZAX and the pivotal Studies A05C and PX01 is that the flortaucipir PET images in these studies were interpreted only by Lilly expert readers rather than independent external readers. The read outcomes of the flortaucipir scans were captured identically across the 4 studies. As for clinical measurements, in Study A05E, the CDR was not conducted at baseline, thus CDR-SB baseline data were unavailable. However, for both Studies

A05E and TZAX it was possible to evaluate the relationship between visual interpretation of the baseline flortaucipir PET image and change as assessed by MMSE, ADAS-Cog11, and FAQ.

The pooled results come from a separate analysis in comparison to the individual study analyses, modified to account for differences in the data collected across studies (for example, ANCOVA was used rather than MMRM to analyse change from baseline CDR-SB because there was only 1 post-baseline timepoint for CDR-SB in Study TZAX); years (continuous) of education was not included as a cofactor in the pooled model because it was not available in Study PX01.

The table below shows risk ratios (RR) or HRs, in accordance with the pre-defined analysis for each study, confidence limits and p-values for each of the variables evaluated, in each of the 4 trials. The last column shows risk ratios for patients pooled across the trials. There is a consistent pattern across these trials for elevated risk of deterioration in the τAD++, as compared to the non-τAD++ subjects. Significant p-values were obtained in the pooled analysis for all progression endpoint measures. Similar results were obtained pooling just Studies A05C and PX01 with independent readers. Notably, all p-values in the pooled analyses remain significant after correction for multiple analyses using the Bonferroni correction procedure.

In the pooled analysis of Studies A05E, A05C, TZAX, and PX01 (Table 50), risk ratios were all nominally statistically significant and ranged from approximately 1.5 to 2. Although this level of risk prediction may seem modest, they must be taken in context with other individual risk factors patients will have for progression, and the fact that this represents only 18 months of change in a disease that has a 20–30-year natural history (Jack et al. 2013). In fact, this level of risk ratio has historically been demonstrated to be valuable in clinical practice. Evidence of such value comes from the cardiovascular and diabetes literature (Wilson et al. 1998; Parry et al. 2015; Elder et al. 2016; Skrtic et al. 2017) as well as from epidemiological data on cognitive progression in elderly patients (Deckers et al. 2015). Since progression is highly variable among individual patients due to differing degrees of risk factors, clinical presentation, and co-morbid illnesses (Swanwick et al. 1998; Doody et al. 2014), it is difficult for clinicians to predict near-term progression. Flortaucipir provides prognostic information that is potentially valuable to patient management.

Table 50: Hazard Ratio or Risk Ratio for Progression in Studies A05E, A05C, TZAX and PX01 Individually and Pooled

	RR[1]	p-value	HR[2]	p-value	RR/HR[3]	p-value	RR[4]	p-value	RR	p-value
CDR-SB 1-pt change			1.581	**.067**	1.762	.0034	1.355	.0313	1.535	<.0001
CDR-G change >0			2.371	.0039	3.33	.0016	1.283	**.2814**	1.920	<.0001
MMSE 3-pt change	2.891	.0006	2.506	.007	2.907	<.0001	1.347	**.0833**	1.974	<.0001
ADAS 4-pt change	1.486	**.0726**	1.359	**.2934**	2.554	.0003	1.771	.0141	1.907	<.0001
FAQ 3-pt change	2.110	.0007	2.785	<.0001	2.339	.0006	1.322	**.0639**	1.851	<.0001

Abbreviations: ADAS = Alzheimer's Disease Assessment Scale; CDR-SB = Clinical Dementia Rating Scale Sum of Boxes; CDR-G = Clinical Dementia Rating Global; FAQ = Functional Activities Questionnaire; HR = hazard ratio; MMSE = Mini-Mental State Examination; RR = risk ratio.

[1] Calculated in Study A05E as crude (no cofactors) relative risk (PPV/1-NPV).

[2] Cox proportional HR adjusted for baseline age, ANART, and baseline score.

[3] In the TZAX study, CDR was only performed at baseline and endpoint so RRs were calculated, whereas HRs were calculated for MMSE, ADAS and FAQ. RR and HR were adjusted for treatment arm, age, years of education, and baseline test score.

[4] PX01 RRs were adjusted for treatment arm, age, years of education, and baseline test score.

Table 51: Pooled risk ratio analysis for studies A05C and PX01

Assessment	RR (95% CI)	P-value
CDR-SB	1.396 (1.109, 1.757)	.0045
MMSE	1.683 (1.223, 2.315)	.0014
ADAS	1.729 (1.229, 2.432)	.0017
FAQ	1.742 (1.349, 2.248)	<.0001
CDR Global	1.599 (1.125, 2.273)	.0089

Abbreviations: ADAS = Alzheimer's Disease Assessment Scale; CDR-SB = Clinical Dementia Rating Scale Sum of Boxes; FAQ = Functional Activities Questionnaire; MMSE = Mini-Mental State Examination; RR = risk ratio.

Similar results were obtained when pooling just Studies A05C and PX01, the 2 studies with independent readers (Table 52). All differences for the pooled analyses remained significant after correction for multiple analyses using the Bonferroni correction procedure.

Table 52: Pooled MMRM analysis for studies A05C and PX01

Test	**тAD++ LS Mean (± SEM)**	**Non-тAD++ LS Mean (± SEM)**	**p-value**
CDR-SB change	2.282 ± 0.1733	0.978 ± 0.243	<.0001
MMSE change	-4.13 ± 0.3046	-1.27 ± 0.4218	<.0001
FAQ change	6.152 ± 0.3959	2.103 ± 0.5448	<.0001
ADAS	6.388 ± 0.507	2.295 ± 0.7085	<.0001

Abbreviations: ADAS = Alzheimer's Disease Assessment Scale; CDR-SB = Clinical Dementia Rating Scale Sum of Boxes; FAQ = Functional Activities Questionnaire; MMSE = Mini-Mental State Examination. Note: p-value is for comparing adjusted mean change from baseline values between тAD++ and non-тAD++

CDR Global Stage Change

In study A05C risk ratios were used to evaluate the relative probability of a deterioration in CDR Global in the тAD++ and non-тAD++subjects. An alternative approach used in Studies A05C, PX01, and TZAX is to evaluate not only the relative proportions of patients that deteriorate, but the relative proportion of patients that improve. This analytic approach takes advantage of all of the data available in the test, and in essence provides some correction for the time-to-time variability of the test measure. As shown in Table 53 across all 3 studies Mantel-Haenszel tests of trend showed statistically significant difference in the distribution of CDR Global scores, consistent with an overall shift away from improvement and towards worsening in the тAD++ relative to the non-тAD++ subjects.

Table 53: CDR Global stage change as a function of baseline flortaucipir PET classification for studies A05C, PX01, TZAX, and pooled

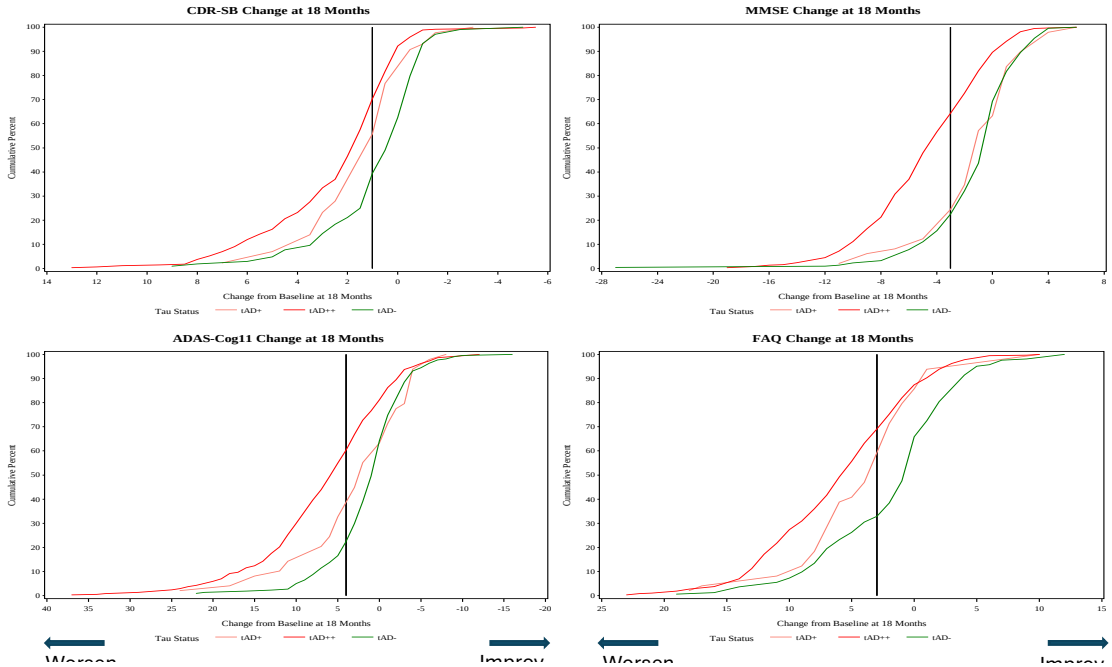
				Number of Subjects (% within each scan classification)
	A05C	**PX01**	**TZAX**	**Pooled**
Decrease (improve)				
Non-тAD++	8 (12.9%)	3 (6.98%)	6 (14.3%)	17 (11.6%)
тAD++	1 (2.1%)	3 (1.85%)	4 (3.0%)	8 (2.33%)
Increase (worsen)				
Non-тAD++	15 (24.2%)	16 (37.2%)	6 (14.3%)	37 (25.2%)
тAD++	23 (47.9%)	71 (43.8%)	61 (45.5%)	155 (45.2%)
Unchanged				

				Number of Subjects (% within each scan classification)
	A05C	**PX01**	**TZAX**	**Pooled**
Non-τAD++	39 (62.9%)	24 (55.8%)	30 (71.4%)	93 (63.3%)
τAD++	23 (47.9%)	88 (54.3%)	69 (51.5%)	180 (52.5%)
p-value	.0056	.2095	<.0001	<.0001

Abbreviations: CDR = Clinical Dementia Rating Scale; PET = positron emission tomography.

Clinical Progression Amongst τAD++, τAD+, and τAD- Subjects

For clinical purposes it is often useful to stratify patients into discrete risk groups. For the purposes of the label-enabling analysis, patients were stratified as having an Advanced AD tau pattern (τAD++) versus those that did not have an Advanced AD flortaucipir PET pattern (non τAD++). However, it is recognised that the biology of tau accumulation in AD is a continuous process that is thought to begin with aggregation in the medial temporal lobe followed by amyloid-enabled spread to the lateral temporal lobe and then beyond to the rest of the neocortex (Jack et al. 2018b). In this context it is notable that the non τAD++ group in the primary analysis included subjects that did have an intermediate pattern of AD tau (τAD+), >90% of whom were also amyloid positive by florbetapir scans. It might be speculated that these subjects could show cognitive and functional deterioration that is intermediate between that of the τAD++ and τAD- subjects. The following Figure 13 shows the cumulative frequency of deterioration on the CDR-SB, MMSE, ADAS-Cog11, and FAQ plotted separately for the τAD++, τAD+, and τAD- subjects. As expected for 3 of the 4 cognitive and functional scales, τAD+ curve (orange) and associated mean change fell between that of the τAD++ (red) and τAD- (green). For MMSE, the τAD+ subjects were indistinguishable from the τAD- subjects.



Abbreviations: ADAS-Cog11 = Alzheimer's Disease Assessment Scale – Cognitive 11; CDR-SB = Clinical Dementia Rating Scale Sum of Boxes; FAQ = Functional Activities Questionnaire; MMSE = Mini-Mental State Examination.

Figure 8: Cumulative frequency of deterioration on the CDR-SB, MMSE, ADAS-Cog11, and FAQ plotted separately for the τAD++, τAD+, and τAD- subjects

CHMP is of the view that it is not appropriate to pool the results from studies with completely different designs, completely different objectives and completely different populations.

The analysis on the diagnostic properties to predict cognitive decline from studies A05C and PX01 show completely different results for sensitivity and specificity. Whereas sensitivity is lower than specificity in Study A05C, the contrary is given in Study PX01 with an extremely low specificity. Hence it appears that conditions for predicting worsening were much stricter in study A05C than in Study PX01 which also contributes to the fact that pooling of both studies appears meaningless.

Apart from that as rightly pointed out by the applicant cognitive decline in AD is influenced by a variety of factors including genetics, concomitant disease and age-related changes and variable treatments, cognitive reserve which may positively or negatively modulate the effects of underlying AD pathology. Therefore, it is not agreed that the level of risk prediction derived from an 18-month period based on Tau imaging alone can be used for a prognostic claim.

2.6.5.7. Supportive study(ies)

It should be noted that the description and assessment of the supportive studies that follow in the sections below are a selection of the most important ones used by the applicant for the justification of the initially submitted claims.

Indication I: "…estimate the density and distribution of the aggregated tau neurofibrillary tangles (NFTs) of Alzheimer's disease (AD) in adult patients who are being evaluated for AD. Flortaucipir (18F) can be used to help establish a diagnosis of AD…"

Histelide Analysis of Study A16 Frontrunner Cases (Pontecorvo et al. 2020)

The first 3 patients to come to autopsy were excluded from the primary analysis of Study A16; however, the cases were used in an exploratory, post-hoc analysis to compare flortaucipir PET signal quantitation regionally with regional histopathology (referred to as the "histelide analysis" as described

and reported in Pontecorvo et al. 2020). The purpose of this analysis was to provide supporting evidence that the distribution and density of neocortical tau pathology in the brain is accurately reflected on the flortaucipir PET scan.

By pathology, all 3 cases had B3 (Braak Stage V or VI) NFT and CERAD moderate or frequent neuritic plaques. Case 1 and Case 3 met criterion for high ADNC, with Thal amyloid stage of 4, but for Case 2 Thal stage was just 3.

The tissue blocks showed tau pathology ranging from absent to frequent. Within neocortical regions, flortaucipir SUVr correlated significantly with the tau immunoassay (Pearson r = 0.81; p<0.001). In contrast, after adjustment for correlation between Aβ and tau tissue concentrations, there was no correlation between flortaucipir SUVr and Aβ tissue concentration.

There was a moderate-to-strong correlation between limbic (hippocampal/entorhinal and amygdala) flortaucipir SUVr and the histelide estimate of tau tissue concentration (Pearson r=0.52; p=.080). The slope of the relationship appeared steeper than for neocortical measures, suggesting a possibly reduced sensitivity of flortaucipir SUVr to tau in these limbic regions. There was no correlation (Pearson r=0.27, p=.280) between phosphorylated tau and flortaucipir SUVr in subcortical regions including ventromedial globus pallidus and caudate, putamen and thalamus.

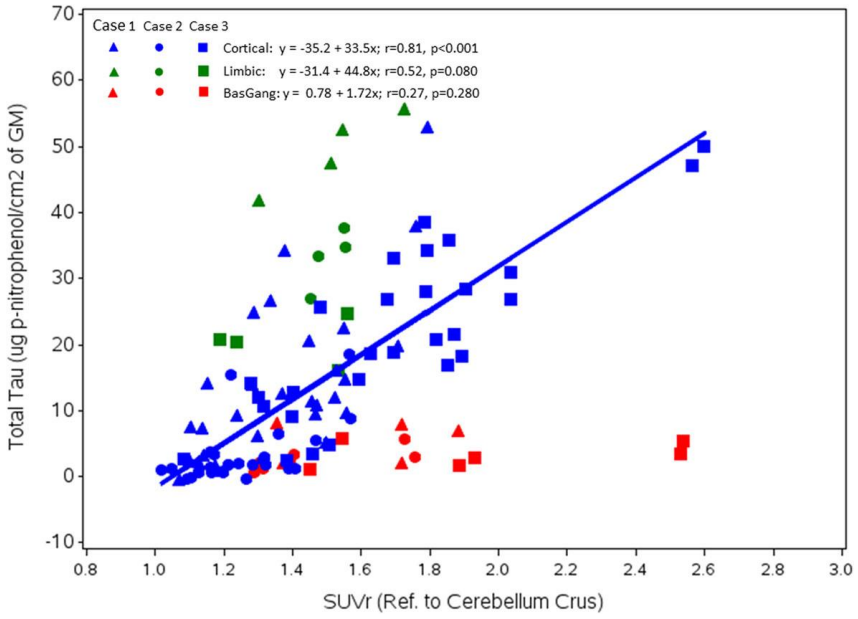


Figure 9: *Relationship between flortaucipir PET signal and regional phosphorylated tau concentration (Study A16 front-runners)*

Abbreviations: SUVr = standardised uptake value ratio.

Associations shown between regional flortaucipir SUVr and phosphorylated tau total tissue concentration (Histelide) values at autopsy. Regression line shows the relationship between cortical regional SUVr and cortical regional phosphorylated tau. Cortical regions are shown in blue, limbic in green, and basal ganglia in red. Symbols differentiate individual subjects.

Further published Studies Relating Flortaucipir PET Signal to Pathology at Autopsy

The following summary was provided to support the claim that flortaucipir (18F) can evaluate the density of tau presence in the brain.

Josephs et al. 2022a investigated flortaucipir PET ability, at quantification, in discriminating between patients with an autopsy-confirmed FrontoTemporal Lobar Degeneration (FTLD) versus patients with an autopsy-confirmed low/intermediate (Braak ≤IV, B0-2) or high ADNC (Braak≥V, B3) pathology (N=143 participants). The authors investigated several quantitation pipelines and showed that flortaucipir PET could achieve optimum discrimination between autopsy-confirmed FTLD vs. high ADNC, reaching an AUC of 0.99 (95%CI 0.94-1.00), a sensitivity of 94% and a specificity of 95%.

Additional smaller scale studies have provided similar evidence showing that elevation in flortaucipir PET in key regions associated with AD was associated with AD pathology at autopsy (Soleimani-Meigooni et al. 2020; Lowe et al. 2020).

Other studies adopted a different design, investigating the association between regional flortaucipir PET signal and regional quantitative measures of tau pathology at autopsy, for example via immunohistochemistry. When performed on or including patients with established AD pathology, all these studies showed significant associations between regional flortaucipir SUVr and regional tau pathology at autopsy, reinforcing the notion that flortaucipir PET signal *in vivo* accurately reflects density and distribution of AD-like tau pathology (Smith et al. 2019a; Lowe et al. 2020; Martersteck et al. 2022; Coughlin et al. 2022).

Flortaucipir (18F) PET in non-AD dementia and other conditions

Four studies evaluated the pattern and density of flortaucipir PET signal in patients with presumed non-AD tauopathies/traumatic encephalopathy. Study A09 evaluated flortaucipir in Aβ- subjects with clinically defined Progressive Supranuclear Palsy (PSP) and corticobasal degeneration (CBD). Study A19 evaluated flortaucipir in subjects with fronto-temporal dementia (FTD). Studies A07 and A11 evaluated flortaucipir in subjects at risk for chronic traumatic encephalopathy (CTE).

Study A09 was a Phase 1, longitudinal observational study that evaluated imaging characteristics of flortaucipir in males and females 50 to 85 years of age, including patients with PSP (n=20) or CBD (n=6), and cognitively healthy volunteers (n=3). Diagnosis of PSP was established based on the criteria by Bensimon et al. 2009 and of CBD based on the criteria by Armstrong et al. 2013. Both PSP and CBD subjects must have had an MMSE score between 14 and 30 (inclusive). Patients were excluded from the PSP/CBD cohorts if they were amyloid positive by florbetapir scan, met NIA-AA clinical criteria for probable AD, or had conditions/diagnosis other than PSP or CBD that could account for their deficits (for example, genetic mutations, PD, DLB, cerebrovascular disease).

In this study there was a small but significant increase in composite, globus pallidus, and dentate SUVr in the PSP patients and to a lesser extent in CBD patients compared to healthy volunteers. Two of the CBD subjects were classified as τAD with 1 case showing minimal uptake in PLT and the other significant uptake in PLT/occipital/parietal as well as some frontal region. In both cases, coincident AD and CBD pathology cannot be ruled out.

Study A19 was a Phase 1 study that evaluated imaging characteristics of flortaucipir in subjects with FTD (n=16), defined as clinical syndromes associated with FTD pathology, including behavioural-variant FTD, FTD with motor-neuron disease, non-fluent/agrammatic and semantic variants of primary progressive aphasia, PSP, and CBS.

Patients with amyloid were excluded and one patient discontinued, so that 14 patients were in the ITT set. All flortaucipir PET scans were read as τAD-. Mean flortaucipir AD-signature VOI SUVr was 1.0726 ± 0.2500 in the ITT population (N = 14).

Studies A07 and A11 were designed to explore the use of flortaucipir as a biomarker for brain injury related to repetitive head trauma; and to examine the relationship between clinical presentation and tau deposition as measured by flortaucipir uptake in subjects at risk for CTE. Flortaucipir images were

interpreted qualitatively by Lilly expert readers for overall level of uptake (ranging from 0=no to 3=intense uptake), but given the minimal level of uptake observed, were not read for tau AD pattern.

Subjects in Study A07 included a total of 28 high-risk CTE subjects and 11 age-matched controls. There was no significant difference in the level of flortaucipir uptake, based on 4 category visual read assessment, between the CTE High-Risk group and age-matched control group. One subject was considered to have mild uptake of flortaucipir, and this subject was Aβ+.

Subjects in Study A11 included a total of 30 subjects: 5 cognitively/functionally impaired and 5 CN active fighters plus 13 cognitively/functionally impaired and 7 CN retired fighters. Higher mean flortaucipir uptake as measured by SUVr was observed in the frontal region of active fighters vs. retired fighters. However, with respect to visual read, only 2 retired fighters with cognitive and functional impairment showed mild uptake; no subjects in the active fighters group demonstrated visual uptake. Thus, there is no conclusive evidence that flortaucipir PET provides useful diagnostic information on potential brain injury in active or retired professional fighters.

Flortaucipir (18F) PET in atypical AD and in Down Syndrome

Flortaucipir has been used to investigate binding in atypical variants of AD, including logopenic variant primary progressive aphasia, posterior cortical atrophy, and the behavioural AD phenotype (Gorno-Tempini et al. 2011; Crutch et al. 2017; Ossenkoppele et al. 2022).

The topography of flortaucipir signal patterns varies across phenotypes, with a topography consistent within the clinical presentation. Patients with logopenic variant primary progressive aphasia may present with asymmetric (L>R) temporo-parietal flortaucipir signal (Ossenkoppele et al. 2016; Dronse et al. 2017; Xia et al. 2017; Nasrallah et al. 2018; Josephs et al. 2018; Tetzloff et al. 2018; Sintini et al. 2019a; Sintini et al. 2019b; Iaccarino et al. 2021; Martersteck et al. 2021; Phillips et al. 2022). Posterior cortical atrophy may present with greater-than-usual posterior parietal/occipital flortaucipir signal, especially in lateral and inferior aspects (Ossenkoppele et al. 2015; Ossenkoppele et al. 2016; Day et al. 2017; Dronse et al. 2017; Xia et al. 2017; Nasrallah et al. 2018; Tetzloff et al. 2018; Nedelska et al. 2019; Sintini et al. 2019a; Sintini et al. 2019b; Iaccarino et al. 2021; La Joie et al. 2021; Visser et al. 2022; Phillips et al. 2022). As for behavioural AD, flortaucipir signal has been shown to be more heterogeneous but still consistent with the typical AD temporo-parietal pattern (Ossenkoppele et al. 2016; Dronse et al. 2017; Li et al. 2020; Singleton et al. 2021). In addition, flortaucipir signal patterns in atrophy-defined subtypes of AD (for example, "typical", "limbic-predominant", "hippocampal-sparing", and "mild atrophy") also differ, with the greatest neocortical signal observed in hippocampal-sparing and typical patients, while limbic-predominant patients showed particularly high entorhinal signal (Ossenkoppele et al. 2019a; Ossenkoppele et al. 2020a; Vogel et al. 2021).

