CENTER FOR DRUG EVALUATION AND RESEARCH

APPLICATION NUMBER:

022454Orig1s000

STATISTICAL REVIEW(S)

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research
Office of Translational Sciences
Office of Biostatistics

Statistical Review and Evaluation

CLINICAL STUDIES

NDA/Serial Number:	22-454 /N0000, 22-454 /N0005
Drug Name:	Datscan ([I123]ioflupane injection) for intravenous use
Indication(s):	Visualization of the dopamine transporter (DaT) distribution within the striatum by single photon emission computed tomography (SPECT) imaging in patients presenting with symptoms or signs suggestive of dopaminergic neurodegeneration.
Applicant:	GE Healthcare
Date(s):	NDA receipt: 9 March 9, 2009 Review: 26 August 2009 PDUFA: 9 September 2009
Review Priority:	Priority
Biometrics Division:	Division of Biometrics 5
Statistical Reviewer:	Mark Levenson, Ph.D.
Concurring Reviewers:	Jyoti Zalkikar, Ph.D. Aloka Chakravarty, Ph.D.
Medical Division:	Division of Medical Imaging and Hematology Products
Clinical Team:	Phillip Davis, M.D. (Medical Officer) Louis Marzella, M.D. (Medical Team Leader) Dwaine Rieves, M.D. (Medical Division Chief)
Project Manager:	James Moore

Keywords: sensitivity, specificity, neurology, radiopharmaceutical

Table of Contents

LIST OF TABLES

LIST OF ABREVIATIONS

AD: Alzheimer's disease
DLB: Dementia with Lewy bodies
ET: Essential tremor
DaT: Dopamine transporter
MSA: Multiple system atrophy
PD: Parkinson's disease
PS: Parkinson's syndrome
PSP: Progressive supranuclear palsy
SDD: Striatal dopaminergic deficit
SOT: Standard of truth
VaD Vascular dementia

1. EXECUTIVE SUMMARY

GE Healthcare (Sponsor) has submitted a New Drug Application for Datscan ([I123]ioflupane injection) for intravenous use as a medical imaging agent. The NDA was reviewed under the indication:

> Datscan is a radiopharmaceutical containing [123I]ioflupane, indicated for visualization of the dopamine transporter (DaT) distribution within the striatum by single photon emission computed tomography (SPECT) imaging in patients presenting with symptoms or signs suggestive of dopaminergic neurodegeneration.

The presence or absence of the DaT protein may be an indirect marker of the functioning of the nigrostriatal neurons. The determination of the functioning of these neurons may assist in the diagnosis of certain neurological diseases, such as Parkinson's disease, essential tremor, dementia with Lewy bodies, and Alzheimer's disease.

In the NDA submission, the Sponsor had originally proposed an indication for detecting loss of functional nigrostriatal dopaminergic neurons. FDA did not feel that such an indication could be supported by the submission and suggested the above indication as the basis of the review. The Sponsor agreed to the revised indication for the basis of the review.

1.1 Conclusions and Recommendations

Overall, the Phase III program does not support for the approval of Datscan. The Phase III program did not meet the standard of providing two adequate and well controlled trials. Additionally, there are questions concerning the validity of the standards of truths in two studies and the image reading in one study.

The clinical utility of Datscan is questionable. The performance measures of Datscan were notably lower in patients with less certain disease states. It is likely that these patients would have the greatest need for such a product.

The label should make clear the context of the use of the drug. The proposed indication is not a diagnostic indication. However, the assessment that the drug provides is only useful in the context of making clinical diagnosis. Since the loss of functional nigrostriatal dopaminergic neurons is associated with multiple diseases, it is important to state that differentiation between certain diseases is not possible. For example, the diseases dementia with Lewy Bodies and Parkinson's with dementia cannot be distinguished with Datscan, because they are both associated with loss of functional nigrostriatal dopaminergic neurons.

A review of the adverse events from the Phase III studies did not reveal any safety issue.