In Down syndrome, the trisomy 21 causes a duplication of the Amyloid Precursor Protein gene, leading to over-production of amyloid beta (Fortea et al. 2020). Flortaucipir studies in Down syndrome have reported elevated signal spatially consistent with observed findings in typical AD (Rafii et al. 2017; Mak et al. 2019; Tudorascu et al. 2020; Zammit et al. 2021; Hartley et al. 2022) and that neocortical flortaucipir signal is associated with cognitive decline (Rafii et al. 2017; Hartley et al. 2022), even with only subthreshold amyloid accumulation (Zammit et al. 2021).

Flortaucipir PET in Older Subjects with Normal Cognition

Study A20 was a Phase 1, open-label, single-centre study that evaluated imaging characteristics of flortaucipir in subjects enrolled in the Biomarkers of Cognitive Decline Among Normal Individuals (BIOCARD) study. By design, more than two-thirds of participants had a family history of AD. This study planned to recruit approximately 100 subjects from the BIOCARD study with the aim of having 80% Aβ+ subjects and 20% Aβ- subjects. However, the study was terminated early due to lower-than-

expected enrolment (n=23). Flortaucipir PET scan images were visually interpreted as consistent (τAD+/++) or not consistent with a tau AD pattern (τAD-), using the same criteria as in the pivotal studies. Visual interpretation of flortaucipir PET images showed that:

- all subjects with flortaucipir PET scans interpreted as τAD++ (consistent with AD pattern) were Aβ+, and
- all subjects who were Aβ- were τAD- (not consistent with AD pattern).

Flortaucipir PET in the patients with MCI

With the responses to the day 120 LoI the detection rates for positive tau-PET image in the patients with MCI were provided. According to the analysis of integrated efficacy datasets from Studies A01, A04, A05, A08, A09, PX01, and TZAX (substudy of subjects who received flortaucipir PET scans in therapeutic Study H8A-MC-LZAX [LZAX]), 387 out of 441 (87.8%) Aβ+ participants living with dementia were tau PET positive (AD pattern) by visual read. Among Aβ+ MCI, 136 out of 178 (76.4%) were tau positive (tau-AD pattern) by visual read. Among 404 Aβ+ participants with a baseline MMSE score ≥20 (equivalent to early symptomatic AD) 350 (86.6%) of them were tau positive (tau-AD pattern) by visual read.

Among 100 amyloid PET negative (Aβ-) MCI subjects only 2 (2%) were considered flortaucipir visually positive (tau-AD pattern).

Impact on diagnosis and patient management

Altmore et al., 2021:

Purpose: Assess the individual and combined diagnostic value of amyloid-PET and tau-PET in a memory clinic population.

Methods: Clinical reports of 136 patients with cognitive complaints were randomly assigned to two diagnostic pathways: AMY-TAU, amyloid-PET is presented before tau-PET; and TAU-AMY, tau-PET is presented before amyloid-PET. Two neurologists independently assessed all reports with a balanced randomised design, and expressed etiological diagnosis and diagnostic confidence (50–100%) three times: (i) at baseline based on the routine diagnostic workup including clinical and neuropsychological assessments and MRI, which were collected within an ongoing prospective research study, (ii) after the first exam (amyloid-PET for the AMY-TAU pathway, and tau-PET for the TAU-AMY pathway), and (iii) after the remaining exam. The main outcomes were changes in diagnosis (from AD to non-AD or vice versa) and in diagnostic confidence.

Image assessment for tau PET was based on a structured scheme depicting four regions of interest (ROIs) corresponding to the neuropathological stages defined by Braak for tau deposition. The visually estimated Braak stage was defined considering the hemisphere showing the highest tau positivity in the ROIs evaluated and tau positivity was defined when the rater score was Braak stage ≥ IV. ROIs corresponding to Braak IV included anterior, inferior, mid and superior temporal regions and frontal inferior region; Braak stage V - parietal and frontal lobes; and Braak stage VI - sensory-motor areas and visual primary cortex. Atypical presentation with medial temporal lobe sparing was considered.

Results*:*

Patients were on average 74 (70–78) (median, 25th–75th percentile) years old, males and females were equally distributed. Among them, 51% (69/136) had MCI, 32% (43/136) had dementia, and 18% (24/136) had subjective cognitive decline (SCD). As a result of the randomisation, there were no significant differences in demographic and clinical features between ARM1 and ARM2 patients.

The baseline etiological diagnosis was inconsistent with the amyloid-PET result in 28% (38/136) of cases (i.e., 28 amyloid-negative AD, and 10 amyloid-positive non-AD patients), and with the tau-PET result in 37% (51/136) of cases (i.e., 46 tau-negative AD, and 5 tau-positive non-AD patients).

Amyloid-PET and tau-PET, when presented as the first exam, resulted in a change of etiological diagnosis in 28% (p = 0.006) and 28% (p < 0.001) of cases, and diagnostic confidence increased by 18% (p < 0.001) and 19% (p < 0.001) respectively, with no differences between exams (p > 0.05).

Table 54: *Overview of the primary outcomes (changes in diagnosis and changes in diagnostic confidence) in the whole sample and in the three cognitive stage groups (SCD, MCI, dementia)*

Primary outcomes		Change in diagnosis (%, *p* value)				Change in diagnostic confidence (%, *p* value)			
		Whole sample *n*=136	SCD *n*=24	MCI *n*=69	Dementia *n*=43	Whole sample *n*=136	SCD *n*=24	MCI *n*=69	Dementia *n*=43
First exam	Amyloid-PET	**28%, *p* =0.006**	17%, *p*=1.00	**35%, *p* =0.008**	23%, *p*=0.343	**+18%, *p* <0.001**	**+23%, *p* <0.001**	**+19%, *p* <0.001**	**+14%, *p* <0.001**
	Tau-PET	**28%, *p* <0.001**	17%, *p*=0.617	**39%, *p* <0.001**	**16%, *p* =0.023**	**+19%, *p* <0.001**	**+26%, *p* <0.001**	**+22%, *p* <0.001**	**+10%, *p* <0.001**
Second exam	Amyloid-PET	6%, *p*=0.077	0%	6%, *p*=0.134	9%, *p*=0.617	**+4%, *p* <0.001**	+3%, *p*=1.00	+3%, *p*=0.114	**+5%, *p* =0.005**
	Tau-PET	9%, *p*=0.149	0%	12%, *p*=0.289	9%, *p*=0.617	**+5%, *p* <0.001**	**+6%, *p* =0.007**	**+4%, *p* =0.005**	**+ 6%, *p* =0.001**

Bold *p* values denote statistical significance ($p < 0.05$)

Stronger impact of a negative amyloid-PET versus a negative tau-PET (p = 0.014) was observed. When added as the second exam, amyloid-PET and tau-PET resulted in a further change in etiological diagnosis in 6% (p = 0.077) and 9% (p = 0.149) of cases, and diagnostic confidence increased by 4% (p < 0.001) and 5% (p < 0.001) respectively, with no differences between exams (p > 0.05).

The amyloid-PET and tau-PET results were reciprocally discordant in 20% (27/136) of cases, with 25 patients with positive amyloid-PET and negative tau-PET (Braak stage = 0 in 14 patients, Braak stage = I–III in 11 patients), and 2 cases having patients with negative amyloid-PET and positive tau-PET (Braak stage = V in both cases). Among these discordant patients, the final (T2) diagnosis was consistent with the amyloid-PET result in 55% (15/27) of cases for the first rater and in 67% (18/27) for the second rater, and with the tau-PET result in the remaining cases. The table below includes only those discordant cases, where diagnosis was also changed.

Table 55: *Detailed information on the diagnostic impact of the second PET scan in the two diagnostic pathways (AMY-TAU and TAU-AMY)*

Diagnostic pathway	N of patients	T0 diagnosis	First PET scan	T1 diagnosis	Second PET scan	T2 diagnosis	Observations
AMY-TAU	8	AD	AMY+	AD	TAU–	Non-AD	• Discordant PET results • Same diagnosis at T0 and T1
	1	Non-AD	AMY+	AD	TAU–	Non-AD	• Discordant PET results • T1 diagnosis changed after first PET scan
	1	AD	AMY–	Non-AD	TAU+	AD	• Discordant PET results • T1 diagnosis changed after first PET scan
	1	Non-AD	AMY–	Non-AD	TAU+	AD	• Discordant PET results • Same diagnosis at T0 and T1
	1*	AD	AMY–	Non-AD	TAU–	AD	• Concordant PET results • T1 diagnosis changed after first PET scan
TAU-AMY	1	AD	TAU–	AD	AMY–	Non-AD	• Concordant PET results • Same diagnosis at T0 and T1
	4	Non-AD	TAU–	Non-AD	AMY+	AD	• Discordant PET results • Same diagnosis at T0 and T1
	2	AD	TAU–	Non-AD	AMY+	AD	• Discordant PET results • T1 diagnosis changed after first PET scan
	1	Non-AD	TAU+	Non-AD	AMY+	AD	• Concordant PET results • Same diagnosis at T0 and T1

AMY+/–: amyloid positive/negative, *TAU+/–*: tau positive/negative

This table includes only patients who changed diagnosis after amyloid-PET ($n = 8$) or tau-PET ($n = 12$) presented as the second exam

*The only case in which the diagnosis changed after a consistent PET result was a patient with a T1 diagnosis of MCI not due to AD based on a negative amyloid-PET whose diagnosis changed to AD at T2 after a negative tau-PET scan (Braak stage = I–III)

The authors concluded that Amyloid-PET and tau-PET significantly impacted diagnosis and diagnostic confidence in a similar way, although a negative amyloid-PET has a stronger impact on diagnosis than a negative tau-PET. Adding either of the two as second exam further improved diagnostic confidence.

Further published evidence to support the role of flortaucipir PET in the diagnostic work-up of AD include 2 studies investigating the diagnostic value of tau PET (Smith et al. 2023; Shimohama et al. 2023). Additionally, the latest IWG recommendations for a clinical diagnosis of AD (Dubois et al. 2021), as well as the recently presented update from the NIA-AA workgroup (NIA-AA Revised Clinical Criteria [AAIC 2023]), and the Appropriate Use Criteria also recently presented (Rabinovici et al. 2023) have been presented.

Indication II: Flortaucipir (18F) can be used to inform risk for progression of cognitive and functional impairment due to AD

Study 18F-AV-1451-A18 (Study A18)

Title: An Open-Label, Multicenter Study Evaluating the Imaging Characteristics of a Follow-Up 18F-AV-1451 Scan in Subjects That Participated in the Confirmatory Cohort of 18F-AV-1451-A05.

Enrolled: 79 (Alzheimer's disease [AD], n = 25; mild cognitive impairment [MCI], n = 54)

Treated (at least 1 dose): 79 Completed: 79

A subgroup of the subjects who completed A05C were followed up.

The primary objective of this protocol was to assess longitudinal change of tau deposition as measured by flortaucipir F18 uptake over time. The study did not include secondary objectives.

Outcomes/endpoints A18

The primary efficacy measure was flortaucipir standardised uptake value ratio (SUVr) change from baseline (CFB) of Study 18F-AV-1451-A05, as measured by the multiblock barycentric discriminant analysis (MUBADA) method (Abdi et al. 2012), which was used to define a voxel-weighted volume of interest (MUBADA VOI; Devous et al. 2018). The MUBADA SUVr was calculated as a ratio of the MUBADA target region to a white matter-based parametric estimate of reference signal intensity (PERSI reference region) as described by Southekal et al. (2017).

Other efficacy variables measured were Mini-Mental State Examination (MMSE), Clinical Dementia Rating Scale – Sum of Boxes (CDR-S), and brain volume by magnetic resonance imaging.

Randomisation and blinding (masking) A18

No randomisation occurred in the study. The investigational product was used open label.

Avid personnel were blinded to the MRI and flortaucipir F 18 PET scans performed on subjects with some exceptions. Avid personnel remained blinded to the MRI and flortaucipir F 18 PET scans until the Study A05 database was locked.

Study A18 - Results

Study A18 is a Phase 2, open-label, outpatient study that evaluated longitudinal change of tau deposition as measured by flortaucipir uptake over time. The participants received the same dose as in study A05C. The only objective of this study A18 was the follow-up of the subjects who participated in the Confirmatory cohort of study A05. This study has been characterised as a safety study for the follow-up of the participants from study A05C but it is rather a longitudinal follow-up of change in flortaucipir PET in Study A05C subjects.

The statistical analysis is acceptable, in general. However, no multiplicity adjustment for the different comparisons was planned, resulting in an explorative assessment also due to the lack of a control group.

Study A18 was one year follow-up of subjects from the study A05C and a participant's flow is not available since no new subjects were enrolled in this study. No subjects discontinued for any reason. 25 AD and 54 MCI subjects were included and competed the study.

As presented and summarised in the following Table 56, overall, there was an increase in flortaucipir SUVr in Aβ+ subjects (LSM change = 0.0303 ± 0.01108, p=0.0080). The increase in flortaucipir SUVr in the Aβ+ group appears to be driven by the SUVr increase in MCI/Aβ+ subjects (LSM change = 0.0352 ± 0.01185, p=0.0046), as the SUVr did not increase significantly in AD/Aβ+ subjects (LSM change = 0.0205 ± 0.01871, p=0.2846). In addition, there was a decrease in flortaucipir SUVr in Aβ- subjects (LSM change = -0.037 ± 0.01611, p=0.0245).

The directionality of the changes in the Aβ+ and Aβ- groups resulted in a significant difference in LSM change difference in flortaucipir SUVr between these groups (LSM change difference = -0.0673 ± 0.02083, p=0.0019). There were no differences in flortaucipir SUVr CFB between the AD group (CFB = 0.0024 ± 0.10211) and the MCI group (CFB = 0.0328 ±0.07378) in subjects who were Aβ+ (LSM change difference = -0.0154 ± 0.02599, p=0.5565) or those who were Aβ- (LSM change difference = -0.0095 ± 0.01264, p=0.4601).

Table 56: Flortaucipir SUVr change from baseline as a function of diagnosis group by amyloid status

	AD		**MCI**		**Total**	
	Aβ+ N = 19	**Aβ- N = 6**	**Aβ+ N = 26**	**Aβ- N = 26**	**Aβ+ N = 45**	**Aβ- N = 32**
CFB	0.0024 ± 0.10211	-0.0339 ± 0.04203	0.0328 ± 0.07378	-0.01 ± 0.02721	0.02 ± 0.08711	-0.0145 ± 0.03117
LSM CFB	0.0205 ± 0.01871	-0.0912 ± 0.03619	0.0352 ± 0.01185	-0.0124 ± 0.01185	0.0303 ± 0.01108	-0.037 ± 0.01611
P-valuea	p = 0.2846	p = 0.0199	p = 0.0046	p = 0.3006	0.0080	0.0245
P-valueb	p = 0.0166		p = 0.0108		0.0019	
P-valuec			p = 0.5565	p = 0.4601		

Abbreviations: Aβ = beta amyloid; AD = Alzheimer's disease; CFB = change from baseline; LSM = least squares mean; MCI = mild cognitive impairment.

a P-value is the result of ANCOVA testing whether the respective amyloid groups' LSM Change, within each diagnosis separately, is equal to 0.

b P-values are the result of ANCOVA models testing if the difference between LSM Changes of Aβ+ and Aβ- within each diagnosis group and overall are equal to 0. The models comparing within group differences was run separate from the model comparing overall differences.

c P-values are the result from ANCOVA models run separately for each amyloid group testing if the difference between LSM Changes of AD and MCI is equal to 0.

According to the applicant, overall, there was a statistically significant increase in flortaucipir SUVr in Aβ+ subjects, a significant decrease in Aβ- subjects, and a significant difference in CFB SUVr between the Aβ+ and Aβ- groups.

Within the either the Aβ+ or Aβ- group, there was no significant difference in CFB SUVr between the MCI and AD subjects.

There was no difference in SUVr between AD and MCI Aβ- subjects.

The correlation between CFB in SUVr and CFB in cognition was not significant.

The applicant's conclusions with respect to change from baseline SUVr and relationship between change from baseline and cognitive change may be limited by technical difficulties in quantitation of the white matter reference region in the subjects with the highest baseline SUVr, perhaps due to the high flortaucipir retention in adjacent grey matter or due to tissue loss and increased ventricular volume.

According to the applicant, overall, there was an increase in flortaucipir SUVr in Aβ+ subjects and a decrease in Aβ- subjects, and a significant difference in SUVr between the Aβ+ and Aβ- groups. However there were no statistically significant differences in CFB between AD and MCI subjects. This study demonstrates expected directional longitudinal changes in flortaucipir SUVr based on the presence of AD pathology.

The study may indicate some differences between Aβ+ and Aβ- within each diagnosis group. However, with no comparator group, it is difficult to interpret these results. In addition, in this supportive study

as opposed to the pivotal studies, scans were evaluated quantitatively (standard uptake value ratio [SUVr]). Apart from that, the findings are not supportive of the applicant´s position that tau scan is predictive of amyloid beta status.

Study A05E

Two supportive trials, Study 18F-AV-1451-A05E (A05E) and Study 18F-AV-1451-TZAX (TZAX), were completed to assess relationship between flortaucipir PET signal and 18-month cognitive decline. Since A05E was a hypothesis generating study for A05C within study A05 this study phase is described in sections above and not repeated here.

H8A-MC- LZAX Tau Imaging Addendum (Study TZAX)

H8A-MC-LZAX (6.2) Tau Imaging Addendum (TZAX): A Post-Hoc Evaluation of the Relationship between Flortaucipir PET Signal and Cognitive Change in Subjects with Mild Alzheimer's Disease Participating in the Solanezumab Expedition 3 Clinical Trial 'Effect of Passive Immunization on the Progression of Mild Alzheimer's disease: Solanezumab (LY2062430) Versus Placebo'

Study TZAX – Methods

Study LZAX was a multicentre, randomised, double-blinded, placebo-controlled, Phase 3 study comparing 400-mg solanezumab with placebo given as infusion once every 4 weeks over 76 weeks in approximately 2100 outpatients with mild AD.

Similar to Study PX01, Study TZAX was designed to take advantage of a therapeutic trial cohort that had flortaucipir PET scans and up to 18-month cognitive follow-up to further evaluate the relationship between flortaucipir PET signal and cognitive and functional decline. The subjects in Study TZAX were recruited to the tau imaging addendum of the H8A-MC-LZAX (LZAX) therapeutic clinical trial (hence Study TZAX) evaluating the anti-amyloid antibody solanezumab (LY2062430).

Objectives

The primary objective of the study was to determine whether baseline visual read patterns predicted higher risk of clinically meaningful cognitive and functional deterioration within 18 months of scan for the CDR-SB or CDR Global change.

Participants

Number of Subjects: Planned: 500, Enrolled: 224, Treated (at least 1 dose): 206, Completed: 181

The patients who participated in Study LZAX were males and females 55 to 90 years old with AD, as demonstrated by meeting National Institute of Neurological and Communicative Disorders and Stroke/AD and Related Disorders Association criteria for probable AD; having a Modified Hachinski Ischaemia Scale score of ≤4, an MMSE score of 20 through 26 at Visit 1, and a Geriatric Depression Scale short-form score of ≤6; having an MRI or CT scan performed within the past 2 years with no findings inconsistent with a diagnosis of AD; and having screening florbetapir PET scan or CSF results consistent with the presence of amyloid pathology (Aβ+ classification). A subset of LZAX sites that participated in the tau addendum contributed the subjects that were considered Study TZAX. A total of 204 subjects enrolled in the tau addendum study with evaluable tau baseline scans, including 97 subjects in the placebo group and 107 subjects in the treatment group. The efficacy population included all subjects regardless of treatment assignment with valid, interpretable PET images and the relevant clinical/cognitive and functional assessment (CDR-SB, MMSE, FAQ, ADAS -Cog11), at the relevant post-baseline time-point.

Treatment

Flortaucipir F 18 (18F-AV-1451), 240 MBq [6.5 mCi] as an intravenous (IV) bolus administration. No comparator was administered in this study. The dose in this study (240MBeq) is different from the one recommended for routine use (370 MBeq).

Study TZAX – Results

There was a statistically significant increase in risk (Risk Ratio = 1.76, p = .0034) for τAD++ vs. non-τAD++ subjects to have a clinically meaningful change (CDR-SB change >1) at 80 Weeks. LS mean change in CDR-SB for τAD++ was 2.4, while for non-τAD++ it was 0.7 (p<0.0001).

Similarly, for the CDR Global Score the analysis showed that baseline tau read predicted clinically meaningful changes in τAD++ versus non-τAD++ with a Risk Ratio of 3.3330 (p=0.0016).

Significant ratios (lower confidence limits > 1.0 indicating a greater proportion of τAD++ subjects passed the pre-specified threshold for clinically significant worsening) were also obtained on other indices shown in the table below, including CDR global (change greater than 0), MMSE (≥3 point decrease), FAQ (≥3 point increase), and ADAS-Cog11 (≥4 point increase).

Table 57: Risk or Hazard Ratios by Flortaucipir PET Classification

Test	RR/HR	CL	p-value
CDR-SB change >1	1.762	1.206, 2.576	0.0034
CDR-SB change ≥ 2.5	2.859	1.306, 6.260	0.0086
CDR Global change >0	3.333	1.576, 7.048	0.0016
MMSE change, ≥ 3 point decrease	2.907	1.675, 5.044	0.0001
FAQ change, ≥ 3 point increase	2.339	1.440, 3.798	0.0006
ADAS-Cog change, ≥ 4 point increase	2.554	1.543, 4.226	0.0003

There were also greater mean changes for τAD++ versus non-τAD++ with respect to in MMSE, ADAS-Cog11 and FAQ. At 80 weeks, LS Mean ΔMMSE for τAD++ was -4.1, while for non-τAD++ it was -0.7 (p <.0001). Similarly, for ADAS-Cog11 LS Mean changes at 80 weeks were 6.0 for τAD++ and 0.3 for non-τAD++ (p <.0001), and for FAQ they were 5.5 for τAD++ and 2.3 for non-τAD++ (p=.0004). Similarly, time to event analyses showed a large preponderance of events at or before 80 weeks for MMSE, ADAS-Cog11 and FAQ in the τAD++ group versus the non-τAD++ group.

In the secondary objective analysis, the diagnostic performance of baseline tau status in predicting CDR-SB clinically meaningful change at 80 weeks showed good sensitivity (84.7%), but relatively low specificity (38.5%). The positive PPV was 70.1, but the NPV was only 59.5. The relationship of quantitative measures of flortaucipir uptake (MUBADA SUVr) with respect to itself and to cognitive/functional outcome was examined. The MUBADA SUVr increased more in τAD++ than in non-τAD++ over 80 weeks. Further, there was a significant relationship between baseline MUBADA SUVr and 18-month changes in MMSE, ADAS-Cog11, FAQ and CDR-SB.

Similar as the PX01 study this study was a post-hoc evaluation of a parent study, in this case the Ely Lilly proprietary Solanezumab Expedition 3 study (LZAX). Also here the inclusion criteria were different to those of the A05 study, as all patients included were positive for amyloid and were diagnosed with mild AD (https://beta.clinicaltrials.gov/study/NCT01900665).

The generalizability of the results is questionable.

18F-AV-1451-A04 (Study A04)

An open-label, multicentre study, evaluating the imaging characteristics of a follow-up 18F-AV-1451 scan in subjects who have had a prior 18F-AV-1451 scan.

This Phase 2, open-label multicentre study was designed as an opportunistic study to evaluate longitudinal change in flortaucipir/tau accumulation in cognitively impaired (CI) and cognitively normal healthy volunteer (HV) male and female subjects ≥50 years of age who had previously received flortaucipir PET scans.

Study A04 was designed to provide an early preliminary estimate of longitudinal change in flortaucipir PET signal. Subjects who had previously received a flortaucipir PET scan as part of either Studies A01/T807000, A03, or an independent investigator-sponsored study, were recruited to return for a follow-up flortaucipir scan and cognitive and functional testing from 0.59 and 2.34 years after the initial scan.

Objectives

The primary objective of this study was to assess the rate of change of tau deposition as measured by 18F-AV-1451 uptake over time.

The secondary objective of this study was to expand the 18F-AV-1451 safety database.

The exploratory objective of this study is to assess the relationship among 18F-AV-1451 uptake, cognitive function measurements, and other collected biomarkers at baseline and over time.

Participants

Number of subjects: Planned: up to 100 subjects; enrolled: 44; treated (at least 1 dose): 37 active drug; completed: 37 active drug.

Patient amyloid status was classified by the investigator based on a previous PET scan with any amyloid PET tracer when available. Patients for whom amyloid status was not known underwent a florbetapir PET scan (370 MBq; 80-100 min post-injection). None of the subjects had CDR scores at baseline, and only 4 had ADAS-Cog11 scores; 24 had baseline MMSE, but in many cases the MMSE was performed and scored using local materials and methods and might not strictly be comparable to the MMSE performed in Lilly studies.

Treatment

All subjects received a single intravenous (IV) bolus administration of approximately 370 MBq (10 mCi) of flortaucipir F 18 injection. Subjects whose amyloid status was unknown received a single IV bolus administration of approximately 370 MBq (10 mCi) of florbetapir F 18 injection. No comparator was administered in this study.

Study A04 – Results

Flortaucipir PET SUVrs were calculated, but visual interpretations of flortaucipir PET scans were not performed during the study. A Lilly expert did perform a visual interpretation after the completion of the study. However, because of the small sample size, variable follow-up interval, and limited longitudinal cognitive and functional test data, the flortaucipir PET-interpretation/cognitive-change results were not integrated with the primary and supportive trials above.

In this population, there was a statistically significant difference in baseline flortaucipir SUVr between CI (n=16) and HV subjects (n=21). Flortaucipir PET SUVr in the HV group was low for both Aβ+ and Aβ- subjects (1.0079 ± 0.0486 and 1.0274 ± 0.0223 respectively). Consistent with the hypothesis that the spread of tau may be implicated in the onset and progression of cognitive and functional impairment, in the subgroup of subjects for whom MMSE data was available at baseline (n=24), there was a significant correlation (r = -0.803, p<0.0001) between baseline SUVr and MMSE.

Baseline flortaucipir SUVr was statistically significantly higher in CI vs. HV subjects. Only one HV subject (Aβ+ by historical PiB PET), had an elevated global (MUBADA) flortaucipir SUVr. Only one Aβ- subject (based on a historical PiB PET) had an elevated flortaucipir SUVr (SUVr = 1.2675). Together with previous Avid studies (Pontecorvo et al. 2017) and published reports (Brier et al. 2016; Johnson et al. 2016; Schöll et al. 2018) these results suggest that the presence of amyloid neuritic plaques (i.e., positive amyloid PET scan) is necessary for an elevated MUBADA flortaucipir SUVr, indicative of widespread flortaucipir retention and presumably widespread tau NFTs. In the subgroup of subjects for whom MMSE data were available at baseline, there was a significant correlation (r = -0.803, p<0.0001) between baseline SUVr and MMSE, and a smaller, nonsignificant correlation (r = -0.392, p = 0.0709) between baseline SUVr and change in MMSE. As already noted by the applicant, although there was a numerically greater change from baseline (CFB) standardised uptake value ratio (SUVr) for the CI group compared with the HV group, the difference in SUVr slope (change across patients as a function of duration of follow up) was not statistically significant. This latter finding could be affected by the fact that baseline MMSE values were obtained using varied site-specific procedures and materials, whereas follow-up values were obtained with standardised procedures and materials provided by the sponsor, and the follow-up times were not uniform across patients. Other correlations investigated were not significant.

Regarding change in SUVr over time, there was a numerically greater CFB SUVr for the CI group (0.0347) compared with the HV group (-0.0008); however, the difference in SUVr slope (change across patients as a function of duration of follow up) was not statistically significant (p = 0.2837).

It can be agreed that overall, the results should be interpreted with caution. This was an opportunistic study. The sample size was relatively small (total N = 37 including only 2 Aβ- CI) and the elapsed follow-up duration was variable and short for some subjects, requiring statistical adjustment (slope methods) that may have affected the estimate of between group differences.

The applicant postulated that that elevated flortaucipir PET SUVr signal is associated with the presence of Aβ neuritic plaques and may be implicated in the onset and progression of cognitive impairment. However, this could not be confirmed statistically.

18F-AV-1451-A08 (Study A08)

18F-AV-1451 PET Imaging in the Preclinical, Prodromal and Dementia Phases of Alzheimer's Disease.

This was a Phase 1 open-label, outpatient study that evaluated the imaging characteristics of flortaucipir in the preclinical, prodromal, and dementia phases of AD. The study compared the density and distribution of flortaucipir uptake by visual interpretation and MUBADA SUVr among 86 men and women >60 years of age with objective cognitive impairment (MCI and AD) and without objective cognitive impairment (CN and SMC). In the efficacy population, 86 subjects received a baseline flortaucipir scan and 76 completed a 12-month follow-up scan.

Objectives

The primary objective of this study was to compare the density and distribution of flortaucipir uptake by visual interpretation and standard uptake value ratio [SUVr] among men and women aged over 60

years with objective cognitive impairment (mild cognitive impairment [MCI] and Alzheimer's disease [AD]) and without objective cognitive impairment (cognitively normal older subjects [CN] and subjects with subjective memory complaints [SMC]).

Participants

Number of subjects: Planned: 110 subjects, Enrolled: 89; treated (at least 1 dose): 86; completed: 76 (4 AD, 10 MCI, 41 SMC and 21 CN) had a follow up scan 12 months later.

Treatment

Flortaucipir F 18 (NLT 74 MBq/mL) as an intravenous (IV) bolus administration. No comparator was administered in this study.

Study A08 – Results

Although the sample of objectively impaired subjects (AD/MCI) was small, there was a significantly higher MUBADA SUVr in impaired than non-impaired subjects. No difference was observed between SMC and CN. Flortaucipir PET images were interpreted by visual examination using the criteria described below.

Read Outcome		**Objective Image Features**
Not consistent with AD pattern (τAD-)		No increased neocortical activity, or increased neocortical activity isolated to the mesial temporal, anterolateral temporal, and/or frontal regions.
AD pattern (τAD)	τAD+	In either hemisphere, increased neocortical activity in the posterolateral temporal (PLT) or occipital region(s).
	τAD++	In either hemisphere, increased neocortical activity in the parietal/precuneus region(s), or frontal region(s) with increased uptake in the PLT, parietal, or occipital region(s).

As already noted by the applicant, there was a significantly higher MUBADA SUVr in impaired (MCI and AD) than non-impaired (CN and SMC) subjects. However the number of the objectively impaired subjects (AD and MCI) was very small (n=14) compared to the non-objectively impaired (SMC and CN) (n=76) leading to a MUBADA SUVr mean difference of only 0.16 between these groups.

The proportion of τAD++ subjects was numerically greater in Aβ+ vs Aβ- subjects (42% vs 9%). There was no significant association between τAD++ and Aβ status across cognitive status groups. Similarly the difference in mean SUVr value between Aβ+ and Aβ- was not significant. There was a significant correlation between baseline MUBADA SUVr and MMSE change from baseline (CFB) ($r=-0.613$; $p<0.0001$). Subjects with a visual read status of τAD++ exhibited significantly higher MMSE CFB compared to both τAD+ and τAD- groups ($p<0.05$). However, the results were not consistent since in the case of the CDR-SB the change at 12 months was not statistically significant (0.249, $p=0.0550$).

Risk ratios for change from baseline as a function of visual read were high. However, given the small number of subjects who actually had threshold change within one year, these ratios should be interpreted with caution.

With respect to visual reads and likelihood of clinical progression, there was no significant association of scan pattern indicating likelihood of progression (τAD++ versus non-τAD++) with cognitive status,

overall, or within each amyloid group. We also tested whether a relationship exists between all subjects with an AD pattern (τAD+ and τAD++ versus τAD-) and cognitive status. In this analysis, there was a trend for an association of likelihood of progression (τAD++ versus non-τAD++) with cognitive status, and overall.

In this study, there was a significant difference in the prevalence of τAD+/τAD++ versus τAD- only in the Aβ- group, favouring more prevalent τAD- status. Note that, unlike other clinical trials in the flortaucipir development programme reported to date where the frequency of Aβ- among τAD+/τAD++ subjects is small (typically <5%), 32% of the τAD+/τAD++ group was Aβ-. It may be relevant that the A08 sample is dominated by SMC subjects (11/17 Aβ+ are τAD+/τAD++; 6/8 Aβ- are τAD+/τAD++). Further, the amyloid status in A08 was not determined by the same criteria or with the same tracer as in other trials. These methodological and small sample size issues may have contributed to this unusual finding.

Examination of change in cognition over time showed a significant decline in MMSE after 12 months in the τAD++ group, but no significant decline was detected in the τAD+ or τAD- groups. However, none of the groups worsened significantly from baseline in the CDR-SB after 12 months, nor did the CFB differ between groups. The 12-month follow up period is shorter than most of the other longitudinal studies in the flortaucipir development programme and likely played a role in the limited findings with respect to change in cognition over time. The τAD++ group did have a significantly higher CDR-SB at baseline than the other 2 groups, who did not differ from each other. Moreover, as noted above, this dataset is dominated by SMC subjects, without objective impairment, which may confound the results when compared with other flortaucipir studies where there was higher representation of objectively confirmed cognitive impairment (AD and MCI) subjects.

A comparison of MMSE CFB as a function of cognitive status demonstrated an overall effect of impairment group on LSmean change in MMSE, but no significant effect of cognitive status group on CDR-SB CFB was observed, nor did any single cognitive status group demonstrate a significant CFB for CDR-SB over 12 months. Similarly, risk ratios for progression to clinically meaningful changes for τAD++ versus non-τAD++ were significant for MMSE, but not for CDR-SB.

The relationship between baseline or change in MUBADA SUVr and cognitive change was explored. There was a significant correlation (adjusted for age and baseline cognitive score) between baseline MUBADA SUVr and CFB for MMSE, and there was a trend for CDR-SB.

There was also a significant correlation for MMSE, but not for CDR-SB with MUBADA SUVr CFB. MUBADA SUVr CFB did not differ between Aβ+ and Aβ- groups.

With respect to the prevalence of τAD+/τAD++ versus τAD- only in the Aβ- group, there was a significant difference with other trials. There were methodological (different criteria for amyloid status by local readers and not AVID expert readers) and sample size issues (dominated by SMC) which may have contributed to unusual findings. Hence, the strength of the contribution of these results as supporting evidence is weak.

2.6.6. Discussion on clinical efficacy

Dose finding

No dedicated dose-finding has been done. However, majority of the studies (including the key clinico-pathological pivotal study) utilised the same dosing and acquisition regimen as proposed in the SmPC. The total expected radiation exposure from the recommended dose (effective dose: 9.6 mSv) is

slightly higher than that of the beta-amyloid PET products (i.e., about 7 mSv in a 70 kg person), but is within the acceptable range. Duration of the image acquisition of 20 min is in the range of the general standard times. The total mass dose of the cold product is limited to 20 µg that is in micro-dose range and the maximum allowed volume of the injected product is 10 mL. The proposed dosing and image acquisition regimens appear reasonable and are accepted.

Evidence of efficacy relevant for the claimed indication

The updated indication claim is: "…Flortaucipir (18F) is a radiopharmaceutical indicated for Positron Emission Tomography (PET) imaging of the brain to assess the neocortical distribution of aggregated tau neurofibrillary tangles (NFTs) in adult patients with cognitive impairment who are being evaluated for Alzheimer's disease (AD). Flortaucipir (18F) is an adjunct to clinical and other diagnostic evaluations.

For limitations of use, see sections 4.4 and 5.1."

Key evidence of efficacy for the indication is derived from the two "pivotal" studies (A16 and FR01) evaluating the diagnostic and technical performance of flortaucipir PET. Supportive evidence contains mostly the published evidence on the ability of flortaucipir to detect AD-related tau, on the impact of the PET on diagnosis and diagnostic thinking and own and published studies evaluating flortaucipir uptake in atypical AD, elderly cognitively normal patients, in clinically defined Progressive Supranuclear Palsy (PSP) and corticobasal degeneration (CBD), in subjects with fronto-temporal dementia (FTD), and in subjects at risk for chronic traumatic encephalopathy (CTE). With the responses to the day 120 LoI pooled analysis of rates of positive flortaucipir PET in the patients with AD-related MCI were submitted to support inclusion of MCI population in the indication.

The A16 study was a clinico-pathological study and the FR01 study was a retrospectively planned reader study. As a retrospectively planned study, the FR01 study is not considered a pivotal study. The study A16 has acceptable design and TS and can be considered a pivotal study. Thus, MAA in this indication relies on one pivotal study and the guideline on a single pivotal study in support of marketing authorisation should be followed (CPMP/EMA/2330/99). Further, the EMA "Guideline on Clinical Evaluation of Diagnostic Agents" (CPMP/EWP/1119/98/Rev 1) and Appendix 1 on Imaging Agents (EMEA/CHMP/EWP/321180/2008) also apply.

Of note, initially a broader indication was targeted that, besides the "diagnostic" claim (so called Indication I), also included a "prognostic" claim (so called Indication II). However, CHMP considered the data for a prognostic claim insufficient and the applicant decided to withdraw the Indication II, as well as a number of claims included in the Indication I. For ease of reading, the discussion of the efficacy data related to the Indication II is presented separately.

Design and conduct of clinical studies

Studies in this submission were completed primarily in the US, but also included subjects from Australia, Canada, Japan, France, Poland, Belgium, Germany, and the Netherlands. The readers in the 4 main studies were all from the US. Additional supporting data presented from the biomarker studies came from global clinical studies of AD treatments. The study population is considered to be representative of the intended use population in the EU because, as discussed, the diagnostic criteria and clinical practice currently used in the evaluation of cognitively impaired patients are consistent between the US and the EU. There is also no evidence that European subjects would have different types of NFT and thus different brain uptake, or different flortaucipir binding compared to US subjects.

From experience with amyloid tracers and the literature, there is no expected difference between US and EU readers with respect to accuracy.

The pivotal Study A16

Subjects with and without cognitive impairment and with or without suspected AD were included. Inclusion of patients with suspected AD (and with cognitive impairment) is endorsed. Enrolment of subjects without cognitive impairment is accepted, as these patients provide "true negative" cases, in order to assess specificity. However, this population is not expected to be the primary target for flortaucipir PET in clinical reality. Selection of end-of-life population is acceptable, since no other option of collecting autopsy samples would have been reasonable/feasible. Projected survival duration of 6 month is acceptable, as no relevant AD-related changes in the brain can be expected in this period. Overall, inclusion/exclusion criteria were not suitable to recruit a population fully representative of the target population. However, this is acceptable given the challenges of conducting a study in "autopsy" population.

Ongoing monitoring of recruitment by the sponsor, i.e., by the DMB (monitoring of images) and pathology expert board (monitoring of the number of the patients with and without histologically confirmed AD) triggered concerns due to the potential risks of selection bias and unblinding. Also, 3 protocol amendments with seemingly relevant changes including the definition of the reference truth (B3 instead of B2) was introduced later in the study. The applicant has provided detailed information about the dates of approval/finalisation of various relevant documents (charters, manuals, etc.) and analyses and confirmed that the established committees/expert groups worked independently from the sponsor and under blinded conditions. Also, the decision of the truth standard definition was made earlier than it was implemented. The applicant concluded that the data integrity/reliability could not had been affected. This is accepted.

The chosen dosing regimen was similar to the one recommended in the PI and is thus acceptable.

Use of histology as the universal standard of truth reference is endorsed.

Both study objectives, assessment of diagnostic performance and testing of the reliability of the reads are clinically relevant objectives.

Co-primary endpoints of the primary analyses, sensitivity and specificity, tested against the histology TS are consistent with the requirements of the EMA guideline on diagnostic agents.

Two primary endpoint analyses were conducted comparing the test against the histological proof of NFTs (primary analysis 1) and of the AD-specific neuropathological changes (ADNC) in the brain (primary analysis 2). The primary analysis 1 reflects the ability of flortaucipir to bind to tau and its estimation in the neocortex. This is the pivotal endpoint that is relevant for the claimed indication.

The primary analysis 2 was planned to support the initially raised claim on confirmation of neuropathological change in the Brain. As this claim has now been removed from the indication, the primary analysis 2 is only considered as supportive.

Gold standard criteria/scores (Hyman et al., 2012) for assessment of the AD-NFTs (B-scores/Braak stages) and ADNC (various levels of change) in autopsy were applied to establish TS for the primary analyses (1 and 2). B-scores/Braak stages characterise only the distribution of the AD-type of NFTs in the brain, but does not capture the quantitative aspects. Thus, quantitative component of estimation of NFTs was not reflected in the primary endpoint/confirmatory analysis. This is acceptable since the modified indication does not include the claim that flortaucipir PET can estimate density of NFTs.

Three types of the pattern of flortaucipir uptake in the brain were defined to categorise PET images as "positive" (tAD+ and tAD++ patterns, consistent with AD diagnosis) and "negative" (tAD- pattern, not

consistent with the diagnosis of AD). These patterns considered topography of flortaucipir uptake in the neocortex and were defined as follows: tAD+ (moderate AD pattern on PET) - increased neocortical activity in the PLT or occipital region(s), tAD++ (advanced AD pattern on PET) - increased neocortical activity in the parietal/precuneus region(s), or increased activity in the frontal region(s) accompanied by increases in the PLT, parietal, or occipital region(s), and tAD- - no increased neocortical activity or increased neocortical activity isolated to the mesial temporal, anterolateral temporal, and/or frontal regions.

On autopsy, Braak stages V-VI/B3 were considered as positive histological proof of NFTs in the brain (TS-positive) and Braak stages 0-IV were considered as negative TS.

It is noted that the definitions of "positive/negative" TS and of "positive/negative" PET signal do not fully reflect the neuropathological criteria of presence/absence of AD. E.g., the Braak stages III and IV are termed the "limbic" stage and refer to progression of NFTs to the limbic structures (amygdala and hippocampus), and subsequently the thalamus. Clinically, in more than half of the individuals the spread of NFTs to stages III and IV is associated with at least mild dementia (Hyman et al., 2012) and this stage of NFT aggregation is considered consistent with early AD neuropathological diagnosis. However, in the study the TS and visual read criteria qualified such cases as not consistent with the diagnosis of AD ("negative" cases) and aimed to diagnose the progressed/severe/end-stage cases of AD only. Thus, the chosen visual read criteria cannot inform about a "positive" or "negative" neuropathological diagnosis of AD, but rather the presence or absence of progressed/severe/end-stage AD. This is a limitation. As the same visual read criteria are recommended for image interpretation in the SmPC, the warning that a negative flortaucipir scan does not rule out the diagnosis of AD or the presence of earlier NFT pathology, e.g., B2, has been added in section 4.4 of the SmPC.

Further, the brain regions selected for PET image interpretations (see the list of various regions in the CRF) and those for assessment of neuropathology/autopsy sampling were close but not identical. When looking at the visual read criteria, high-intensity signal in the following areas of neocortex were regarded as positive AD-pattern: posterolateral temporal (PLT), occipital, or parietal/precuneus region(s), with or without frontal region(s) activity. Histology samples were collected in the following areas of the brain, which seem to overlap with the visual read regions: middle and superior temporal gyrus, middle frontal gyrus, inferior parietal lobule, peristriate occipital cortex. When comparing these, it seems that the regions supportive of positive AD pattern in the visual read cover the histological locations for occipital, parietal and frontal regions. Visual read results for the temporal region might not have been correctly matched with histological findings. Clinical relevance of this finding cannot be estimated in the absence of detailed information. It may be that the presented diagnostic performance tested against B3 is not fully accurate.

Sample size was justified with an expected lower confidence limit but not with a probability of 80% or 90% of that limit being larger than 0.5, nor did it correspond to the success criterion of at least 3 readers having a lower confidence limit being larger than 0.5. In addition, between reader variability was ignored. In that sense a sample size of 14 cases per TS group would have been much too small. Finally, more subjects were included in the final analysis, potentially due the number of subjects for whom autopsies were ongoing. However, any result driven stopping the study should be excluded. Otherwise, the procedure could have resulted in a "camouflaged" sequential procedure without proper adjustment. With the response to the Day 120 LoQ the applicant described the actually applied procedure with respect to study stop and provided evidence that study stopping was unlikely to be informed by study results.

As exploratory endpoints inter-reader agreement and testing of the quantitative measurement vs TS were done. Both are regarded relevant endpoints.

The number of readers (5 in total) and the reading, blinding and data handling procedures targeted to keeping blind seem overall acceptable. However, recruitment and patient data (PET images and histology outcomes) were monitored by a dedicated expert groups with the possibility to report back to the sponsor. Also, PET reading procedure/methodology was completed during the study. These factors raise question on how well the blinding and data integrity could have been kept intact. Nevertheless, the applicant provided the detailed information about the dates of approval/finalisation of various relevant documents (charters, manuals, etc.) and analyses and confirms that the established committees/expert groups worked independently from the sponsor and under blinded conditions. Only the data of 3 front-runner patients were shared. Overall, the presented argumentation and the conclusion that the data integrity should not have been affected, appears acceptable.