1.2 Brief Overview of Clinical Studies

For the confirmation and support of efficacy, the Sponsor provided three Phase III studies: DP008-003, PDT3004, and PDT301. Studies DP008-003 and PDT3004 concerned patients with movement disorders. Study PDT301 concerned patients with dementia. The Sponsor also provided an independent investigator exploratory study ("Walker" study) of dementia patients.

None of these studies were conducted under a United States Investigative New Drug application. Each of these studies was conducted entirely in Europe.

The original objective of these studies did not directly relate to the proposed indication of the visualization of the DaT protein or the loss of functioning nigrostriatal neurons. The objectives of the three Phase III studies (DP008-003, PDT3004, and PDT301) concerned the diagnostic potential of Datscan images for various classes of diseases. Studies DP008-003 and PDT3004 addressed patients with movement disorders. The primary diagnostic distinction for these studies was between Parkinson's syndrome diseases and other movement disorders principally essential tremor. Studies PDT301 addressed patients with dementia. The primary diagnostic distinction for this study was between dementia with Lewy Bodies and other forms of dementia, such as Alzheimer's disease and vascular dementia. For all studies, the diagnostic distinction separates diseases associated with degeneration of nigrostriatal dopaminergic neurons and disease without such degeneration. Through the association of the degeneration with the various diseases, the Sponsor proposed to use the diagnostic studies to support the DaT protein visualization indication.

Study PDT301 was the only study with a prespecified statistical analysis plan and prespecified statistical hypotheses. The standard of truth was central blinded clinical diagnosis based on information as baseline. However, the clinical diagnosis for these dementia diseases at baseline as a standard of truth may not have high validity. Both the standard of truth determination and the image reads appeared adequately blinded.

Study DP008-003 had a prespecified statistical analysis plan but no prespecified statistical hypotheses. The original plan was modified for the US submission in several important ways, in some cases to make use of more rigorous data. However, the post-hoc change in plan does diminish the statistical rigor of the analysis. The original primary statistical analysis was based on unblinded reads, whereas the revised analysis included the use of five blinded reads. These blinded readers were chosen among the site investigators.

Study PDT3004 had a prespecified statistical analysis plan but no prespecified statistical hypotheses. The study concerned movement disorder patients with early features of Parkinson's syndrome and not existing diagnoses as in study DP008-003. The standard of truth was central blinded clinical diagnosis based on clinical information and video assessment up to 36 months. The image review was based on three central blinded readers. Both the standard of truth determination and the image reads appeared adequately blinded.

The Walker study was an independent investigator, exploratory, single site study. It was a follow-up study with a small sample of 22 patients. No statistical analysis plan was available for the study. Because of the small sample size, the estimates were very variable. The study evaluated baseline clinical diagnosis and baseline Datscan imaging for distinguishing between dementia with Lewy Bodies from other dementia. The standard of truth was neuropathology disease diagnosis.

1.3 Statistical Issues and Findings

Table E1 provides the sensitivities and specificities estimates and confidence intervals for the three Sponsor studies.

Study PDT301 was the only study with prespecified statistical hypotheses. For this study, the primary endpoints were sensitivity and specificity for distinguishing probable dementia with Lewy bodies from non-dementia with Lewy bodies. The point estimates for the three blinded readers for the sensitivities ranged from 75 to 80% and for the specificities ranged from 89 to 91%. The prespecified thresholds for sensitivity and specificity were 65% and 75% respectively. The study statistically demonstrated that the endpoints exceeded these thresholds for 2 out of the 3 central blinded readers and missed the threshold for sensitivity for the third reader by a very small amount. However, as noted, the clinical diagnosis for these dementia diseases at baseline as a standard of truth may not have high validity. Using a standard of truth at 12 month did not improve the performance measures. The sensitivities for distinguishing probable or possible dementia with Lewy bodies from non-dementia with Lewy bodies were notably lower (point estimate range: 61 to 64%).

Study DP008-003 had a prespecified statistical analysis plan but no prespecified statistical hypotheses. Overall, the point estimates of sensitivity ranged from 92% to 97% and for specificity from 74% to 96%, based on the five blinded readers. These sensitivities may have been high because the patient population consisted of patients with existing diagnoses.