The success criterion (threshold of 50% for lower limit of 95% CI) defined by the applicant appears rather weak. Considering, e.g., a real-world scenario in which the diagnostic test would result in a sensitivity or specificity of less than 50% for half the readers, the chances of obtaining a result corresponding to success in a study with 5 specific readers could be rather large, depending on the between-reader variability. With other words, study success could be compatible with a scenario in which for a randomly selected reader the probability of insufficient diagnostic properties would be 50%. These considerations depend on the heterogeneity between readers. Actually, in the current study an important heterogeneity was observed with respect to specificity. In that sense the success criterion appears insufficient to guarantee with a large probability that for any future reader both sensitivity and specificity would be larger than 50%. Hence, for a useful diagnostic procedure it should rather be expected that in all readers out of 5 readers both, sensitivity and specificity would be larger than 50%. The chances of study success could be large in scenarios of larger between-reader heterogeneity where an important part of the readers would fail to obtain sufficient diagnostic properties. Overall, the criterion still appears to be weak allowing for an authorisation of a diagnostic procedure where a randomly selected reader may fail to provide reliable diagnostic results with a large probability. Hence, study results will be assessed independently of the success criterion considered to be too weak.

Regarding the secondary analysis, a majority read remains difficult to interpret for real life application, since the corresponding procedure does usually not correspond to clinical practice.

No procedure was defined in case of missing PET scans, i.e., these subjects were excluded from the analysis, implicitly assuming a completely-missing-at-random mechanism. It remains doubtful that the missingness can be considered as independent of the readability of the scan, especially in case a scan was rejected due to failed QC. Apparently, two subjects with an autopsy were excluded due to missing or failed scan. Nevertheless, the applicant provided additional sensitivity analyses, were these 2 subjects are replaced as differential worst cases (i.e. as a positive test result for true negatives and as a negative test result for true positives) which were considered to be conservative and resulted in small reductions in sensitivity and specificity of 1.8% and 1.9% only.

The methodology applied for quantitative measurements appears appropriate. The presence and density of NFT/NT (categories: none, rare, sparse, moderate, frequent, not assessed) and presence of atypical tau pathology (categories: present, absent, not applicable) was evaluated in the study in accordance with the study documentation. However, quantitative measurement parameters (e.g., SUVr) were not compared vs. density/quantity of NFTs/AD-typical tau on autopsy in the ITT population. This is acceptable, since the claim that flortaucipir PET can estimate NFT density in the brain has been removed from the indication wording. A published study was submitted by the applicant, which described correlation of SUVrs on Tauvid-PET vs. density of tau in autopsy in the 3 front-runner patients from the A16 study. This and further publications are discussed below.

To summarise, the study is acceptable in terms of applied methodology, although, the detected limitations and weaknesses need to be considered when interpreting the results.

The Study FR01

The FR01 study was a retrospective study that aimed 1. to replicate the primary results of the A16 study in a larger set of autopsy population including the front-runner and supplemental cases (which were excluded from the primary endpoint analysis in the A16 study) and 2. to test inter- and intra-reader agreement in the population resembling the targeted patient population, including patients with MCI from the A5C study. The study also tested efficacy of more stringent definition of positive PET, i.e., tAD++ in terms of its ability to accurately diagnose AD cases compared to autopsy, and in terms of its reliability as assessed by means of inter- and intra-reader agreement. The latter analyses were done in the patients with MCI separately (target population in which no autopsy data are available) and pooled with the patients from the A16 study (autopsy population). The aim of this analysis was to compare technical performance (inter-reader agreement) of flortaucipir (18F) PET in the MCI and mixed MCI plus autopsy population. The applicant argues that similar technical performance of the test across populations can substantiate extrapolation of the A16 study data to the target indication/broad population. This argumentation is not supported. Assessment of technical performance (i.e., inter-reader variability) alone in the MCI population is insufficient to substantiate the broad indication claim and similar technical performance across populations does not translate into similar diagnostic performance, so that data in diagnostic performance are needed (or justification on their extrapolation) and data on inter-reader agreement alone are insufficient.

Overall, the study design and procedures are acceptable. The hypotheses to be tested defining study success for reader agreement, referring to a Kappa of 0.6 appears weak. The deficiencies/limitations identified in the A16 study also apply to this study, as long as the same procedures are concerned.

As mentioned above, the study and the objectives are regarded exploratory and as partially supportive, but not confirmatory.

Altmore et al., 2021

This was a large, appropriately designed, independently-conducted, counter-balanced head-to-head crossover study comparing impact of flortaucipir PET and amyloid PET on diagnostic thinking. The study design is considered appropriate.

The applicant's studies and published studies on the correlation of flortaucipir (18F) signal vs. autopsy and on flortaucipir uptake in other tauopathies, traumatic encephalitis, older cognitively normal patients, atypical AD, etc.

Multiple studies presented are exploratory studies utilising quantitative measurements with variable methodologies to evaluate the presence and correlation of flortaucipir signal in PET mostly without autopsy as a reference standard. Also, flortaucipir uptake in atypical AD and multiple non-AD conditions has been studied utilising quantitative measurements. These studies have limited supportive value, as quantitative measurements are not proposed for flortaucipir PET interpretation. Also, the methodologies applied for quantitative measurements varied and none of these have been validated thus far. Additionally, the applicant has removed the claim that flortaucipir PET can estimate NFT density from the indication wording.

Efficacy data and additional analyses relevant for the claimed indication

Study A16

The study was conducted outside Europe, in the USA and Australia. There is no scientific reason (different pathophysiology of the disease, major differences in the diagnostics/patient characteristics, patient management at the time of the study conduct, etc.) to believe that the data collected in the US/Australian population cannot be extrapolated to the European one. The involved sites were mostly specialised clinics. This is acceptable. Recruitment period was quite long, but appears reasonable for the chosen study design.

The size of the recruited population (n=64 primary analysis) is relatively small, but is in the same range that was already accepted for beta-amyloid PET products.

Majority of the patients were elderly and with dementia (49 of 64 subjects, i.e., 85% in the EFF1) and showing advanced stage of NFT aggregation in the brain (42, i.e., 65.6% with Braak stage V-VI). The patients with clinical manifestation of mild cognitive impairment (MCI) were practically not represented (only 1 patient). However, there were 17 subjects with the Braak stages III and IV (B2), corresponding to intermittent stage of AD in histology. Absence of data in the MCI population is a limitation and the study population is not fully representative of the targeted one. However, the challenges of conduct of an autopsy study, especially in subjects with MCI, are acknowledged and these limitations/deficiencies have been mentioned in the SmPC.

Notably, only part of the patients had baseline levels of cognition and functionality evaluated, which is a limitation. The MMSE and IQCODE scores (where collected) indicate that the patients with dementia were in progressed/severe stage of their condition, while there seems to be hardly any difference between the CN and the single MCI subjects in the cognitive/functioning status.

Of note, part of the study population (16 of 49 subjects with dementia) had non-AD type of dementia and part of the subjects (n=29) had atypical tau on autopsy. LBD, vascular changes, Parkinson's disease (reported in the results further), frontotemporal dementia, mixed form of AD and "other" (not specified) were the non-AD diagnoses reported. It is unclear on which evidence these diagnoses were based and why the numbers of the patients with non-AD diagnoses and non-AD tau do not match. It is assumed, that the difference is related to the additional evidence gained through autopsy. Also, the status of cognition in the patients with the intermediate stage of NFT (B2) on autopsy is not clear.

The lack of information on the status of cognition at baseline is considered less relevant as the main aim of this study was to evaluate the ability of flortaucipir PET to accurately reflect histological findings/progressed Braak stages. The coexistence of other diseases/conditions that may contribute to dementia in people with Alzheimer's disease is well known.

The study met its predefined primary endpoints for both analyses (against NFT and ADNC). Good level of sensitivity was achieved with somewhat lower and variable levels of specificity for both analyses.

For the first analysis (vs. NFT) sensitivity (95% CI) ranged between 92.3% (79.7, 97.3) and 100% (91.0, 100) with lower limits of CIs above 79.7 in all readers. Specificities were lower and more variable (95% CI) ranging from 52% (33.5, 70.0) to 92% (75.0, 97.8). Only 3 of 5 readers passed the 50% success threshold pre-defined for the lower limit of 95% CI and lower limits of 95% CI for 3 of 5 readers were below 60% (the threshold that would have been more acceptable as minimally clinically relevant), which indicates the test's difficulty in correct diagnosis of absence of the disease. This means that the test was more likely to produce false positive results (risk of overdiagnosis), and the negative PET scans were more likely to describe true condition (absence of AD). This was explained in the SmPC (section 4.4).

In the second primary analysis (vs ADNC) the data were very similar with sensitivity (95% CI) ranging between 94.7% (82.7, 98.5) and 100% (90.8, 100.0) with lower limits of CIs above 82.7 and lower and variable levels of specificities (95% CI) ranging from 50% (32.1, 67.9) to 92% (75.9, 97.9) and lower limits of 95% CI for 3 of 5 readers below 60%.

With the responses to the Day 180 LoOI the applicant provided post hoc analysis of the likelihood ratios including 95% CIs. For the primary analysis 1, data showed that the LR+, which was true positive rate divided by the false-positive rate (sens/(1-spec)), ranged from 2.08 (95% CI: 1.46, 3.8) to 11.54 (95% CI: 3.71, 152) across the 5 readers. The LR-, which was the false-negative rate divided by the true negative rate, ranged from 0 (95% CI: 0, 0.17) to 0.1 (95% CI: 0.02, 0.27) across the involved 5 readers. There is no universally accepted interpretation of what level of LR can be considered relevant. However, since the likelihood ratios represent the factor by which the odds of being diseased (or not diseased) change after obtaining the test result, it is generally accepted that the further away the LR values are from "1" (LR+ > than 1; LR- < than 1), the more informative (more likely to show true finding) the test is. In general, a small LR+ (or a large LR -) would mean that the amount of information obtained by the test result is limited.

To summarise, in terms of diagnostic performance, this was formally a positive study with some uncertainties concerning the test's ability to exclude AD in its absence, i.e. to correctly identify AD-negative patients as being negative, as part of these patients may be erroneously diagnosed as AD positive (i.e., risk of over-diagnosis). Again, the data were collected in a very restricted population and generalisability of the data to broader population especially to MCI is limited.

Inter-reader agreement has reached acceptable levels when overall agreement was evaluated. For comparisons across individual readers the range of agreement was quite broad. Obviously, specificity varied considerably between readers. Comparing, e.g. readers 2 and 5, reader 5 had higher rate of false positive findings (by 40%) compared to reader 2. Apparently, reader 5 was much more inclined to interpret a scan as positive, than reader 2, resulting in a somewhat higher sensitivity at the cost of a much lower specificity. A warning that flortaucipir should be interpreted only by a trained reader has been added in the SmPC.

The studied subgroups by age and gender are too small and/or variable in size to allow for proper comparison across the subgroups and draw firm conclusions. However, the applicant speculates that decreased diagnostic performance in the elderly population of the age above 75 may be due to the changes in the brain, which take place due to AD. While flortaucipir uptake increases with age in cognitively normal/non-AD population, it decreases in the patients with AD. The study data seem to be in line with this observation.

Subgroup analysis conducted in the subjects with dementia showed roughly similar sensitivity as the primary analysis population, but with worsened specificity. These data are practically not informative given the small size of the subgroups and the fact that the outcomes of main analysis are dominated by the population with dementia.

The conducted quantitative analysis showed that SUVr 1.126 was the optimal cut-off for definition of increased uptake of flortaucipir (18F). The absolute majority of the amyloid and tAD++ subjects had SUVr above this cut-off and were distributed across Braak V and VI stages on histopathology. In contrast to this, the patients with tAD+ and amyloid had SUVrs below this cut-off value. Obviously, only few of the patients with tAD+ were included in the study, which does not allow to draw firm conclusions. However, it is obvious, that tau distribution as described by the PET (i.e., tAD+, tAD++) did not correlate fully with the distribution described in histology (Braak V, Braak VI). It is expected that tAD+ should reflect Braak V and tAD++ Braak VI stage. The applicant was requested to discuss the correlation of PET findings with histology, but did not provide such analysis. Since quantitative analysis of Tauvid PET is not suggested in the SmPC, and the submitted data are limited, the claim that flortaucipir PET evaluates density in the Section 4.1 of the SmPC was not supported and the applicant has removed this claim.

The conducted assessment of diagnostic performance in the population including supplemental data set is considered not supportive, as the analysis was not defined in the protocol these data are regarded to have limited supportive value.

Sensitivity analysis of the primary endpoints utilising more stringent visual read criterion showed improved specificity, but this analysis is not considered very relevant, as it restricts the diagnosis to even more limited subpopulation.

FR01 study

This was a reader study. Assessment of diagnostic performance was done in the subgroup of population from the A16 study ("autopsy") and inter- and intra-reader was evaluated in the broader population more representative of the target indication (studies A16 and A05C together or A05C population separately).

Assessment of diagnostic performance showed similar or slightly improved diagnostic performance of flortaucipir PET for the same endpoints as in the A16 study. Slight improvement may be explained by the reader training, that appears to have been more specific/better in this study and partly by the modified patient population.

Inter-reader agreement also showed roughly similar results as the study A16, but intra-reader agreement is suboptimal. Inter-reader agreement in the populations including only patients with MCI and mixed MCI plus cognitively normal plus patients with dementia were quite similar and the applicant concludes that this can be used as proof of similar efficacy of the test across various populations. As already mentioned above, this argument is not supported.

Impact on patient management and on diagnostic confidence

Collection of data on the impact on patient management in difficult given that there is no approved therapy available at this point in time. Impact on the diagnosis/diagnostic confidence can be derived from recommendations of the various expert groups (Dubois et al. 2021; the recently presented update from the NIA-AA workgroup (NIA-AA Revised Clinical Criteria from 2023).

Additional data supporting the claim that flortaucipir PET can estimate density of NFTs

To support the claim that flortaucipir PET can estimate density of AD-related NFTs, data from the 3 front-runner patients from the A16 study (published) have been summarised and reference to few further autopsy-based published studies was included in the dossier. Front-runner patient data: The evaluation showed very good correlation between the PET signal (SUVr) and density of tau in the autopsy in the neocortical area. Limbic area (limbic (hippocampal/entorhinal and amygdala) (that is speculated to be associated with the intermediate stage of tau/NFT distribution in the AD) showed lower correlation and the remaining regions in the brain showed no correlation at all. The applicant speculates, that it is not clear whether lower correlation of flortaucipir SUVr and the histelide estimate of tau tissue concentration in the limbic area was due to PET signal loss subsequent to atrophy and the resultant partial volume effects on the PET scans (Johnson et al. 2016), or whether this reflects a lower affinity of flortaucipir for the types of tau found in medial temporal lobe of patients with AD (Lowe et al. 2016). Absence of correlation between phosphorylated tau and flortaucipir SUVr in subcortical regions including ventromedial globus pallidus and caudate, putamen and thalamus is expected considering data in the literature and the flortaucipir development studies (Study A05, Brier et al. 2016; Shcherbinin et al. 2016; Pontecorvo et al. 2017). These regions show apparent off-target flortaucipir signal in both Aβ+ AD patients and Aβ- CN elderly. The applicant states that the differing kinetics of flortaucipir PET in these regions (Shcherbinin et al. 2016) suggests the presence of a different lower-affinity, high-density binding site that is probably not related to NFTs. The applicant states, that this off-target uptake is not relevant to flortaucipir PET visual interpretation. Overall, the

totality of submitted evidence suggests that flortaucipir PET is likely able to estimate density of NFTs in a general sense. However, the SmPC of Tauvid does not provide any instructions how to measure density and does not recommend estimation of the density to support image interpretation. Thus, density is not anchored as an additional parameter for image interpretation. Moreover, there is no established/validated method (including thresholds) of quantitative assessments available nowadays, that could be recommended. Consequently, the claim on "density" remains a general statement and was consequently removed from section 4.1. This is agreed.

Data supporting the claim that flortaucipir PET can confirm the presence of AD neuropathology

The claim that Tauvid PET can confirm the presence of AD neuropathology (ADNC) is not substantiated. This claim implies that positive Tauvid PET automatically confirms pronounced aggregation of beta-amyloid in the brain, and could thus replace beta-amyloid PET. This is not agreed. Although, the A16 study indeed showed association between the presence of beta-amyloid, of NFTs and of dementia, the study population was not fully representative of the target population. It has not been demonstrated that similar levels of correlation may be achieved in the target population as in the study. Therefore, the respective claim in the indication wording was removed. This is endorsed. Notably, a number of statements/data related to this claim are still included in the SmPC. These should be removed to avoid confusion and because these data may be perceived as promotional. (OC; reference is made to the annotated SmPC)

Generalisability of data to target population and diagnostic value of the proposed visual read criteria

Study population tested in the pivotal autopsy study was not fully representative of the target indication including also patients with MCI. The applicant has presented a post hoc exploratory analysis in the patients with MCI that shows that the amyloid-positive patients with MCI had positive flortaucipir PET in about 75% of cases, arguing that amyloid-positive MCI patients may have NFT aggregation that has progressed to neocortex and that flortaucipir PET may contribute to the diagnostic work-up in this population. Major limitation of these data is that the accuracy of this evaluation cannot be assessed because of the absence of standard of truth/composite standard of truth. Published literature indicates that clinical diagnosis of MCI and changes in the neocortex do not fully coincide (revised NIA-AA 2023) and that some patients with advanced stage of NFT aggregation may have only mild changes in cognition (Hyman et al., 2012).

Since flortaucipir PET reflects AD-related neuropathological changes in the brain, i.e., aggregation of NFTs in the neocortex, it seems plausible that flortaucipir PET can detect NFTs regardless of clinical symptoms, as long as NFTs are located in the relevant areas of neocortex considered in visual read criteria. However, accuracy of the test in MCI and may be lower. To address these uncertainties, respective warnings have been added in the SmPC section 4.4. Additionally, a cross-reference to sections 4.4 and 5.1 of the SmPC was included in section 4.1 to point out the limitations of data extrapolation.

Use of flortaucipir PET is only recommended in conjunction with other diagnostic methods (see section 4.1 of the SmPC). Considering the above, the broad indication (the claim "adults with cognitive impairment" that also includes patients with MCI) is acceptable, provided that wording adaptations to the related warnings in the SmPC are accepted (see the annotated document).

The applicant has confirmed that the recommended visual read criteria do not identify earlier Braak stages (III-IV), but only the progressed stage of NFT aggregation in the neocortex (Braak V and VI). This has been emphasised in the SmPC section 4.4. This is agreed.

Further evidence relevant for efficacy and labelling

Extensive scientific evidence (the applicant's studies and published studies) shows that quantitative measurement of flortaucipir signal may help to differentiate AD-NFTs against non-AD conditions. Also, the data showed specific patterns of uptake in the patients with atypical AD, as well as uptake of flortaucipir (mostly mild signal intensity) in the patients with Down syndrome, in elderly cognitively normal, in non-AD tauopathies, etc.

The applicant argues that uptake of flortaucipir by non-AD tau vs. AD typical tau can be distinguished by means of quantitative measurements and/or signal topography. This may be agreed. However, no concrete recommendations on the use of quantitative measurements are included in the SmPC and the applicant explains that none of the methods tested is validated and that the ongoing work should lead to the development of a validated methodology after approval. This is accepted.

Flortaucipir binding has been reported in non-AD tauopathies and, as indicated by the applicant, this could be misinterpreted as AD pattern (although rarely). *In vivo* tau PET binding has been reported in patients with progressive supranuclear palsy (PSP) (Whitwell et al., 2017, Passamonti et al., 2017), those with corticobasal degeneration (CBD) (Josephs et al., 2016, McMillan et al., 2016) and microtubule associated protein tau (MAPT) mutation carriers (Jones et al., 2016 and 2018, Smith et al., 2016, Spina et al., 2017). Off-target uptake has been reported also in meningiomas (Chen et al., 2023). Appropriate warning regarding possible risks of erroneous image interpretation has been included in the SmPC.

The type of the tauopathy common for AD (3R+4R) is not specific for AD, but is also observed in Primary age-related tauopathy and chronic traumatic encephalopathy (Zang et al., 2023). Considering the above, the statement "… (NFT) of AD…" was removed from the indication wording as requested.

Cross section pooled analysis of 4 clinical studies showed some association between the presence of beta-amyloid and tau identified by the respective PET imaging and level of cognitive impairment and an association between the absence of amyloid and negative flortaucipir PET in the MCI population. Concretely, flortaucipir PET was positive only in 2% of the amyloid-negative patients with MCI and this may not even suggest AD profile. Therefore, flortaucipir PET is not informative in amyloid-negative patients with MCI, and in these patients conduct of flortaucipir PET is not recommended. The warning that flortaucipir PET is not informative in the patients with negative beta-amyloid PET is appropriately reflected in the SmPC.

Further, distribution of the flortaucipir signal is to be analysed within the neocortex. Signal analysis outside of the recommended areas poses the risk of misinterpretation due to the off-target uptake of flortaucipir (see below). Consequently, general statement that flortaucipir can estimate distribution of NFTs in the brain was not supported and the indication was adapted to specify the target region assessed by flortaucipir PET, i.e., the neocortex. This is agreed.

Indication II (No longer claimed): "… and to inform risk for progression of cognitive and functional impairment due to AD."

This indication is no longer claimed so the MO and other concerns related to the studies submitted for this indication were no longer pursued. The description of these studies is kept for information.

For the second part of the indication the applicant has submitted two pivotal studies (A05C and PX01) and five supportive studies (A05E, A18, TZAX, A04 and A08).

Pivotal studies

Study A05C was the confirmatory phase of the open-label, multicentre study A05, evaluating the safety and imaging characteristics of flortaucipir PET in cognitively healthy volunteers, subjects with MCI, and subjects with AD. The confirmatory phase (A05C) was informed by the results of the exploratory phase (A05E) but included only cognitively impaired subjects with a neurodegenerative cause defined by an MMSE score of ≥20 and ≤27. The primary objective was to confirm the relationship between flortaucipir uptake in the brain as measured by positron emission tomography (PET) and the subsequent rate of cognitive decline observed over longitudinal follow up of 18 months. The secondary objective of this Confirmatory Phase (A05C) was to assess the diagnostic performance of baseline tau positivity according to a flortaucipir scan visual interpretation, for predicting subjects' clinically meaningful cognitive and functional deterioration within 18 months of scan, as measured by the CDR-SB scales.

Flortaucipir F 18 (18F-AV-1451), 370 MBq [10 mCi] and Florbetapir F 18 370 MBq [10 mCi] were administered as an intravenous (IV) bolus injections. No comparator was administered in this study.

Study PX01 is a retrospective analysis of the PET substudy from the AZES therapeutic study and similarly to study A05C evaluated the relationship between baseline flortaucipir PET signal and cognitive change in a follow-up of 205 subjects up to 2 years. Approximately 90 of these subjects completed a CDR assessment at the 24-month visit. However, the primary analysis was conducted at 18 months to be comparable with the analysis of study A05C. No new subjects were recruited and no drug was administered in this study. The scans for all 205 subjects (AD 141 and MCI 64) were evaluated. At each flortaucipir imaging visit in parent study AZES, all subjects received a single IV bolus administration of approximately 240 MBq (6.5 mCi) of flortaucipir F 18 Injection.

The applicant clarified that a total of 68 subjects received 20 mg of lanabecestat (20 subjects with MCI and 48 subjects with AD), 64 subjects received 50 mg of lanabecestat (22 subjects with MCI and 42 subjects with AD), and 73 subjects received placebo. The applicant indicated that the primary analysis for Study PX01 included a lanabecestat treatment group as a covariate, showed as a non-statistically significant factor.