Study PDT3004 had a prespecified statistical analysis plan but no prespecified statistical hypotheses. Overall, the point estimates of sensitivity were all approximately 78% and the specificities were all approximately 97%. These sensitivity estimates were notably lower than those for study DP008-003, perhaps because the patient population had less clear disease state.

In the Walker study, the estimate for sensitivity of Datscan was 78% (95% CI: 40%, 97%) and the estimate for specificity was 85% (95% CI: 55%, 98%). For the baseline clinical diagnosis, the estimate for sensitivity was same as for the Datscan image, 78% (95% CI: 40%, 97%). The estimate of the specificity of baseline clinical diagnosis was notably lower, 46% (95% CI: 19%, 75%). The low specificity of the baseline clinical diagnosis calls into question the use of clinical diagnosis of dementias as a standard of truth in study PDT301.

Likely because all studies were conducted in Europe, the patients in the three Sponsor studies were overwhelmingly Caucasian and may not be representative of the US patient population. Because of the small sample size, difference among age and gender subgroups could not be determined.

On a population-level, the inter-reader variability of the Datscan image reads appeared small. Additionally, in the three Sponsor studies, the number of non-evaluable images was small.

There were 35 serious adverse event and 5 deaths, none of these events were related to the drug according to the investigator. Four of the 5 deaths occurred in a long-term follow-up period and the remaining death came from a pneumonia event.

The statistical rigor of the Phase III program was not strong. According to the FDA regulations and guidance, generally at least two adequate and well-controlled studies are needed to establish effectiveness. Only one study (PDT301) had prespecified statistical hypothesis and this study had a questionable standard of truth. For study PD008-003, the statistical plan had notable deficiencies. The standard of truth evaluation was not strong and the primary analysis was based on unblinded reads of the images.

The clinical utility needs to be evaluated with respect to both the studied patient population and the resulting performance measures. The performance of Datscan varied between and within the studies depending on the studied patients. Study PDT301 included patients with probable and possible dementia with Lewy bodies. The inclusion of the possible dementia with Lewy bodies patients in the analysis lowered the performance measures. A similar result was apparent in the comparison of studies PD008-003 and PDT3004. Both these studies concerned patients with movement disorders. The study PD008-003 patient population consisted of patients with existing diagnoses, whereas the study PDT3004 patient population consisted of patients with early signs of disease. The performance measures of Datscan were notably lower in the study of patients with early signs of disease.

Table E1: Sponsor Studies: Sensitivity and Specificity for Primary Efficacy.

	Sensitivity (95 % CI)	Specificity (95 % CI)
DP008-003		
Reader A N=185	93.0 (87.9, 96.5)	96.3 (81.0, 99.9)
Reader B N=185	96.8 (92.8, 99.0)	74.1 (53.7, 88.9)
Reader C N=185	96.2 (91.9, 98.6)	85.2 (66.3, 95.8)
Reader D N=185	92.4 (87.1, 96.0)	92.6 (75.7, 99.1)
Reader E N=185	94.3 (89.5, 97.4)	92.6 (75.7, 99.1)
PDT3004		
Reader A N=102	77.5 (66.0, 86.5)	96.8 (83.3, 99.9.)
Reader B N=99	77.9 (66.2, 87.1)	96.8 (83.3, 99.9)
Reader C N=101	78.6 (67.1, 87.5)	96.8 (83.3, 99.9)
PDT301		
Reader A N=216	79.8 (69.2, 88.0)	91.2 (85.2, 95.4)
Reader B N=216	75.3 (64.2, 84.4)	88.5 (82.0, 93.3)
Reader C N=218	80.3 (69.9, 88.3)	90.5 (84.3, 94.9)

Notes: See text related to and notes on Table 6.