Randomisation did not occur in study PX01 either and the readings were also blinded as in A05C. It is unclear on how the GCP could be assured as the data for PX01 study were collected as part of another terminated prematurely study. The applicant confirmed that studies were conducted in accordance with ethical principles that have their origin in the Declaration of Helsinki and are consistent with the ICH GCP applicable regulatory requirements. Further, the applicant indicated that for the AZES study 19 site audits and 7 Contract Research Organisation/Laboratory audits were conducted by the sponsor. No further information regarding auditing or GCP inspection was provided. As the applicant agreed to the withdrawal of the prognosis claim from further consideration in the present procedure some issues are not further pursued.

In addition to design differences from study A05C, the statistical test for the primary analysis in Study PX01 evaluated a risk ratio, rather than a hazard ratio, which was used in Study A05C.

Both studies, A05C and PX01, attempted to test as primary hypothesis the relationship between a specific pattern of flortaucipir PET retention that is an advanced (τAD++) flortaucipir PET AD pattern, and risk for near-term (18-month) cognitive deterioration and functional progression. There were no deviations that could have had an impact on the outcome of the studies.

A number of secondary and exploratory analyses have been performed in both pivotal studies.

Methodological issue for both pivotal studies

In study A05C a hazard ratio was used to compare subjects with flortaucipir scan rated as тAD++ versus subjects with scans rated as non-тAD++ in progressing to a clinically meaningful event (CDR-SB with 1 point or more increase) within 18 months of scan. Similarly (but not identically) in study PX01 a risk ratio was used to compare тAD++ vs. non-тAD++ subjects in progression to clinically meaningful deterioration evaluated by CDR-SB at 18 months.

Likelihood ratios are considered to be more appropriate compared to the ratios (hazard and risk ratio) used by the applicant, since likelihood ratios would be independent of the prevalence (hence severity of the disease) under the assumption that sensitivity and specificity can be transferred. A likelihood ratio for a positive outcome can indicate the factor by which the odds of worsening is increased after obtaining a positive test result. In any case, the difference between likelihood ratios (LR+) and risk ratios is small. In both studies, the amount of information given by the test appears rather weak and the test cannot be considered really informative.

There are a number of methodological limitations of PX01 study including:

- this is retrospective analysis of the data obtained in the sub-study of the failed AZES study which investigates the treatment of lanabecestat in patients with AD.

- PX01 study only included the data from patients who completed this sub-study study and does not include drop-outs.

- the dose used in this study was lower as compared to the proposed authorisation dose

- 4 of the 5 readers in this study had previously been trained on the read method as part of Study A05C therefore it is not clear if these readers could be considered as independent readers in the context of the current study.

- the study population of AZES study and subsequently in the PX01 study only included patients who were amyloid positive (which is a narrower population as compared to the proposed target population)

Taking into consideration these methodological limitations, it is considered that study PX01 does not fulfil robustness conditions for a confirmatory clinical trial and therefore it can only be considered as a supporting trial.

Supportive studies

Study 18F-AV-1451-A05E was the Exploratory/hypothesis generating phase 2 of the Phase 2/3 cross-sectional and longitudinal, open-label, multicentre observational study A05. The primary objective was to compare flortaucipir imaging results among subjects with Alzheimer's disease (AD), mild cognitive impairment (MCI) and cognitively healthy older individuals.

The primary objective of the Exploratory Phase longitudinal component as specified in the protocol was to assess the rate of change of tau deposition as measured by flortaucipir uptake over time. An additional objective was to explore associations between baseline flortaucipir retention with clinical and functional measures, in order to generate hypotheses regarding the predictive relationships between flortaucipir and cognitive decline that could be tested in the Confirmatory Phase of the study.

Study A18 was a Phase 2, open-label, outpatient study that evaluated longitudinal change of tau deposition in the subjects who participated in the Confirmatory phase of study A05, as a follow-up. This study has been characterised as a safety study for the follow-up of the participants from study A05C but it is rather a longitudinal follow-up of change in flortaucipir PET in Study A05C subjects.

The statistical analysis of study A18 is acceptable, in general. However, no multiplicity adjustment for the different comparisons was planned, resulting in an explorative assessment also due to the lack of a control group. In this supportive study as opposed to the pivotal studies, scans were evaluated quantitatively (standard uptake value ratio [SUVr]). Therefore, this study cannot contribute to the proposed indications because as stated in the SmPC for this indication visual and not quantitative assessment is recommended. Further, quantitative assessment was not investigated as a primary endpoint in any pivotal study.

Similar to Study PX01, **study TZAX** was designed as a tau imaging substudy to take advantage of a therapeutic trial cohort that had flortaucipir PET scans and up to 18-month cognitive follow-up to further evaluate the relationship between flortaucipir PET signal and cognitive and functional decline. The parent study LZAX or Engage 3 was a multicentre, randomised, double-blinded, placebo-controlled, Phase 3 study comparing 400-mg solanezumab with placebo given as infusion once every 4 weeks over 76 weeks in approximately 2100 outpatients with mild AD.

Study A04 was a Phase 2, open-label multicentre opportunistic study, designed to provide an early preliminary estimate of longitudinal change in flortaucipir PET signal. Subjects (cognitively impaired (CI) and cognitively normal healthy volunteer (HV) male and female subjects ≥50 years of age) who had previously received a flortaucipir PET scan as part of either Studies A01/T807000, A03, or an independent investigator-sponsored study, were recruited to return for a follow-up flortaucipir scan and cognitive and functional testing from 0.59 and 2.34 years after the initial scan. From the 44 subjects enrolled, 37 completed the study.

Study A08 was a Phase 1 open-label, outpatient study that compared the density and distribution of flortaucipir uptake by visual interpretation and MUBADA SUVr among 86 men and women >60 years of age with objective cognitive impairment (MCI and AD) and without objective cognitive impairment (CN and SMC). In the efficacy population, 86 subjects received a baseline flortaucipir scan and 76 subjects completed a 12-month follow-up scan.

Indication II - Results:

A total of 160 subjects (AD, n = 62 and MCI, n = 98) were enrolled in the pivotal study A05C, 159 were scanned and 111 completed the study. At baseline, 78 patients were τAD++, 11 were τAD+, and 70 were τAD- by majority read. From these 159 subjects 101 were Aβ+ and 58 were Aβ-.

A large number of patients were terminated from the study in the AD group (27/62) and in the MCI group (22/98) resulting in a percentage of 44% and 22% for the AD and MCI group, respectively and with the total number being as high as 49/111 (44%).

In light of the relatively small number of patients included in the confirmatory phase compared to the potential patients with AD and MCI and the large number of discontinuations, the size of the study is of concern for the support of an indication to inform risk for progression of cognitive and functional impairment due to AD.

In the pivotal study A05C, five imaging physicians, blinded to all clinical information, independently visually interpreted the flortaucipir PET scans as 1 of 3 read patterns as not consistent with an AD pattern (τAD-), consistent with a Moderate AD pattern (τAD+), or consistent with an Advanced AD pattern (τAD++).

The hazard ratio for the primary analysis in A05C was 1.581 of at least a 1-point worsening on the CDR-SB with a lower bound of the 95% confidence limit of 0.968. The chosen CDR-SB change of 1 point can be considered a potential measure of response in therapeutic trials. However, the primary analysis did not meet the pre-specified success criteria of lower 95% confidence limits >1.0. The hazard ratios in other indices including CDR-SB change > 2.5, CDR Global stage change (change

greater than 0), MMSE (≥3 point decrease), and Functional Activities Questionnaire (FAQ, ≥3 point increase) were nominally statistically significant apart from the ADAS-Cog (≥4-point increase). Since the primary analysis failed these results do not render the study successful.

It should be pointed out that not only was the secondary objective of study A05C not formally met due to the failed primary analysis., but the method also had performance limitations as the sensitivity did not reach the 50% success criteria for any of the readers (please see Table 37 (from Table A05.11.18. of the Body of the Clinical Study Report). The relative risk ratios ranged from 1.391 to 1.708, and the upper limits of the 95% CI of the relative risk ratios were all above 2.0 ranging from 2.055 to 2.568. Considering the lower limit of the classical (asymptotic) 95% confidence interval, the values for all readers are close to 1 and for reader 5 even lower than 1. In case the argument of the applicant were that some of the readers show a LR+ larger than 1 (but not all), a multiplicity issue would arise. Adjusted confidence intervals for the likelihood ratio (see above) would all include 1. The applicant does not intend to make claims regarding the diagnostic performance of flortaucipir for prediction of progression for individual patients and they did not discuss these results. However, this is important information if a prognostic claim is made in the indication.

It is noted that sensitivity is lower than specificity, with the latter having a mean value of only 65.5% for every reader. For all 5 readers the lower confidence limits of specificity only just exceeded the target of 50% (52.7%), but the lower confidence limits for sensitivity did not exceed 50% for any of the readers, ranging from 36.9% to 46.1% (please see Table 37 (from Table A05.11.18. of the Body of the Clinical Study Report). The confirmatory phase of the study A05 can be considered as not being able as a pivotal trial to support a claim to inform risk for progression of cognitive and functional impairment due to AD.

Study PX01 was designed testing a similar hypothesis as in Study A05C. In study PX01, the primary analysis showed a statistically significant risk ratio of 1.36 (p=0.0313) for subjects in the τAD++ group as compared to those in the non-τAD++ group (τAD- and τAD+) in terms of progression to clinically meaningful deterioration (CMD) within 18 months. The effect size RR 1.355 (1.028, 1.785) is very small and therefore it is unlikely to be clinically relevant.

A basic limitation of study PX01 was that it concerned a retrospective analysis of a population (N=205) that had a higher baseline risk for AD in comparison to study A05C, with the exclusion of drop-outs, probably creating a selection bias and leading to results which would not be generalizable across a broader population or representative of the population proposed in the indication. The applicant clarified that the main reason for studies A05C and PX01 being considered as pivotal is that the reads for Studies A05C and PX01 were performed by independent external physicians trained on the proposed visual read method for clinical use, blind to all patient data other than the PET scan. In other studies (i.e. A05E and TZAX), the cases were subsequently read blind to clinical information by Lilly readers but because the results from some cases were used in designing methods and because the reads were performed by Lilly staff rather than independent experts, these studies were considered supportive rather than pivotal. As the applicant agreed to withdrawal of the prognosis claim from further consideration in the present procedure, this issue is not further pursued.

Flortaucipir F 18 in study PX01 was administered at a different concentration as a 240 MBq from the recommended single intravenous dose for flortaucipir (18F) injection of 370 MBq (10 mCi) for the intended use of Tauvid. The lower dose in the parent lanabecestat study I8D-MC-AZES (AZES) of study PX01 was used to minimise the toral radiation exposure to study participants. To compensate for the lower radioactive dose, these studies typically used flortaucipir PET scan acquisition times longer than the recommended 20 minutes to obtain sufficient image counts, with the timing of the PET scan centred around the same mean time post-injection (that is, 90 minutes). Since the proposed dosing scheme is 370 MBq with 20-minute acquisition, the results with a different dosing (240 MBq) and

different scanning schedule (approximately 90 minutes) could have been only supportive for a prognostic claim, in case they were successful.

With the secondary analysis in PX01 for the risk of progression to clinically meaningful deterioration (CMD) at 18 months, only ADAS-Cog11 change ≥4 reached statistical significance. For the other endpoints/indices MMSE ≤-3 point change, FAQ ≥3 point change and CDR global >0 point -change the values did not achieve statistical significance. Thus, the results do not show consistency across all endpoints/indices.

In study PX01, during the exploratory analyses for the Risk Ratio of τAD++ vs. Non-τAD++ subjects in progression to CMD evaluation by CDR-SB, MMSE, ADAS-Cog11, FAQ, and CDR Global at 18 and 24 Months, CDR-SB and ADAS were statistically significant at 18 months. However, the lower limits of the confidence intervals were close to 1, which does not provide reassurance for the robustness of the results. Furthermore, MMSE ≤-3 point change, FAQ ≥3 point change and CDR global >0 point change did not achieve statistical significance for subjects in the τAD++ group as compared to those in the non-τAD++ group (τAD- and τAD+), as in the case of the above-mentioned secondary analysis. According to the applicant, comparable trends were observed at 24 months, suggesting an increased risk of deterioration in subjects with Advanced (τAD++) flortaucipir PET patterns. However, this cannot be agreed upon, since at 24 months the values of all the endpoints did not reach statistical significance.

With respect to the exploratory analysis of diagnostic performance in predicting clinical deterioration, the sensitivity was moderately high (ranging from 81.6% to 87.5%), but specificity was very low (22.9% - 32.2%). It has been previously considered that the positive signature of CSF biomarkers, i.e. a low Aβ1-42 and high Tau are qualified to predict the evolution to dementia in patients diagnosed as MCI. The sensitivity will be at least 80% and the specificity 60%; however it should be noted that these CSF biomarker results were obtained with different outcomes (conversion from MCI to AD vs. clinically meaningful change threshold in the present study), differences in the enrolled populations (MCI only vs. MCI and AD dementia in the present study), the time of follow-up (3+ years vs. 18 months in the present study) (Qualification Opinion of Alzheimer's Disease Novel Methodologies/biomarkers for BMS-708163).

Similar to Studies A16 and FR01, inter-reader agreement in the prognosis studies (A05C and PX01) was assessed primarily using Fleiss' kappa with its 95% CI for scan reads dichotomised to tAD++ (Advanced AD pattern, expected to progress clinically) vs. non-tAD++ (Moderate AD pattern or negative for AD, not expected to progress clinically). There were no pre-specified hypotheses tested for inter-reader reliability. In the case of A05C the read/re-read kappa was greater than 0.79 in all cases, but the lower limit of the 95% CI for two of the readers was below 0.60 (0.5180 and 0.5424). In the case of PX01, the overall Fleiss Kappa was 0.754 (just above the 0.75 excellent agreement cut-off point). However, the lower limit of 95% CI was below 0.60 (0.582 and 0.561). These taken into consideration together with the specificity and sensitivity analyses point towards a weakness in the reading procedures of the scans and the inter-reader reliability.

As the applicant stated, the 24-month follow-up risk ratios in study PX01 for the τAD vs τAD- comparison were somewhat smaller and failed to achieve significance when the Risk Ratio of τAD (τAD+ and τAD++) vs. τAD- subjects in progression to Clinically Meaningful Deterioration was evaluated by CDR-SB, MMSE, ADAS-Cog11, FAQ, and CDR Global. There is no continuity in the results from 18 to 24 months, which does not provide reassurance for the "predictive abilities" of flortaucipir.

Subgroup analyses in study PX01 also could not allow any firm conclusions due to the small numbers of subjects in the subgroups.

Regarding the exploratory phase A05E, out of the 223 enrolled subjects (51 AD, 98 MCI, 58 CN and 16 OCN), 167 completed the study. Of the 223 enrolled subjects, 1 did not receive flortaucipir, leaving 222 in the primary analysis (flortaucipir as a function of diagnosis) which were 51 Alzheimer's Disease (AD), 97 mildly cognitively impaired (MCI), 58 older cognitively normal (OCN) and 16 young cognitively normal (YCN) subjects.

The confirmatory phase (A05C) included an independent population restricted to subjects clinically diagnosed as MCI or dementia with a suspected neurodegenerative cause with an MMSE score between 20 and 27, inclusive.

According to the applicant the results of the Exploratory Phase longitudinal component demonstrate that flortaucipir PET signal increased from baseline to 18 months in Aβ+ but not Aβ- subjects, and that the magnitude of the change was predicted most strongly by the baseline flortaucipir PET SUVr value. It can be agreed with the applicant that the results of the Exploratory Phase (A05E) suggest that there may be a pattern of flortaucipir retention that is unique to patients with AD pathology, which is not observed in Aβ- subjects. The analysis was used to compare the mean SUVr values between diagnostic groups (AD, MCI, and OCN) within amyloid beta status (Aβ+, Aβ-), which was performed on the MUBADA SUVr as well as the SUVr for each brain region. It should be noted that the YCN group served as a negative control to facilitate the understanding of a true negative flortaucipir scan and was not included in the analysis due to the obvious age difference between the groups and small numbers. The older Aβ- cognitively normal (OCN) individuals who, as expected, did not accumulate tau pathology over 18 months were an adequately robust control group for the Study A05E longitudinal analysis objectives.

Overall, the exploratory phase (A05E) was hypothesis generating.

In study A18, 79 subjects enrolled and completed the study A18. There were no differences in flortaucipir SUVr CFB between the AD group (CFB = 0.0024 ± 0.10211) and the MCI group (CFB = 0.0328 ±0.07378) in subjects who were Aβ+ (LSM change difference = -0.0154 ± 0.02599, p=0.5565) or those who were Aβ-. There was a statistically significant increase in flortaucipir SUVr in Aβ+ subjects, a significant decrease in Aβ- subjects, and a significant difference in CFB SUVr between the Aβ+ and Aβ- groups. Study A18 may indicate some differences between Aβ+ and Aβ- within each diagnosis group. However, with no comparator group, it is difficult to interpret these results. Apart from that, the findings are not supportive of the applicant´s position that tau scan is predictive of amyloid beta status. Both are needed to support the diagnosis of suspected AD and that should be reflected in the wording of the indication.

In the study TZAX, 224 subjects were enrolled and 181 of them completed the study. The dose in study TZAX (240MBeq) is different from the one recommended for routine use of flortaucipir (370 MBeq).

In a pooled analysis of Studies A05E, A05C, TZAX, and PX01, risk ratios for progressing to the clinically meaningful event as determined by the indices CDR-SB 1-pt change, CDR-G change >0, MMSE 3-pt change, ADAS 4-pt change, FAQ 3-pt change within 18 months, according to the applicant, were all nominally statistically significant and ranged from approximately 1.5 to 2. Although this level of risk prediction may seem modest, they must be taken in context with other individual risk factors patients will have for progression, and the fact that this represents only 18 months of change in a disease that has a 20–30-year natural history.

The applicant also claimed that this level of risk ratio had historically been demonstrated to be valuable in clinical practice. Evidence of such value comes from the cardiovascular and diabetes literature as well as from epidemiological data on cognitive progression in elderly patients. However, the pooling of such heterogeneous results with different patient populations and healthy volunteers, with different

doses administered, several methodological issues and inconsistent results is not considered appropriate. For example, one of the most important differences between the supportive Studies A05E and TZAX and the pivotal Studies A05C and PX01 was that the flortaucipir PET images in these studies were interpreted only by applicant's expert readers rather than independent external readers. Furthermore, hazard or risk ratios for individual indices were not all statistically significant in studies A05E, A05C and PX01. The only point that can be agreed with the applicant is that progression is highly variable among individual patients due to differing degrees of risk factors or protective factors, clinical presentation, and co-morbid illnesses.

In study A04, there was a numerically greater change from baseline (CFB) standardised uptake value ratio (SUVr) for the CI group compared with the HV group. However, the difference in SUVr slope (change across patients as a function of duration of follow up) was not statistically significant, possibly due to the relatively small sample size and short follow-up duration for some subjects, requiring statistical adjustment, which may have had an impact of the estimate of differences. Together with previous Avid studies and published reports the results in A04 suggest that the presence of amyloid neuritic plaques (i.e., positive amyloid PET scan) is necessary for an elevated MUBADA flortaucipir SUVr, indicative of widespread flortaucipir retention and presumably widespread tau NFTs.

In study A08 the number of the objectively impaired subjects (AD and MCI) was very small (n=14) compared to the non-objectively impaired (SMC and CN) (n=76) leading to a MUBADA SUVr mean difference of only 0.16 between these groups, which is a significantly higher MUBADA SUVr in impaired (MCI and AD) compared to non-impaired (CN and SMC) subjects. In the case of the CDR-SB the change between baseline MUBADA SUVr and at 12 months was not statistically significant (0.249, p=0.0550). Methodological and sample size issues could have contributed to the unusual findings in A08.

Assessment of paediatric data on clinical efficacy

Not applicable.

2.6.7. Conclusions on the clinical efficacy

The finally accepted wording of the indication, i.e.;

> *This medicinal product is for diagnostic use only.*
>
> *Flortaucipir (18F) is a radiopharmaceutical indicated for positron emission tomography (PET) imaging of the brain to assess the neocortical distribution of aggregated tau neurofibrillary tangles (NFTs) in adult patients with cognitive impairment who are being evaluated for Alzheimer's disease (AD). Flortaucipir (18F) is an adjunct to clinical and other diagnostic evaluations.*
>
> *For limitations of use, see sections 4.4 and 5.1.*

is supported by the available evidence and is acceptable, provided that the CHMP requested improvements to the relevant limitations are followed.

The totality of data can substantiate the modified target indication.
Relevant limitations of use and the possibility of reduced accuracy in subjects with MCI have been reflected in the SmPC.

Further relevant adaptations to the SmPC to align the contents with the modified indication still remain to be made.

For the prognostic indication (Indication II) two pivotal studies A05C and PX01 were presented for the prognostic claim however, there were important differences with respect to included populations and doses of flortaucipir applied.

In study A05C the primary analysis did not meet the pre-specified success criterion for flortaucipir PET scan visual read of τAD++ pattern to predict a higher risk of subjects´ clinically meaningful cognitive and functional deterioration defined as at a least 1-point worsening on the CDR-SB compared to baseline (HR 1.581; CI 0.986, 2.581; p = 0.067). The secondary objective of study A05C was the diagnostic performance of baseline tau positivity for predicting subjects´ clinically meaningful cognitive and functional deterioration and was not met either. For the criterion outcome CDR-SB change of 1 point the lower 95% confidence limit of specificity (52.7%), but not sensitivity (36.9% to 46.1%) just exceeded the targeted 50% for all 5 readers.

Study PXO1 met the primary endpoint, however the study was a retrospective analysis from a tau imaging substudy of the AZES therapeutic trial in a population that had higher baseline risk for AD. It did not show consistency across endpoints. With respect to the exploratory analysis of the inter-reader reliability, the sensitivity was moderately high, but specificity was very low (22.9% - 32.2%).

The conditions for predicting worsening were different between these two studies also contributing to the fact that pooling the results from both pivotal studies appears not meaningful. Overall, the risk ratios /hazard ratios appear modest in the range of 1.535 to 1.974 in the pooled analysis across studies A05E, A05C, TZAX and PX01.

The supportive studies A05E (exploratory phase of A05), A18 (follow-up of subjects from A05C), TZAX, A04 and A08 had limitations. These consisted of small sample sizes and differences in the population (leading to unusual findings), differences in dosing regimen, lack of comparator groups, methodological and design issues, in some cases absence of statistical significance and inconsistency of results. In general they cannot provide sufficient supporting evidence.

The totality of the results did not support the claim that flortaucipir (18F) can be used for a prognostic claim to inform risk for progression of cognitive and functional impairment due to AD.

The applicant accepted to withdraw the prognosis claim from further consideration in the present application procedure.

The CHMP considers the following measures necessary to address issues related to efficacy:

1. The CHMP highlighted the importance of specific and proper training for use of Tauvid to minimise the risk of Tauvid PET imaging interpretation errors and the responsibilities of the MAH in this respect. The CHMP recommends the MAH to create effective training material for training of health care professionals and to support independent educators to implement such training. These need to be completed before marketing of the product in the respective country.
2. The MAH is recommended to put in place adequate post-authorisation measures, e.g. monitoring image interpretation errors/pitfalls in image interpretation in the literature and reporting via PSURs, updating of SmPC chapter "image interpretation errors" based on the latest information. These measures are to be planned and conducted as ongoing activities.