2. INTRODUCTION

2.1 Overview

GE Healthcare (Sponsor) has submitted a New Drug Application (NDA) for Datscan ([I123]ioflupane injection) for intravenous use as a medical imaging agent. The NDA contained the proposed indication:

> Datscan is a radiopharmaceutical containing [123I]ioflupane, indicated for detecting loss of functional nigrostriatal dopaminergic neurons by single photon emission computed tomography (SPECT) imaging in patients presenting with symptoms or signs suggestive of dopaminergic neurodegeneration

In the FDA briefing package for the August 11, 2009 Peripheral and Central Nervous System Drugs Advisory Committee Meeting, FDA proposed another indication. During a teleconference on 5 August 2009 with the Sponsor concerning the Advisory Committee meeting, the Sponsor and FDA agreed that the revised indication would be the basis of Committee discussion and the review of the drug. The revised indication is

> Datscan is a radiopharmaceutical containing [123I]ioflupane, indicated for visualization of the dopamine transporter (DaT) distribution within the striatum by single photon emission computed tomography (SPECT) imaging in patients presenting with symptoms or signs suggestive of dopaminergic neurodegeneration.

Datscan is a radionuclide tracer version of a ligand intended to bind to the dopaminergic transporter (DaT) protein in the nigrostriatal region of the brain. The DaT protein participates in dopamine reuptake. The DaT protein may be an indirect marker of the functioning of the nigrostriatal neurons. The determination of the functioning of these neurons may assist in the diagnosis of certain neurological diseases. Loss of functioning nigrostriatal neurons is associated with certain neurologic diseases such as Parkinson's disease and dementia with Lewy bodies but not other neurologic diseases, which may have similar clinical symptoms. The sponsor used the term striatal dopaminergic deficit (SDD) to refer to loss of functional nigrostriatal dopaminergic neurons.

The FDA guidance--Developing Medical Imaging Drugs and Biological Products--provides a framework for evaluating the Datscan and describes various types of medical imaging agent indication. The binding to the DaT protein may be considered a biochemical assessment. The biochemical assessment may be common to several diseases and may not be diagnostic of any particular disease.

According to the Guidance, to establish efficacy, the clinical trials should compare the agent to a reference with high validity. Subjective endpoints such as diagnostic confidence should not be used as primary endpoints. The guidance also recommends that the agent be evaluated with appropriate representation of subjects and that the particular inclusion/exclusion criteria are justified.

The Guidance emphasizes that the assessment should clinically useful. Clinical usefulness is described as helping in making accurate diagnoses, contributing to beneficial clinical outcomes, or providing accurate prognostic information. In some cases, the clinical utility may be self evident.

For the confirmation and support of efficacy, the sponsor provided four Phase III or IV studies: DP008-003, PDT3004, PDT301, and PDT408. The sponsor also provides an independent investigator, exploratory, single-site study: "Walker" study. None of these studies were conducted under a United States Investigative New Drug application. Each of these studies was conducted entirely in Europe. The study PDT408 was not considered in this review, because (1) the study generally evaluated clinical management without consideration to clinical utility and (2) the standard of truth evaluation for this study was not blinded to the Datscan image results. The Walker study will be considered separately because its small size limits statistical inference.

For the US submission, the sponsor created new versions of the clinical study reports, which primarily affected the focus of the reports. However, there were some changes of statistical consequences. The changes will be addressed in the review of the individual studies. The sponsor also created CDISC versions of the study and analysis datasets. Unless otherwise noted, all statements and findings refer to the original European clinical study reports, protocols, and amendments.

Table 1 summarizes the key statistical aspects of the three Sponsor studies, DP008-003, PDT3004, and PDT301. Each study was a non-comparative in which Datscan imaging results were evaluated relative to a clinical diagnostic standard of truth. The objectives of the three studies did not directly relate to the proposed indication for Datscan, visualization of the dopamine transporter (DaT) distribution. Studies DP008-003 and PDT3004 concerned movement disorder diseases. Study PDT301 concerned dementia diseases.

The objectives concerned the diagnostic potential of Datscan for classes of diseases. For example, the objective of study PDT3004 was

> To determine the predictive value of Datscan SPECT to differentiate between subjects with early features of Parkinsonism, and other causes of tremor (mainly essential tremor), and healthy volunteers.

The sponsor connected the objectives of three studies with the proposed by associating each disease with the presence or absence of striatal dopaminergic deficit (SDD). Thus, clinical diagnoses of the various diseases acted as the standard of truth for evaluating the Datscan assessment of the dopamine transporter (DaT) distribution.