2.6.8. Clinical safety

2.6.8.1. Patient exposure

Like other diagnostic radiopharmaceuticals it is usually administered as a single use.

Generally, safety data were collected from the time of consent through 48 hours postdose via a follow-up phone call. The applicant's argumentation that due to the rapid elimination of flortaucipir this time-period is adequate is endorsed.

No safety data specific to repeated administration of flortaucipir has been identified.

Flortaucipir was administered open-label as a single IV injection and there are no placebo or active comparator data. AEs were analysed for the overall Safety Population as well as by the subgroups diagnosis (Dx) and biomarker (BM) studies and by separate studies and summary statistics have been provided overall and by cognitive status cognitively impaired (CI) and cognitively normal (CN).

The overall Safety Population contains data of 4652 subjects who received at least 1 dose; 63 subjects (1.35%) discontinued due to AEs. Of these 63 subjects, 2 subjects (0.043%) discontinued due to an AE that occurred within 2 days of flortaucipir administration (TEAE) and 22 subjects (0.5%) discontinued due to death.

Table 58 Subject disposition by study type by cognitive status and overall enrolled population

	Diagnosis Studies			Biomarker Studies			
	CI	CN	Total	CI	CN	Total	Overall
Enrolled Population							5985
No Study Dose Injection							1333
With Study Dose Injection (Safety Population)	730	408	1138	3364	150	3514	4652
Safety Population							
Completed Study	556 (76.2%)	358 (87.7%)	914 (80.3%)	1282 (38.1%)	0	1282 (36.5%)	2196 (47.2%)
Discontinued	174 (23.8%)	50 (12.3%)	224 (19.7%)	2082 (61.9%)	150 (100%)	2232 (63.5%)	2456 (52.8%)
Primary Reason for Discontinuation							
Protocol Deviation	40 (5.5%)	1 (0.2%)	41 (3.6%)	3 (0.1%)	0	3 (0.1%)	44 (0.9%)
Adverse Event	1 (0.1%)	0	1 (0.1%)	60 (1.8%)	0	60 (1.7%)	61 (1.3%)
Modified Adverse Event*a	0	0	0	2 (0.1%)	0	2 (0.1%)	2 (0.0%)
Consent Withdrawn	72 (9.9%)	9 (2.2%)	81 (7.1%)	224 (6.7%)	0	224 (6.4%)	305 (6.6%)
Administrative Decision	38 (5.2%)	36 (8.8%)	74 (6.5%)	437 (13.0%)	0	437 (12.4%)	511 (11.0%)
Lost to Follow-up	12 (1.6%)	4 (1.0%)	16 (1.4%)	22 (0.7%)	0	22 (0.6%)	38 (0.8%)
Death	9 (1.2%)	0	9 (0.8%)	13 (0.4%)	0	13 (0.4%)	22 (0.5%)
Technical Problems	1 (0.1%)	0	1 (0.1%)	2 (0.1%)	0	2 (0.1%)	3 (0.1%)
Other	0	0	0	37 (1.1%)	0	37 (1.1%)	37 (0.8%)

Note: Percentages are based on the total number of subjects in each cognitive status using the safety population. CI: Cognitively Impaired; CN: Cognitively Normal
Enrolled population includes all subjects who have signed informed consent, regardless of whether or not they received study medication.
Safety population includes all subjects who received at least one flortaucipir.
*a Modified Adverse Event: defined as AEs that happened within 2 days following a flortaucipir scan, or was attributed to flortaucipir scan, and resulted in a subject's discontinuation from study.
Table summarises and displays results of the studies where assessments were performed.

For the overall Safety Population, the mean (SD) age was 72.9 ± 8.62 years (range 21 to 104 years) and the majority of subjects were ≥65 years of age. Flortaucipir is not indicated for use in the paediatric population.

The most frequently occurring illnesses in the overall Safety Population have been reported and together with the concomitant medications taken in the study population are considered to be in line with the expected target population. The trend towards higher rates of concomitant disease in the Dx (particularly CI_Dx) versus the BM population is most likely attributable to the Dx population consisting of subjects being sicker than in the BM population. Also, concomitant medication use was generally higher in the CI_Dx group, most notably for anticholinesterases, other anti-dementia drugs, and SSRIs.

Radiation exposure

Detailed data are provided in the section on dosimetry.

The recommended 370 MBq dose corresponds to the effective dose of 9.6 mSv that is somewhat higher than that for other 18F amyloid PET radiopharmaceuticals (around 7 mSv). The effective dose, as well as relevant warnings regarding generic risks of radiation exposure have been included in the product information.

2.6.8.2. Adverse events

The analysis of the TEAEs indicates that the incidence of most AEs was infrequent and that AEs were usually mild. In the overall Safety Population, 303 subjects (6.5%) experienced 421 TEAEs. The most frequent AEs occurring in ≥0.2% of participants were headache (0.9%), injection site pain (0.6%), diarrhoea (0.3%), blood pressure increased (0.3%), dizziness (0.2%), hypertension (0.2%), and nausea (0.2%). As to be expected, rates of TEAEs were lower in the cognitively impaired (CI), comprised of subjects with memory problems, than in the cognitively normal (CN) group.

Table 59 Summary of overall treatment-emergent adverse events by study type, by cognitive status safety population

	Diagnosis Studies			Biomarker Studies			
	CI (N=730)	CN (N=408)	Total (N=1138)	CI (N=3364)	CN (N=150)	Total (N=3514)	Overall (N=4652)
Number of Subjects Who Had a TEAE	81 (11.1%)	54 (13.2%)	135 (11.9%)	153 (4.5%)	15 (10.0%)	168 (4.8%)	303 (6.5%)
Number of TEAEs	105	81	186	217	18	235	421
Number of Subjects Who Had at least one Serious TEAE	2 (0.3%)	2 (0.5%)	4 (0.4%)	6 (0.2%)	0	6 (0.2%)	10 (0.2%)
Number of Serious TEAEs	2	2	4	6	0	6	10
Number of Subjects Who Had at Least One Moderate or Severe TEAE	15 (2.1%)	9 (2.2%)	24 (2.1%)	44 (1.3%)	4 (2.7%)	48 (1.4%)	72 (1.5%)
Number of Moderate or Severe TEAEs	16	16	32	53	5	58	90
Number of Subjects Who Had a Treatment-Related TEAE	32 (4.4%)	24 (5.9%)	56 (4.9%)	25 (0.8%)	2 (1.3%)	27 (0.8%)	83 (1.9%)
Number of Treatment-Related TEAEs	37	33	70	27	2	29	99
Number of Subjects Who Had a Procedure-Related TEAE	38 (5.3%)	21 (5.1%)	59 (5.2%)	84 (2.7%)	3 (2.0%)	87 (2.6%)	146 (3.3%)
Number of Procedure-Related TEAEs	45	33	78	122	3	125	203
Number of Subjects Who Had a TEAE Leading to Study Discontinuation	0	0	0	2 (0.1%)	0	2 (0.1%)	2 (0.0%)
Number of TEAEs Leading to Study Discontinuation	0	0	0	2	0	2	2
Number of Subjects Who Had a TEAE with an Outcome of Death	1 (0.1%)	1 (0.2%)	2 (0.2%)	0	0	0	2 (0.0%)
Number of TEAEs with an Outcome of Death	1	1	2	0	0	0	2

Note:
1. Treatment emergent AEs are undesirable experiences, signs or symptoms that begin or worsen in intensity or frequency <= 48 hours after the flortaucipir F 18 dose injection (or <= 2 days if AE onset time was not collected).
2. Moderate or Severe TEAE include TEAE with missing intensity.
3. Serious TEAE include TEAE with missing seriousness.
4. Treatment-related TEAE are those TEAE with a relationship or missing a relationship to study medication. The relevant information was not collected in studies TZAX thus this summary excludes study TZAX.
5. Procedure-related TEAE are those TEAE with a relationship or missing a relationship to injection procedure. The relevant information was not collected in studies TZAX and T807000 thus this summary excludes studies TZAX and T807000.
6. Adverse event terms are coded using MedDRA version 25.0. Percentages are based on the total number of subjects in each column except for treatment related and procedure related TEAEs, which are based on total number of subjects excluding studies TZAX and/or T807000, respectively.
7. CI: cognitively impaired; CN: cognitively normal.
8. Table summarises and displays results of the studies where assessments were performed.

Table 60 Treatment-emergent adverse events (≥0.2% of participants) by preferred term (in descending order by frequency), by study type, and by cognitive status safety population

Preferred Term	Diagnosis Studies			Biomarker Studies			
	CI (N=730)	CN (N=408)	Total (N=1138)	CI (N=3364)	CN (N=150)	Total (N=3514)	Overall (N=4652)
Headache	14 (1.9%)	13 (3.2%)	27 (2.4%)	12 (0.4%)	1 (0.7%)	13 (0.4%)	40 (0.9%)
Injection site pain	19 (2.6%)	6 (1.5%)	25 (2.2%)	4 (0.1%)	0	4 (0.1%)	29 (0.6%)
Diarrhoea	3 (0.4%)	3 (0.7%)	6 (0.5%)	7 (0.2%)	0	7 (0.2%)	13 (0.3%)
Blood pressure increased	5 (0.7%)	4 (1.0%)	9 (0.8%)	3 (0.1%)	0	3 (0.1%)	12 (0.3%)
Dizziness	3 (0.4%)	2 (0.5%)	5 (0.4%)	4 (0.1%)	1 (0.7%)	5 (0.1%)	10 (0.2%)
Hypertension	4 (0.5%)	1 (0.2%)	5 (0.4%)	2 (0.1%)	1 (0.7%)	3 (0.1%)	8 (0.2%)
Nausea	1 (0.1%)	2 (0.5%)	3 (0.3%)	3 (0.1%)	1 (0.7%)	4 (0.1%)	7 (0.2%)
Dysgeusia[a]	0	4 (1.0%)	4 (0.4%)	1 (0.0%)	0	1 (0.0%)	5 (0.1%)

Abbreviations: CI = cognitively impaired; CN = cognitively normal; N = number of subjects in the population.
a Dysgeusia has been included as it has been determined to be an adverse reaction of flortaucipir due to temporal relationship with administration.

Of the 421 TEAEs, 409 (97.1%) were mild or moderate, and the majority was assessed as not related to the study drug. The 12 severe events (2.9%) included one event of T wave abnormal on ECG reported in Study LLCF, but this event was a pre-existing condition inadvertently captured as an SAE. In an unadjusted analysis, subjects who experienced TEAEs had a higher mass dose (~0.1 µg) than those who did not, but after adjusting for study as a confounding factor, no statistically significant difference in LS mean mass dose was seen (p=0.9187) indicating that the difference was most likely attributed to the mass dose variation across pooled studies.

Table 61 Summary of treatment-emergent adverse events by system organ class and preferred term, severity, study type, and cognitive status safety population

System Organ Class Preferred Term Severity	Diagnosis Studies CI (N=730)	 CN (N=408)	 Total (N=1138)	Biomarker Studies CI (N=3364)	 CN (N=150)	 Total (N=3514)	Overall (N=4652)
Total Number of TEAE	105	81	186	217	18	235	421
Mild	89	65	154	164	13	177	331
Moderate	14	14	28	46	4	50	78
Severe	2	2	4	7	1	8	12
Number of Subjects With at Least One TEAE	81 (11.1%)	54 (13.2%)	135 (11.9%)	153 (4.5%)	15 (10.0%)	168 (4.8%)	303 (6.5%)
Mild	66 (9.0%)	45 (11.0%)	111 (9.8%)	109 (3.2%)	11 (7.3%)	120 (3.4%)	231 (5.0%)
Moderate	13 (1.8%)	7 (1.7%)	20 (1.8%)	37 (1.1%)	3 (2.0%)	40 (1.1%)	60 (1.3%)
Severe	2 (0.3%)	2 (0.5%)	4 (0.4%)	7 (0.2%)	1 (0.7%)	8 (0.2%)	12 (0.3%)
Nervous System Disorders	22 (3.0%)	22 (5.4%)	44 (3.9%)	38 (1.1%)	5 (3.3%)	43 (1.2%)	87 (1.9%)
Mild	18 (2.5%)	18 (4.4%)	36 (3.2%)	32 (1.0%)	4 (2.7%)	36 (1.0%)	72 (1.5%)
Moderate	4 (0.5%)	4 (1.0%)	8 (0.7%)	5 (0.1%)	1 (0.7%)	6 (0.2%)	14 (0.3%)
Severe	0	0	0	1 (0.0%)	0	1 (0.0%)	1 (0.0%)
Transient Ischaemic Attack	0	0	0	2 (0.1%)	0	2 (0.1%)	2 (0.0%)
Mild	0	0	0	0	0	0	0
Moderate	0	0	0	1 (0.0%)	0	1 (0.0%)	1 (0.0%)
Severe	0	0	0	1 (0.0%)	0	1 (0.0%)	1 (0.0%)
General Disorders And Administration Site Conditions	29 (4.0%)	12 (2.9%)	41 (3.6%)	14 (0.4%)	0	14 (0.4%)	55 (1.2%)
Mild	27 (3.7%)	10 (2.5%)	37 (3.3%)	10 (0.3%)	0	10 (0.3%)	47 (1.0%)
Moderate	1 (0.1%)	2 (0.5%)	3 (0.3%)	4 (0.1%)	0	4 (0.1%)	7 (0.2%)
Severe	1 (0.1%)	0	1 (0.1%)	0	0	0	1 (0.0%)
Injection Site Pain	19 (2.6%)	6 (1.5%)	25 (2.2%)	4 (0.1%)	0	4 (0.1%)	29 (0.6%)
Mild	17 (2.3%)	6 (1.5%)	23 (2.0%)	3 (0.1%)	0	3 (0.1%)	26 (0.6%)
Moderate	1 (0.1%)	0	1 (0.1%)	1 (0.0%)	0	1 (0.0%)	2 (0.0%)
Severe	1 (0.1%)	0	1 (0.1%)	0	0	0	1 (0.0%)
Investigations	8 (1.1%)	6 (1.5%)	14 (1.2%)	9 (0.3%)	0	9 (0.3%)	23 (0.5%)
Mild	7 (1.0%)	6 (1.5%)	13 (1.1%)	6 (0.2%)	0	6 (0.2%)	19 (0.4%)
Moderate	1 (0.1%)	0	1 (0.1%)	2 (0.1%)	0	2 (0.1%)	3 (0.1%)
Severe	0	0	0	1 (0.0%)	0	1 (0.0%)	1 (0.0%)
Electrocardiogram T Wave Abnormal	0	0	0	1 (0.0%)	0	1 (0.0%)	1 (0.0%)
Mild	0	0	0	0	0	0	0
Moderate	0	0	0	0	0	0	0
Severe	0	0	0	1 (0.0%)	0	1 (0.0%)	1 (0.0%)
Musculoskeletal And Connective Tissue Disorders	4 (0.5%)	3 (0.7%)	7 (0.6%)	15 (0.4%)	1 (0.7%)	16 (0.5%)	23 (0.5%)
Mild	2 (0.3%)	2 (0.5%)	4 (0.4%)	11 (0.3%)	0	11 (0.3%)	15 (0.3%)
Moderate	2 (0.3%)	1 (0.2%)	3 (0.3%)	3 (0.1%)	1 (0.7%)	4 (0.1%)	7 (0.2%)
Severe	0	0	0	1 (0.0%)	0	1 (0.0%)	1 (0.0%)
Muscular Weakness	0	0	0	2 (0.1%)	0	2 (0.1%)	2 (0.0%)
Mild	0	0	0	1 (0.0%)	0	1 (0.0%)	1 (0.0%)
Moderate	0	0	0	0	0	0	0
Severe	0	0	0	1 (0.0%)	0	1 (0.0%)	1 (0.0%)
Psychiatric Disorders	4 (0.5%)	2 (0.5%)	6 (0.5%)	14 (0.4%)	1 (0.7%)	15 (0.4%)	21 (0.5%)
Mild	2 (0.3%)	1 (0.2%)	3 (0.3%)	11 (0.3%)	0	11 (0.3%)	14 (0.3%)
Moderate	2 (0.3%)	1 (0.2%)	3 (0.3%)	2 (0.1%)	0	2 (0.1%)	5 (0.1%)
Severe	0	0	0	1 (0.0%)	1 (0.7%)	2 (0.1%)	2 (0.0%)
Insomnia	1 (0.1%)	0	1 (0.1%)	4 (0.1%)	1 (0.7%)	5 (0.1%)	6 (0.1%)
Mild	1 (0.1%)	0	1 (0.1%)	3 (0.1%)	0	3 (0.1%)	4 (0.1%)
Moderate	0	0	0	1 (0.0%)	0	1 (0.0%)	1 (0.0%)
Severe	0	0	0	0	1 (0.7%)	1 (0.0%)	1 (0.0%)
Agitation	2 (0.3%)	1 (0.2%)	3 (0.3%)	2 (0.1%)	0	2 (0.1%)	5 (0.1%)
Mild	0	1 (0.2%)	1 (0.1%)	1 (0.0%)	0	1 (0.0%)	2 (0.0%)
Moderate	2 (0.3%)	0	2 (0.2%)	0	0	0	2 (0.0%)
Severe	0	0	0	1 (0.0%)	0	1 (0.0%)	1 (0.0%)

System Organ Class Preferred Term Severity	Diagnosis Studies CI (N=730)	 CN (N=408)	 Total (N=1138)	Biomarker Studies CI (N=3364)	 CN (N=150)	 Total (N=3514)	Overall (N=4652)
Injury, Poisoning And Procedural Complications	2 (0.3%)	3 (0.7%)	5 (0.4%)	11 (0.3%)	0	11 (0.3%)	16 (0.3%)
Mild	2 (0.3%)	3 (0.7%)	5 (0.4%)	8 (0.2%)	0	8 (0.2%)	13 (0.3%)
Moderate	0	0	0	1 (0.0%)	0	1 (0.0%)	1 (0.0%)
Severe	0	0	0	2 (0.1%)	0	2 (0.1%)	2 (0.0%)
Fall	1 (0.1%)	0	1 (0.1%)	5 (0.1%)	0	5 (0.1%)	6 (0.1%)
Mild	1 (0.1%)	0	1 (0.1%)	3 (0.1%)	0	3 (0.1%)	4 (0.1%)
Moderate	0	0	0	1 (0.0%)	0	1 (0.0%)	1 (0.0%)
Severe	0	0	0	1 (0.0%)	0	1 (0.0%)	1 (0.0%)
Hand Fracture	0	0	0	1 (0.0%)	0	1 (0.0%)	1 (0.0%)
Mild	0	0	0	0	0	0	0
Moderate	0	0	0	0	0	0	0
Severe	0	0	0	1 (0.0%)	0	1 (0.0%)	1 (0.0%)
Cardiac Disorders	4 (0.5%)	2 (0.5%)	6 (0.5%)	5 (0.1%)	0	5 (0.1%)	11 (0.2%)
Mild	2 (0.3%)	0	2 (0.2%)	4 (0.1%)	0	4 (0.1%)	6 (0.1%)
Moderate	2 (0.3%)	1 (0.2%)	3 (0.3%)	1 (0.0%)	0	1 (0.0%)	4 (0.1%)
Severe	0	1 (0.2%)	1 (0.1%)	0	0	0	1 (0.0%)
Myocardial Infarction	0	1 (0.2%)	1 (0.1%)	0	0	0	1 (0.0%)
Mild	0	0	0	0	0	0	0
Moderate	0	0	0	0	0	0	0
Severe	0	1 (0.2%)	1 (0.1%)	0	0	0	1 (0.0%)
Metabolism And Nutrition Disorders	1 (0.1%)	1 (0.2%)	2 (0.2%)	5 (0.1%)	0	5 (0.1%)	7 (0.2%)
Mild	1 (0.1%)	0	1 (0.1%)	3 (0.1%)	0	3 (0.1%)	4 (0.1%)
Moderate	0	1 (0.2%)	1 (0.1%)	1 (0.0%)	0	1 (0.0%)	2 (0.0%)
Severe	0	0	0	1 (0.0%)	0	1 (0.0%)	1 (0.0%)
Hyperglycaemia	0	0	0	1 (0.0%)	0	1 (0.0%)	1 (0.0%)
Mild	0	0	0	0	0	0	0
Moderate	0	0	0	0	0	0	0
Severe	0	0	0	1 (0.0%)	0	1 (0.0%)	1 (0.0%)
Neoplasms Benign, Malignant And Unspecified (Incl Cysts And Polyps)	0	1 (0.2%)	1 (0.1%)	5 (0.1%)	0	5 (0.1%)	6 (0.1%)
Mild	0	0	0	2 (0.1%)	0	2 (0.1%)	2 (0.0%)
Moderate	0	0	0	3 (0.1%)	0	3 (0.1%)	3 (0.1%)
Severe	0	1 (0.2%)	1 (0.1%)	0	0	0	1 (0.0%)
Neoplasm Malignant	0	1 (0.2%)	1 (0.1%)	0	0	0	1 (0.0%)
Mild	0	0	0	0	0	0	0
Moderate	0	0	0	0	0	0	0
Severe	0	1 (0.2%)	1 (0.1%)	0	0	0	1 (0.0%)
Renal And Urinary Disorders	1 (0.1%)	2 (0.5%)	3 (0.3%)	1 (0.0%)	0	1 (0.0%)	4 (0.1%)
Mild	0	2 (0.5%)	2 (0.2%)	1 (0.0%)	0	1 (0.0%)	3 (0.1%)
Moderate	0	0	0	0	0	0	0
Severe	1 (0.1%)	0	1 (0.1%)	0	0	0	1 (0.0%)
Acute Kidney Injury	1 (0.1%)	0	1 (0.1%)	0	0	0	1 (0.0%)
Mild	0	0	0	0	0	0	0
Moderate	0	0	0	0	0	0	0
Severe	1 (0.1%)	0	1 (0.1%)	0	0	0	1 (0.0%)

Note:
1. Treatment emergent AEs are undesirable experiences, signs or symptoms that begin or worsen in intensity or frequency <= 48 hours after the flortaucipir F 18 dose injection (or <= 2 days if AE onset time was not collected).
2. At each level of summarisation, a subject is counted once for the most severe event if the subject reported one or more occurrences of the same event.
3. If the severity is missing or unknown, the AE was counted as "Severe".
4. The table is listed in a descending order according to the frequency of system organ class and preferred term in the overall column.
5. Adverse event terms are coded using MedDRA version 25.0
6. Percentages are based on the total number of subjects in each column. CI: cognitively impaired; CN: cognitively normal.
7. Table summarises and displays results of the studies where assessments were performed.

Safety results from investigator-led studies were in line with findings from the overall Safety population.

As regards adverse reaction terms to be included in the PI the applicant proposed headache (n=40 [0.9%]), injection site pain (n=29 [0.6%]), blood pressure increased (n=24; [0.5%]; 8 events hypertension, 3 events blood pressure systolic increased, 1 event hypertensive urgency), and dysgeusia (n=5; 0.1%). This is endorsed. There were 3 cases of injection site extravasation reported as related to study medication; despite the high osmolarity of flortaucipir, the risk of injection site reaction and extravasation appears to be low. There were sporadic reports of skin rashes, hypotension, flushing, swelling face. Hypersensitivity has been associated with similar type agents. Statistically significant increases in systolic blood pressure and diastolic blood pressure were seen immediately post

dose and at the end-of-scan time point. The most marked effect was seen on SBP at the end of treatment cycle where the magnitude of the mean increase ranged from 6.4- 8.7 mmHG across three repeat treatments at the end of study endpoint. Mean DBP increases ranged from 1mmHG - 1.3mmHG immediately post dose to 3.1 - 3.3 mm HG at end of study.

Of 23 subjects who experienced 29 PCS increases in blood pressure, the applicant stated 8 are reported as having reported TEAEs of hypertension or blood pressure increased. ISS table 303 (Treatment Emergent Adverse Events by Preferred Term) includes 8 reports of hypertension but also includes 12 reports of blood pressure increased and 3 reports blood pressure systolic increased. Of note, there were no severe or serious cases reported. Ten of these were reported as procedure related.

Although a higher background rate of hypertension is to be expected in this population (22% had hx of CVD, 6% had a hx of vascular disease, 14% were on ACE Inhibitors, 14% were taking selective b-blockers), the 29 PCS changes in systolic blood pressure or diastolic blood pressure included quite marked increases ranging from 20 - 78 mmHg in systolic pressure and 15 - 68mmHg in DBP. However, overall, there is no clinical picture suggestive of increases in BP being directly related to a pharmacodynamic effect of flortaucipir and there were no clinical sequelae associated with these; inclusion of blood pressure increased as an uncommon ADR is endorsed.

Overall, the analysis of the available data currently does not indicate AEs limiting the use of flortaucipir from a safety perspective. However, the analysis is limited by the absence of any type of comparator and of any long-time data as well as missing data on repeated use.

2.6.8.3. Serious adverse event/deaths/other significant events

Deaths

The applicant considered deaths as treatment emergent if they occurred within 7 days after treatment. Two deaths occurred within this 7-day period. One death was due to acute kidney injury in a subject with declining ability to perform activities of daily living including eating and drinking. The other death was due to a malignant neoplasm in a subject with reported end stage liver cancer. Both events occurred in subjects enrolled in study A16 that recruited subjects with a projected life expectancy of ≤6 months at study enrolment.

Twenty-nine deaths have been reported in the safety population, irrespective of the time interval between flortaucipir administration and death. Narratives have been provided for all deaths reported irrespective of the time interval between flortaucipir administration and death. No new concerns have been identified.

Other Serious Adverse Events

Angina pectoris (study A05), myocardial infarction (study A16), and transient ischaemic attack (1 in study LLCF, 1 in study LMDC) have been reported as SAEs within 48 hours after flortaucipir administration. The applicant has also added 3 events as SAEs (hand fracture, agitation, hyperglycaemia) although seriousness was not provided by the investigator. For the cardiovascular events, alternative causes are considered likely.

Other Significant Adverse Events

The two other significant adverse events reported by the applicant, one event of cholelithiasis and one event of cerebral microhaemorrhage, are not considered to be likely related to flortaucipir administration.

2.6.8.4. Laboratory findings

Differences in mean change from baseline to postdose in haematology parameters have been seen with flortaucipir administration, but these changes were not considered clinically meaningful; they were small in magnitude, within normal ranges, and mean changes often not obviously in a detrimental direction. Mean changes also did not demonstrate a common trend suggestive of toxicity across related parameters.

Also, small significant mean changes from baseline to postdose were seen in chemistry and urinalysis parameters, none were considered clinically significant.

As regards vital signs, there were small but significant changes in several parameters from baseline measurements. These changes were not considered clinically relevant, and the events show no clinically meaningful relationship to the mass dose of the study drug. No clinically meaningful or statistically significant correlation was seen between mass dose of flortaucipir and systolic or diastolic blood pressure at either the postdose or end-of-scan time point; mass dose ranged from 0.01 to 13 µg. It is considered likely that the events of increased blood pressure seen in vital signs analysis are rather related to the timing of blood pressure measurements and imaging procedures than to flortaucipir administration.

Small statistically significant changes were seen in multiple ECG parameters at the postdose and end-of-scan time points (n=546; more than). There were increases in QTcB and QTcF at the end-of-scan time point at each imaging visit. Overall, the mean increases in QTcB and QTcF at the end-of-scan time point were 2.34 and 5.14 msec, respectively, both the CI and CN groups showed similar mean changes. No subjects demonstrated an increase in QTcF greater than 60 msec above baseline values or reached post-baseline value of >500 msec either immediately post-dose, or at the end of scan. Only 1 subject had a greater than 60-msec (61.5 msec) increase in QTcB from baseline to the end of scan. Increase in QTcF by >30 msec immediately post-dose and end of scan time was report was observed in 0.4% and 1.7% of the cases. Increase of QTcF above normal range was observed in 1.2% and 1.8% of the cases at post-dose and end of scan ECGs, respectively.

No statistically significant correlation was seen between mean change in QTcB or QTcF at the end-of-scan time point and mass dose. No statistically significant changes in QTc were noted at the time of peak plasma concentration (immediately after injection). No TEAEs related to QT interval prolongation or ventricular arrhythmias were reported. For cardiovascular safety pharmacology assessments see non-clinical assessment.

The four clinical studies included in the safety pooled analysis excluded patients with current clinically significant cardiovascular disease or clinically significant abnormalities on screening ECG (including but not limited to QTc>450 msec) and patients with a history of additional risk factors for Torsades de Pointes (TdP) (e.g., heart failure, hypokalaemia, family history of Long QT syndrome).

While significant changes in temperature, heart rate, blood pressure, and respiratory rate were seen immediately post-dose and at the end-of-scan time point, potentially clinically significant measures were primarily seen at the end-of-scan time point, thus suggesting that these changes were related to the scan procedure rather than the flortaucipir administration.

2.6.8.5. In vitro biomarker test for patient selection for safety

N/A

2.6.8.6. Safety in special populations

The applicant has provided safety analyses for the subpopulations age, gender, race, BMI, pre-existing hepatic or renal impairment, and apolipoprotein E4 (ApoE4) status.

As regards age only an analysis of TEAEs in subjects ≥65 versus <65 years of age has been provided. The applicant argues that due to the small numbers of individual events reported for subjects ≥65 years of age, a further division into age subgroups was not considered to yield useful information. The limited analysis provided did not reveal AE concerns specific to an elder population.

Furthermore, the analyses did not indicate clinically significant difference in pattern or frequency of AEs between female and male subjects, non-white and white subjects, β-amyloid status negative or positive subjects, β-amyloid apolipoprotein (Apo) E4 allele carrier, non-carrier, or unknown status subjects, or subjects with pre-existing hepatic or renal impairment.

A comparison of TEAEs by body mass index (BMI) indicated that subjects with a BMI >32 kg/m^2 experienced more TEAEs than those in the other 2 groups (<19 kg/m^2, ≥19 kg/m^2 - ≤32 kg/m^2), but the small number of patients with a BMI <19 kg/m^2 limits the ability to discern potential differences in TEAE pattern or frequency. Subjects with a BMI >32 kg/m^2 reported headache and injection site pain (5/311 [1.6%] and 8/311 [2.6%], respectively) more frequently than those with a BMI ≥19 and ≤32 kg/m^2 (27/2154 [1.3%] and 19/2154 [0.9%], respectively).

As flortaucipir is excreted through the hepatobiliary and renal systems, subjects with hepatic or renal impairment have a potential of increased radiation exposure, but the radioactive decay of ^{18}F with a $t_{1/2}$ = 109.77 minutes is independent of the flortaucipir elimination and would be 98% decayed at 10 hours regardless of hepatic or renal function. Wording has been added to the SmPC that careful consideration of the benefit-risk ratio in these patients is required since increased radiation exposure is possible. As concerns dosimetry please see '*Pharmacokinetics*' above.

Analysis of TEAEs by cognitive cohort and by subjects taking or not taking concomitant medications for AD showed that subjects not taking concomitant AD medications reported headache and injection site pain more frequently than those taking concomitant AD medications (headache 18/818 [2.2%] versus 6/1195 [0.5%], injection site pain 15/818 [1.8%] versus 9/1195 [0.8%], respectively).

2.6.8.7. Immunological events

Immunological events have not been investigated.

2.6.8.8. Safety related to drug-drug interactions and other interactions

There are no known drug interactions for flortaucipir.

2.6.8.9. Discontinuation due to adverse events

Two (2) subjects were discontinued from clinical studies due to AEs. One (1) AE of cholelithiasis occurred on the day of flortaucipir administration in a subject with a history of cholelithiasis. Furthermore, 2 additional microhaemorrhages were identified on repeat MRI one day after flortaucipir administration in a subject with a history of cerebral microhaemorrhage. None of these AEs were considered related to flortaucipir administration.

2.6.8.10. Post marketing experience

Flortaucipir has been approved since May 2020 in the USA; however, commercial distribution was delayed for business reasons until October 2022. As of 31 January 2023, a total of 11 commercial doses have been sold in the US and no serious or significant postmarketing safety events have been reported.

With the responses to the Day 120 list of questions the applicant has clarified that they have not received any post marketing reports of AEs for flortaucipir as of 22 September 2023.

2.6.9. Discussion on clinical safety

Flortaucipir is an 18F-labeled diagnostic positron emission tomography (PET) radiopharmaceutical designed to image aggregated tau neurofibrillary tangles (NFTs) of Alzheimer's disease (AD) in the brains of patients with cognitive impairment being evaluated for AD and other causes of cognitive decline. Like other diagnostic radiopharmaceuticals it is usually administered as a single use but repeated administration for assessment of cognitive development might occur considering the indications applied for.

Generally, safety data were collected from the time of consent through 48 hours postdose; the applicant's argumentation that due to the rapid elimination of flortaucipir this time-period is adequate is endorsed. No safety data specific to repeated administration of flortaucipir has been identified, but this is acceptable given, that the product is recommended for single use.

Flortaucipir was administered open-label as a single IV injection and there are no placebo or active comparator data. AEs were analysed for the overall safety population as well as by the subgroups diagnosis (Dx) and biomarker (BM) studies and by separate studies. Summary statistics have been provided overall and by cognitive status cognitively impaired (CI) and cognitively normal (CN).

The overall safety population from the 28 clinical trials included in the safety analysis contains data of 4652 subjects who received at least one dose; of these 63 subjects (1.35%) discontinued due to AEs. Of the 63 subjects, 2 (0.043%) discontinued due to an AE that occurred within 2 days of flortaucipir administration (TEAE) and 22 (0.5%) discontinued due to death. The mean (SD) age was 72.9 (± 8.62) years (range 21 to 104 years) and the majority of subjects were ≥65 years of age which is in line with the expected target population. Flortaucipir is not indicated for use in the paediatric population.

The most frequently occurring illnesses as well as the concomitant medications taken in the study population are considered to be in line with the expected target population. The trend towards higher rates of concomitant disease in the Dx (particularly CI_Dx) versus the BM population is most likely attributable to the Dx population consisting of subjects being sicker than in the BM population. In line with this finding, concomitant medication use was generally higher in the CI_Dx group, most notably for anticholinesterases, other anti-dementia drugs, and SSRIs.

The analysis of the TEAEs indicates that the incidence of most AEs was infrequent and that AEs were usually mild. In the overall safety population, 303 subjects (6.5%) experienced 421 TEAEs. The most frequent AEs occurring in ≥0.2% of participants were headache (0.9%), injection site pain (0.6%), diarrhoea (0.3%), blood pressure increased (0.3%), dizziness (0.2%), hypertension (0.2%), and nausea (0.2%). As to be expected, rates of TEAEs were lower in the cognitively impaired (CI), comprised of subjects with memory problems, than in the cognitively normal (CN) group.

Of the 421 TEAEs, 409 (97.1%) were mild or moderate, and the majority was assessed as not related to the study drug. The 12 severe events (2.9%) included one event of T wave abnormal on ECG

reported in Study LLCF, but this event was a pre-existing condition inadvertently captured as an SAE. In an unadjusted analysis, subjects who experienced TEAEs had a higher mass dose (~0.1 µg) than those who did not, but after adjusting for study as a confounding factor, no statistically significant difference in LS mean mass dose was seen (p=0.9187) indicating that the difference was most likely attributed to the mass dose variation across pooled studies.

Safety results from investigator-led studies were in line with findings from the overall Safety population.

As regards adverse reaction terms to be included in the PI, headache, injection site pain, blood pressure increased, and dysgeusia were included. There were 3 cases of injection site extravasation reported as related to study medication. There were sporadic reports of skin rashes, hypotension, flushing, swelling face; hypersensitivity has been associated with similar type agents. There were cases of significant increase in blood pressure reported, but overall, there is no clinical picture suggestive of increases in BP being directly related to a pharmacodynamic effect of flortaucipir and there were no clinical sequelae associated with these; inclusion of blood pressure increased as an uncommon ADR is endorsed.

Overall, the analysis of the available data currently does not indicate AEs limiting the use of flortaucipir from a safety perspective. Safety information was collected in the setting without placebo control, single use and long-term follow-up. However, this is acceptable given that the substance is recommended for single application in a micro-dose range.

As regards deaths, serious AEs, and other significant events the applicant considered deaths as treatment emergent if they occurred within 7 days after treatment. Two deaths occurred within this 7-day period. One death was due to acute kidney injury in a subject with declining ability to perform activities of daily living including eating and drinking. The other death was due to a malignant neoplasm in a subject with reported end stage liver cancer. Both events occurred in subjects enrolled in study A16 that recruited subjects with a projected life expectancy of ≤6 months at study enrolment.

Angina pectoris (study A05), myocardial infarction (study A16), and transient ischaemic attack (1 in study LLCF, 1 in study LMDC) have been reported as SAEs within 48 hours after flortaucipir administration. The applicant has also added 3 events as SAEs (hand fracture, agitation, hyperglycaemia) although seriousness was not provided by the investigator. For the cardiovascular events, alternative causes are considered likely.
The two other significant adverse events reported by the applicant, one event of cholelithiasis and one event of cerebral microhaemorrhages, are not considered to be likely related to flortaucipir administration.
Overall, the analysis of deaths, serious adverse events, and other significant events currently does not indicate unacceptable significant risks associated with the administration of flortaucipir.

As regards laboratory findings and vital signs differences in mean change from baseline to postdose in haematology parameters have been seen with flortaucipir administration, but these changes were not considered clinically meaningful. Also, small significant mean changes from baseline to postdose were seen in chemistry and urinalysis parameters that were not considered clinically significant. As regards vital signs, there were small but significant changes in several parameters from baseline measurements. These changes were not considered clinically relevant, and the events show no clinically meaningful relationship to the mass dose of the study drug. As regards events of increased blood pressure seen in vital signs analysis, there is no clinical picture suggestive of increases in BP being directly related to a pharmacodynamic effect of flortaucipir and there were no clinical sequelae associated with these.

Small statistically but not clinically significant changes were seen in multiple ECG parameters at the postdose and end-of-scan time points. There were increases in QTcB and QTcF at the end-of-scan time point at each imaging visit. Overall, the mean increases in QTcB and QTcF at the end-of-scan time point were 2.34 and 5.14 msec, respectively, both the CI and CN groups showed similar mean changes. However, none of the patients had an increase in QTcF interval considered to correlate with increased risk of torsades de pointes (e.g., increase by >60 msec, or QTcF above 500 msec). QTcF prolongations above 30 msec compared to baseline or above the normal value was reported in < 2% cases. Larger changes observed in the QTcB values are considered of limited informative value since QTcB is less reliable compared to QTcF in terms of its accuracy.

No statistically significant correlation was seen between mean change in QTcB or QTcF at the end-of-scan time point and mass dose. Although there were statistically significant changes in QTc, none were considered large enough to raise serious safety concerns and no TEAEs related to QT interval prolongation or ventricular arrhythmias were reported. Of note, due to the absence of a placebo control the data are not placebo-adjusted (as would be done in a well-designed thorough QT study). Further, the four clinical studies included in the safety pooled analysis excluded patients with current clinically significant cardiovascular disease or clinically significant abnormalities on screening ECG (including but not limited to QTc>450 msec) and patients with a history of additional risk factors for Torsades de Pointes (TdP) (e.g., heart failure, hypokalaemia, family history of Long QT syndrome). Nonetheless, the recently conducted GLP-conform new hERG study indicates an influence of F18-flortaucipir on hERG currents. However, the safety margin of 338fold between *in vitro* IC_{50} and clinical exposure to unbound drug substance is considered high enough to conclude that a torsadogenic potential is unlikely (reference is made to the non-clinical part of this document). Further, the product is intended for single use only and flortaucipir is eliminated from the body quickly. Given the above, the risk of QT prolongation on flortaucipir is regarded as very low in the clinical setting. While significant changes in temperature, heart rate, blood pressure, and respiratory rate were seen immediately postdose and at the end-of-scan time point, potentially clinically significant measures were primarily seen at the end-of-scan time point, thus also suggesting that these changes were rather procedure related than to the flortaucipir administration.

As regards safety in special populations the applicant has provided safety analyses for the subpopulations age, gender, race, BMI, pre-existing hepatic or renal impairment, and apolipoprotein E4 (ApoE4) status. Concerning age only an analysis of TEAEs in subjects ≥65 versus <65 years of age has been provided. The applicant argues that due to the small numbers of individual AEs reported for subjects ≥65 years of age, a further division into age subgroups was not considered to yield useful information. The limited analysis provided did not reveal AE concerns specific to an elder population. Furthermore, the analyses did not indicate clinically significant difference in pattern or frequency of AEs between female and male subjects, non-white and white subjects, β-amyloid status negative or positive subjects, β-amyloid apolipoprotein (Apo) E4 allele carrier, non-carrier, or unknown status subjects, or subjects with pre-existing hepatic or renal impairment. A comparison of TEAEs by body mass index (BMI) indicated that subjects with a BMI >32 kg/m^2 experienced more TEAEs than those in the other 2 groups (<19 kg/m^2, ≥19 kg/m^2 - ≤32 kg/m^2), but the small number of patients with a BMI <19 kg/m^2 limits the interpretation. Subjects with a BMI >32 kg/m^2 reported headache and injection site pain more frequently than those with a BMI ≥19 and ≤32 kg/m^2.

As flortaucipir is excreted through the hepatobiliary and renal systems, subjects with hepatic or renal impairment have a potential of increased radiation exposure, but the radioactive decay of ^{18}F is independent of the flortaucipir elimination and would be 98% decayed at 10 hours regardless of hepatic or renal function. Wording has nevertheless been added to the SmPC to consider the benefit-risk ratio in these patients due to a possible increase in radiation exposure.

The recommended 370 MBq dose corresponds to the effective dose of 9.6 mSv that is somewhat higher than that for other 18F amyloid PET radiopharmaceuticals (around 7 mSv), but much lower than the limit value of 20 mSv in a calendar year defined for protection of occupationally exposed persons ((§ 78(1), Radiation Protection Act), and thus acceptable from the safety perspective. Appropriate warnings regarding the risks of radiation exposure have been included in the PI. As concerns dosimetry please see '*Pharmacokinetics*' above.

Analysis of TEAEs by cognitive cohort and by subjects taking or not taking concomitant medications for AD showed that subjects not taking concomitant AD medications reported headache and injection site pain more frequently than those taking concomitant AD medications.

Immunological events have not been investigated.

There are no known drug interactions for flortaucipir.

As regards discontinuation due to adverse events the AEs of cholelithiasis and cerebral microhaemorrhages leading to the discontinuation of 2 subjects are not considered to be likely related to flortaucipir administration.

From the safety database all the adverse reactions reported in clinical trials have been included in the Summary of Product Characteristics (SmPC).

Assessment of paediatric data on clinical safety

Not applicable

2.6.10. Conclusions on the clinical safety

Safety of flortaucipir has been analysed in 4652 subjects. Overall, the analysis of the available data currently does not indicate AEs limiting the use of flortaucipir from a safety perspective. The identified events at least possibly related to the substance have been included in the product information.

2.7. Risk Management Plan

2.7.1. Safety concerns

None.

2.7.2. Pharmacovigilance plan

No additional pharmacovigilance activities.

2.7.3. Risk minimisation measures

None

2.7.4. Conclusion

The CHMP considers that the risk management plan version 1.0 is acceptable.

2.8. Pharmacovigilance

2.8.1. Pharmacovigilance system

The CHMP considered that the pharmacovigilance system summary submitted by the applicant fulfils the requirements of Article 8(3) of Directive 2001/83/EC.

2.8.2. Periodic Safety Update Reports submission requirements

The requirements for submission of periodic safety update reports for this medicinal product are set out in the Annex II, Section C of the CHMP Opinion. The applicant did request alignment of the PSUR cycle with the international birth date (IBD). The IBD is 28.5.2020. The new EURD list entry will therefore use the IBD to determine the forthcoming Data Lock Points.

2.9. Product information

2.9.1. User consultation

The results of the user consultation with target patient groups on the package leaflet submitted by the applicant show that the package leaflet meets the criteria for readability as set out in the *Guideline on the readability of the label and package leaflet of medicinal products for human use.*

2.9.2. Labelling exemptions

A request to omit certain particulars from the labelling as per Art.63.3 of Directive 2001/83/EC has been submitted by the applicant and has been found acceptable by the QRD Group for the following reasons:

The QRD group considered that the information in the shield label is sufficient to guarantee the correct use of the medicinal product. Also bearing in mind that the availability of the product in the market should not be hampered, and this approach is in line with other radiopharmaceuticals.

The QRD Group accepted the use of minimum particulars for vials (15mL) of both strengths.

Additionally, the QRD group agreed the flexibilities on the label of the vial (both strengths) in relation to volume, amount of radioactivity and expiry date. Particularly, it is accepted to omit the total volume from the immediate label and to use the format *"ToC + (<number of hours depending on the strength>h)"* as expiry date. It was also agreed to use the format *"≤ <maximum activity> MBq at ToC (see outer packaging)"* as amount of activity on the vial label.

The CHMP is in agreement with the QRD group recommendations for the package leaflet (PL).

2.9.3. Additional monitoring

Pursuant to Article 23(1) of Regulation No (EU) 726/2004, Tauvid (Flortaucipir (18F)) is included in the additional monitoring list as it contains a new active substance which, on 1 January 2011, was not contained in any medicinal product authorised in the EU.

Therefore the summary of product characteristics and the package leaflet includes a statement that this medicinal product is subject to additional monitoring and that this will allow quick identification of

new safety information. The statement is preceded by an inverted equilateral black triangle.

3. Benefit-Risk Balance

This section is focused on the approved indication, the withdrawn indication has not been presented and discussed here.

3.1. Therapeutic Context

3.1.1. Disease or condition

Alzheimer's disease (AD) is the most common cause of dementia in the elderly. A recent study estimated 6.9 million persons living with AD dementia in Europe, with an additional 15.2 million persons living with milder cognitive impairment due to AD pathology (Gustavsson et al. 2022). AD and other dementias were reported as the third leading cause of death for people over 70 years of age in the EU, accounting for about 7% of total deaths (IHME 2020). AD is a significant growing socio-economic issue and early diagnosis of AD is estimated to bring socio-economic benefits (Barnett et al. 2014; Dubois et al. 2016).

3.1.2. Available therapies and unmet medical need

The updated International Working Group (IWG) recommendations (Dubois et al. 2021) state that AD can be diagnosed in the presence of both a specific clinical phenotype of AD and evidence of amyloid and tau pathology, supporting its definition as a clinico-biological entity (Dubois et al. 2021).

Three Aβ amyloid PET tracers, i.e., Amyvid (Florbetapir (18F); by the applicant of this MAA; EMEA/H/C/002422; approved in 2013), Vizamyl MA (Flutemetamol 18F; EMEA/H/C/002557; approved in 2014) and Neuraceq (Florbetaben (18F) by Life Molecular Imaging GmbH; EMEA/H/C/002553; approved in 2014) have been approved in the EU for detecting β-amyloid plaque in the brains of patients undergoing evaluations for AD. A negative Aβ PET scan indicates sparse to no neuritic plaques and is inconsistent with a neuropathological diagnosis of AD at the time of image acquisition. However, since amyloid pathology can be present in the absence of AD clinical syndrome, or may not be the cause of the presenting cognitive symptoms, it cannot be used to diagnose AD by itself (Clark et al. 2011, 2012).

Other ways of determining the presence of beta amyloid and tau aggregates include combined amyloid and tau cerebro-spinal fluid (CSF) biomarkers analysis. However, CSF analysis does not inform on the distribution of the NFTs in the brain. Further, lumbar puncture (LP) is not performed in patients, who use contraindicated medications, such as anti-coagulants, have significant spinal/vertebral conditions or surgeries, and lack of capacity to consent. Moreover, a sizable proportion of patients (about 45%) with MCI and dementia refuse to undergo LP (Blazel et al., 2020). Consequently, there is an unmet medical need in these patients, but also in general population with impaired cognitive function, for additional methods of diagnosing the presence of NFTs in the brain.

Other supportive diagnostic tools not directed to the detection of beta-amyloid or tau in the brain are MRI and FDG PET. However, these present non-specific markers of downstream neurodegeneration, and do not provide direct evidence of the underlying pathophysiology (Jack et al. 2018; Dubois et al. 2021).

So far, no imaging diagnostic agent for detection of tau/NFTs in patients undergoing evaluation for AD has been approved in the EU and Tauvid would be the first PET product in this indication.

3.1.3. Main clinical studies

The applicant has submitted two "pivotal" studies (A16 and FR01).

Study A16 was a confirmatory multi-centre, phase III, clinico-pathological study comparing flortaucipir (18F) PET to post-mortem assessment of tau NFT pathology (Braak stage V+VI, or B3 score, Braak 2006) and high level of ADNC (Hyman et al. 2012). In total 156 patients at the end of their lives were enrolled and autopsy samples of 82 subjects were collected (including supplemental set with 16 subjects). Primary analysis was done in 64 subjects: 49 with dementia, 1 with MCI and 14 cognitively normal. Sixteen subjects with dementia were diagnosed non-AD dementia at the baseline and 29 had non-AD tau in the autopsy. Forty one of 66 cases (64 plus 2 frontrunners with interpretable visual reads) were classified as Braak Stages V to VI (B3) NFT pathology; 40 of 41 (97.6%) of these cases had CERAD moderate or frequent amyloid neuritic plaques. Only one of these cases was asymptomatic (cognitively normal) prior to death. In the subjects with Braak Stage III/IV (B2) NFT pathology (n=17), 12 (69.4%) were amyloid positive (CERAD) and 10 (58.8%) were clinically symptomatic.

Visual read criteria were defined as tAD+/tAD++ (corresponding to Braak V and VI, or B3 on autopsy) for "positive" PET. In the autopsy, B3 and "high" ADNC were considered "positive" TS.

Study FR01 was a retrospective reader study. No new data were collected and new evaluation of PET images collected in the studies A16 (n=83 including 1 subject with PET, but without autopsy) and A05C (n=159 in the MCI and AD population) was done to re-assess the diagnostic performance of flortaucipir (18F) PET against histopathological findings in autopsy (n=82) and to analyse inter and intra-reader variability utilising the images collected in the A16 and A05C studies (n=159 with MCI or AD and n=242 pooled population).

In both studies similar image reading methodology with 5 qualified and fully blinded readers analysing the flortaucipir (18F) images was done. Specific training was conducted in the readers prior to the reading procedure in the study FR01. The same visual read criteria and definition of TS were applied as in the A16 study.

3.2. Favourable effects

A16 study: Primary analysis (n=64) against NFT as TS: The sensitivity (95% CI) ranged from 92.3% (79.7, 97.3) to 100.0% (91.0, 100.0) across readers and all readers met success criterion. Median value was 92.3%. Specificity (95%CI) ranged from 52.0% (33.5, 70.0) to 92.0% (75.0, 97.8) and only three of 5 readers met success criteria. For specificity median value (range) was 76.0%.

FR01 Study: Primary analysis (n=82) against NFT as TS: The sensitivity (95% CI) ranged from 87.2% (74.8, 94.0) to 93.6% (82.8, 97.8) across readers and all readers met success criterion. Median value was 89.4%. Specificity (95%CI) ranged from 62.9% (46.3, 76.8) to 91.4% (77.6, 97.0) and only 4 of 5 readers met success criteria. For specificity median value was 77.1%.

Inter-reader agreement in total population (n=241) was 93.8% for overall and ranged from 91.3 to 97.1% when comparing 2 readers. The Fleiss' kappa was 0.87 for overall agreement (95% CI: 0.83, 0.91); $p<0.0001$ for all comparisons. The lower bound of two-sided 95% CI was > 0.6 for all comparisons.

Inter-reader agreement in the target population (MCI and dementia n=159) was 95.1% for overall and ranged from 93.1% to 97.5% for pairs. The Fleiss' kappa was also high: 0.90 for overall agreement; p<0.0001 for all comparisons.

3.3. Uncertainties and limitations about favourable effects

The key limitation of the provided data package is that the pivotal evidence is limited to only one study (A16).

The following limitations have been identified:

- The population included into the A16 study is not fully representative of the target population. Evidence of diagnostic performance in the MCI population is missing. It is uncertain to which extent the data can be extrapolated to broader indication (broader population with dementia and with MCI).

  With their responses, the applicant restricted the indication to "adult patients with cognitive impairment who are being evaluated for AD". This still leads to inclusion of a broader patient population (i.e., patients with various degrees of cognitive impairment from mild cognitive impairment (MCI) to dementia) compared to those included in the pivotal study (i.e., patients with dementia). However, this can be considered acceptable based on the following rationale:

  - Flortaucipir PET detects aggregated tau neurofibrillary tangles (NFTs) at their progressed stage (B3, or Braak stages V and VI), i.e., when NFTs spread to the neocortex. Thus, the product detects neuropathological changes in the brain, whereas cognitive impairment (MCI to dementia) reflects clinical symptomatology of AD.
  - Neuropathological changes and clinical symptoms do not strictly correlate. While published evidence suggests that changes in cognition (clinical picture) correlate with aggregation of NFTs (neuropathological change) to some extent, various stages of NFT aggregation (B1, B2, B3) and levels of impairment in cognition (Mild to dementia) do not fully match. NFTs of the B3 level have been observed in patients with various levels of cognitive impairment, including those with MCI.

  It is therefore plausible that the flortaucipir PET can detect NFTs in the neocortex (i.e., detection of neurophysiological changes in the brain) regardless of clinical symptomatology. This speaks in favour of the broader indication claimed. In addition, the product is to be used only in conjunction with clinical and other diagnostic evaluations. Therefore, a diagnosis of AD is not made on the basis of a positive flortaucipir PET alone.

  Additionally, all relevant limitations (lack of data in the MCI population, uncertainty of data extrapolation, lack of expected benefit in the amyloid-negative MCI patients, etc.) have been reflected in the SmPC (section 4.4) with the respective cross-reference in section 4.1. to reduce the risk of diagnostic failures.

- The visual read criteria developed and applied to detect high-degree changes in the brain (NFT aggregation and neuropathological change) in the A16 study do not allow detection of earlier stage of the NFT aggregation. Flortaucipir PET scan does not rule out presence of AD. This limitation has been clearly described in the PI.

- The success criteria defined for the study (threshold of 50% for the lower limit of 95%CI to pass and requirement that only 3 of 5 readers should fulfil this criterion) are too weak. However, sensitivities across readers were considerably higher than the predefined threshold.

- Variable specificity creates uncertainty and suggests difficulties of the test to correctly exclude AD by identifying truly negative patients. Since specificity is ”poor” the positive predictive value is reduced, which means that a positive test result may incorrectly diagnose a patient as having AD. This limitation has been highlighted in the SmPC, and use of Tauvid PET is recommended only in conjunction with other diagnostic methods (see section 4.1 in the SmPC).
- The patient could have an alternative disease status. Intra-reader agreement in the FR01 study showed suboptimal levels of agreement, suggesting that the test may not be reliable. These limitations should be addressed with proper training of the healthcare professionals.
- Uptake of flortaucipir has been observed in various conditions and situations (non-AD tauopathies, older age, Dawn Syndrome, etc.). Topographical and quantitative assessments are claimed to be effective to distinguish AD from non-AD taupathies. However, no recommendations for quantitative measurements have been included in the SmPC as a validated methodology of quantitative measurements does not exist. Potential mistakes in image interpretation have been described in the SmPC to adequately inform healthcare professionals. Also, Tauvid PET is to be used in conjunction with other diagnostic methods. Further, the product information clearly states that Tauvid images should only be interpreted by readers trained in the interpretation of PET images with flortaucipir (section 4.4 of the SmPC). Careful monitoring of new cases of errors and pitfalls in image interpretation is recommended (e.g., via PSUR) in the post-authorisation phase.
- The key limitation of the FR01 study is that this was a retrospective study and can provide supportive data only.

Overall, although the data package on diagnostic performance is limited, it is considered sufficient by the CHMP.

3.4. Unfavourable effects

The overall safety population from the 28 clinical trials included in the safety analysis contains data of 4652 subjects who received at least one dose; of these 63 subjects (1.35%) discontinued due to AEs. Of the 63 subjects, 2 (0.043%) discontinued due to an AE that occurred within 2 days of flortaucipir administration (TEAE) and 22 (0.5%) discontinued due to death.

In the pooled analysis with 303 subjects with 421 TEAEs, the most frequent AEs occurring in ≥0.2% of participants were headache (0.9%), injection site pain (0.6%), diarrhoea (0.3%), blood pressure increased (0.3%), dizziness (0.2%), hypertension (0.2%), and nausea (0.2%). In general, the rates of TEAEs were considered consistent with what would be expected for a mostly elderly population with varying levels of CI.

In an uncontrolled clinical study setting mean change in QTcF in subjects ≥65 years was 5.34 msec (CN 5.71 and CI 5.24 post dose readings). In subjects <65 years old, the mean change was 4.56 msec.

Small but statistically significant increases in mean systolic and diastolic blood pressure were seen immediately post dose and at the end-of-scan time point. Twenty-nine potentially clinically significant changes in systolic blood pressure or diastolic blood pressure were reported, including quite marked increases ranging from 20-78 mmHg in systolic pressure and 15-68 mmHg in DBP. No clinically meaningful or statistically significant correlation was observed between mass dose of flortaucipir and systolic or diastolic blood pressure at either the postdose or end-of-scan time point.

Small, statistically but not clinically significant changes were seen in multiple ECG parameters at the postdose and end-of-scan time points. Overall, the mean increases in QTcB and QTcF at the end-of-

scan time point were 2.34 and 5.14 msec, respectively. Increase in QTcF by >30 msec and/or above upper limit of normal was reported in 1.8% and 1.7% of the cases after flortaucipir application. No cases of QTcF increase by >60 msec or above 500 msec was reported. No statistically or clinically significant changes in QTc were noted at the time of peak plasma concentration (immediately after injection). No TEAEs related to QT interval prolongation or ventricular arrhythmias were reported.

3.5. Uncertainties and limitations about unfavourable effects

Safety data have been collected in the trials without control that complicates proper assessment of the safety profile.

The four clinical studies included in the ECG analysis excluded patients with current clinically significant cardiovascular disease or clinically significant abnormalities on screening ECG (including but not limited to QTc>450 msec) and patients with a history of additional risk factors for Torsades de Pointes (TdP) (e.g., heart failure, hypokalaemia, family history of Long QT syndrome). It is uncertain to which extent the non-critical findings of QTc in the studies can be extrapolated to broad patient population. However, newly conducted non-clinical hERG study showed large safety margin of 300 that allows to conclude negligible risk of QT prolongation.

It is considered likely that past medical history of hypertension in some cases, timing of blood pressure measurements and especially the imaging procedures per se may have caused the PCS cases of SBP and DBP and TEAEs related to elevated BP, but a contribution of flortaucipir cannot be fully excluded.

It is currently unclear, whether sporadic reports of skin rashes, hypotension, flushing, swelling face, could be the cases of hypersensitivity reactions, which have not been reported as ADRs in the SmPC thus far.

3.6. Effects Table

Table 62. Effects table for Tauvid in the agreed indication (data cut-off: 3/02/2023)

Effect	Short Description	Unit	Treatment	Control	Uncertainties/ Strength of evidence	References
Favourable Effects						
Sensitivity/ specificity – vs. NFT (n=64)	Diagnostic performance of the visual read criteria (tAD+/tAD++ being "positive" PET and tAD- "negative" PET) against TS defined as NFT in autopsy (B3 score – "positive" TS and B0-B2 – "negative" TS)	% (95% CI)	Sensitivity (95% CI) ranged from 92.3% (79.7, 97.3) to 100.0% (91.0, 100.0) across readers. Specificity, the lower bound of the 95% CI was >50.0% for 3 of the 5 readers; for these 3 readers, the specificity (95% CI) ranged from 76.0% (56.6, 88.5) to 92.0% (75.0, 97.8). For all readers specificity ranged from 52.0% (33.5, 70.0) to 92.0% (75.0, 97.8)	NA	Adequately designed study with an acceptable gold standard reference/TS. Data regarded confirmatory. Specific and limited patient population. Data do not cover the target population. Data cannot be fully extrapolated. Number of uncertainties in regards to the data handling/blinding procedures.	Study A16

Effect	Short Description	Unit	Treatment	Control	Uncertainties/ Strength of evidence	References
Inter-reader agreement (n=105)	Assessment of the agreement across 5 readers		Overall agreement among the readers was 89.9% with a Fleiss's kappa of 0.80 (CI: 0.74, 0.86)	NA	Acceptable assessment procedures, but restricted image set. Uncertainty regarding extrapolation to subjects with MCI.	A16
Sensitivity/ specificity – vs. NFT (n=82)	Diagnostic performance of the visual read criteria (tAD+/tAD++ being "positive" PET and tAD- "negative" PET) against TS defined as NFT in autopsy (B3 score – "positive" TS and B0-B2 – "negative" TS)	% (95% CI)	Sensitivity (95% CI) ranged from 87.2% (74.8, 94.0) to 93.6% (82.8, 97.8). For specificity, the lower bound of the 95% CI was ≥50% for 4 of the 5 readers; for these 4 readers, the specificity (95% CI) ranged from 77.1% (61.0, 87.9) to 91.4% (77.6, 97.0). For all readers specificity ranged from 62.9% (46.3, 76.8) to 91.4% (77.6, 97.0)	NA	Larger sample size than in the A16 study. Retrospective study design. Data regarded exploratory. Weak evidence	FR01
Inter-reader agreement (mixed population) (n=242)	Assessment of the agreement across 5 readers		Overall agreement >90% a Fleiss's kappa of 0.87 (95% CI: 0.83-0.91).		Larger sample size than in the A16 study with broad population resembling the targeted indication. Retrospective study design. Data regarded exploratory. Weak evidence	FR01
Inter-reader Agreement (MCI population) (n=159)	Assessment of the agreement across 5 readers		0.90 Fleiss's kappa (95% CI: 0.85-0.95)		Larger sample size than in the A16 study with the MCI population that is part of the targeted indication. Retrospective study design. Data regarded exploratory. Weak evidence	FR01
Unfavourable Effects						
Most frequent TEAEs	% calculated in the pooled population of 303 subjects.	%	Headache (0.9%), injection site pain (0.6%), diarrhoea (0.3%), blood pressure increased (0.3%), dizziness (0.2%), hypertension (0.2%), and nausea (0.2%)	NA	Generally, large safety database over 4000 subjects. Absence control. Moderately strong evidence	Pooled population (n=303)

Abbreviations: AD = Alzheimer's disease; CI = confidence interval; MCI = mild cognitive impairment; n = number of patients with flortaucipir PET scans that were read; NFT = neurofibrillary tangles; PET = positron emission tomography.
Notes: a Cox proportional hazard model adjusted for baseline age, ANART, and baseline CDR-SB score.
b Risk ratios adjusted for treatment arm, age, years of education, and baseline test score.

3.7. Benefit-risk assessment and discussion

3.7.1. Importance of favourable and unfavourable effects

Although, measurement of tau in liquor or plasma is possible, there is no approved imaging tool that would visualise location of NFTs in the brain, to support the diagnostic work-up of AD. Further, lumbar puncture (LP) that is an established method for collecting CSF for analysis of tau, is contraindicated in patients who use contraindicated medications such as anti-coagulants, have significant spinal/vertebral conditions or surgeries, or lack capacity to consent. Moreover, a sizable proportion of patients (about 45%) with MCI or dementia refuse to undergo LP (Blazel et al., 2020). Thus, there is an unmet medical need for an additional diagnostic imaging tool that could contribute to the diagnostic work-up of AD and Tauvid would be the first such product.

One pivotal and one supportive study have been presented to evaluate the diagnostic performance and inter-reader reliability of flortaucipir (^{18}F) PET in detection of NFTs. Both studies met the primary endpoint criteria showing acceptable level of diagnostic efficacy and reliability in the restricted study population.

The submitted evidence allows to conclude with acceptable certainty that flortaucipir (^{18}F) PET (utilising the recommended visual read criteria) is able to detect AD-related neuropathological changes, concretely, progressed aggregation of NFTs in the neocortex (specific sections). It is plausible to assume that this diagnostic ability can be exerted regardless of the status of cognition, which is the clinical diagnosis and does not necessarily coincide with neuropathological changes in the brain. Therefore, inclusion of the MCI population in the indication is acceptable, provided that the lack of adequate data in this population and the uncertainty of diagnostic efficacy of Tauvid is adequately reflected in the SmPC.

Variable specificity of flortaucipir (^{18}F) PET observed in the pivotal study creates uncertainty, as this suggests the difficulty of the test to correctly exclude AD by identifying truly negative patients. In other words, positive PET may represent "false positive" result and AD may be over-diagnosed.

However, the indication states that the product should be used in conjunction with other diagnostic tools only and the identified limitations/weaknesses are addressed in the SmPC to adequately inform physicians. These measures are expected to improve the accuracy of the overall diagnostic work-up of patients with suspected AD. Additionally, the SmPC has been adapted to better specify the limitations of use for flortaucipir (^{18}F).

The safety profile of flortaucipir (^{18}F) appears overall benign. The unfavourable effects were limited and of limited clinical relevance. Observed changes in QTc are considered clinically not significant.

3.7.2. Balance of benefits and risks

There is an unmet medical need for an imaging tool that can accurately detect the presence of NFTs and their distribution in the brain of patients with cognitive impairment being evaluated for the presence for AD. Tauvid has shown good sensitivity in detecting NFTs in the neocortex. The lower and more variable specificity may be acceptable since Tauvid is to be used as adjunct to other diagnostic tools and therefore a diagnosis of AD is not made on the basis of a Tauvid PET scan alone.

The key limitations and deficiencies identified are:

- The confirmatory efficacy data are limited to a single and relatively small study.

- Key efficacy information has been derived from a specific/restricted population at end of life and there are some uncertainties regarding extrapolation of study results, especially to patients with MCI. However, flortaucipir is expected to detect aggregation of NFTs in the neocortex regardless of the status of cognition. Therefore, subjects with MCI can be included in the indication with a warning statement that test accuracy may be lower in these subjects.
- The "validated" and recommended visual read criteria are unable to detect early phases of NFT aggregation and negative flortaucipir (^{18}F) PET does not exclude presence of AD.
- Validated methodology of quantitative measurements as a support to the visual read has not been provided and, thus, a respective claim has been removed from the indication.
- Data suggest that flortaucipir (^{18}F) PET may have difficulty in correctly diagnosing the absence of AD (variable specificity).
- Intra-reader variability was rather high.

During the assessment the applicant was initially recommended to put in place adequate post-authorisation measures to limit the effects of these uncertainties, however after further consideration it was agreed to address these limitations and uncertainties as recommendations and including appropriate warnings in the SmPC.

It is considered that flortaucipir (^{18}F) PET can bring benefit in the diagnostic work-up of patients with AD, if properly applied. The indication is acceptable.

There are no relevant clinical safety concerns with the use of flortaucipir (^{18}F) that would preclude its use.

Therefore, the benefits of flortaucipir (^{18}F) are considered to outweigh the risks.

3.7.3. Additional considerations on the benefit-risk balance

None

3.8. Conclusions

The overall benefit/risk balance of Tauvid is positive, subject to the conditions stated in section 'Recommendations'.

4. Recommendations

Outcome

Based on the CHMP review of data on quality, safety and efficacy, the CHMP considers by consensus that the benefit-risk balance of Tauvid is favourable in the following indication(s):

> *This medicinal product is for diagnostic use only.*
>
> *Flortaucipir (18F) is a radiopharmaceutical indicated for positron emission tomography (PET) imaging of the brain to assess the neocortical distribution of aggregated tau neurofibrillary tangles (NFTs) in adult patients with cognitive impairment who are being evaluated for Alzheimer's disease (AD). Flortaucipir (18F) is an adjunct to clinical and other diagnostic evaluations.*
>
> *For limitations of use, see sections 4.4 and 5.1.*

The CHMP therefore recommends the granting of the marketing authorisation subject to the following conditions:

Conditions or restrictions regarding supply and use

Medicinal product subject to restricted medical prescription (see Annex I: Summary of Product Characteristics, section 4.2).

Other conditions and requirements of the marketing authorisation

- **Periodic Safety Update Reports**

The requirements for submission of periodic safety update reports for this medicinal product are set out in the list of Union reference dates (EURD list) provided for under Article 107c(7) of Directive 2001/83/EC and any subsequent updates published on the European medicines web-portal.

Conditions or restrictions with regard to the safe and effective use of the medicinal product

- **Risk Management Plan (RMP)**

The marketing authorisation holder (MAH) shall perform the required pharmacovigilance activities and interventions detailed in the agreed RMP presented in Module 1.8.2 of the marketing authorisation and any agreed subsequent updates of the RMP.

An updated RMP should be submitted:

- At the request of the European Medicines Agency;
- Whenever the risk management system is modified, especially as the result of new information being received that may lead to a significant change to the benefit/risk profile or as the result of an important (pharmacovigilance or risk minimisation) milestone being reached.

Conditions or restrictions with regard to the safe and effective use of the medicinal product to be implemented by the Member States

N/A

New active substance status

Based on the CHMP review of the available data, the CHMP considers that Flortaucipir (^{18}F) is to be qualified as a new active substance in itself as it is not a constituent of a medicinal product previously authorised within the European Union.