Clinical diagnoses of diseases can serve as a standard of truth for a biochemical indication. However, the diagnosis criteria need to be well validated and the diseases need to have good association with the biochemical assessment. To review this important aspect of the review of the studies, this review presents a succinct discussion of the clinical diagnoses, taken chiefly from the clinical review and a consultation from the Division of Neurology Products (19 May 2009). More information is available in these documents.

The studies involve several neurological diseases including movement disorder diseases and dementia diseases. The movement disorder diseases include Parkinson's disease (PD), essential tremor (ET), multiple system atrophy (MSA), and progressive supranuclear palsy (PSP). Both MSA and PSP are considered Parkinson' syndrome (PS) diseases. The dementia diseases included dementia with Lewy bodies (DLB), Alzheimer's disease (AD), and vascular dementia (VaD). Some of the diseases are associated with the of loss of functional nigrostriatal dopaminergic neurons. These include PD, MSA, PSP, and DLB but not ET, AD, and VaD. The degree of loss varies among the diseases and the time of onset.

Other factors such as medication use may affect the DaT protein. This may include the drug levodopa used in the treatment of Parkinson's disease.

The validity of the clinical diagnoses for these diseases varies by disease, criteria, and time from onset of symptoms. The accuracy of Parkinson's disease diagnosis may good if patients are followed for three years and changes in symptoms are observed. Dementia with Lewy bodies is a difficult to diagnose, particularly as the initial appearance of symptoms. Standard clinical diagnosis criteria for the standard of truth for the studies DP008-003 and PDT3004 have been revised since the conduct of the studies. For these diseases, neuropathology at autopsy provides definitive standard of truth for the determination of the diseases.

Table 1: Key Statistical Aspects of Study Designs.

	DP008-003	PDT3004	PDT301
Study Period	8/1997 – 2/1998	1/1999 – 6/2005	11/2003 – 6/2006
Population	Movement disorders: with diagnosis of PD, MSA, PSP, ET; HV	Movement disorders: early signs of Parkinsonism (PD, ET, other); HV	Dementia: probable DLB, possible DLB, AD, VaD
Centers	6 (all European)	10 (all European)	40 (all European)
Primary SOT; Time	Local, based on referral and baseline diagnosis; baseline	Central, based on video evaluations up to 36m; 36m	Central, based on baseline evaluation; baseline
Primary Image Reads	Local unblinded	Local and 3 central, independent, blinded	3 central, independent, blinded
Other Image Reads	5 independent blinded[1]		
Primary Diagnostic Distinction	PS v. non-PS (excl. HV)	Probable PD, Possible PD v. ET, other (incl. HV)	Probable-DLB v. non-DLB
Primary Endpoints	Sens., spec.	Sens., spec., PPV, NPV, acc.	Sens., spec.
Primary Statistical Analysis	Lower limit of one-sided 95% CI. No thresholds prespecified.	Lower limit 95% CI (not known if one- or two-sided). No thresholds prespecified.	Tests: Sens. > 0.65 Spec. > 0.73. α=0.05 two-sided

Notes:
SOT: Standard of truth, DLB: dementia with Lewy bodies, AD: Alzheimer's disease, VaD: vascular dementia, PD: Parkinson's disease, ET: essential tremor, MSA: multiple system atrophy, PSP: progressive supranuclear palsy, PS: Parkinson's syndrome (PD, MSA, PSP). Sens.: sensitivity; spec.: specificity; PPV: positive predictive value; NPV: negative predictive value; acc.: accuracy.
1. Blinded readers chosen from site investigators

The Walker study consisted of two components: the initial cross-sectional component and a follow-up longitudinal component. Statistical analysis plans have not been obtained for the study and it is not clear how much of the analyses were prespecified. The initial cross-sectional component compared ioflupane image ratios among clinically diagnosed patient groups. The patient groups included DLB, Alzheimer's disease, Parkinson's disease, cortico-basal degeneration, and healthy volunteers. Patients were imaged and clinically evaluated at baseline. The study initiated in June 1996 and the cross-sectional component, including all patient recruitment and baseline diagnosis, was completed by November 1999.

In the follow-up component, some patients were clinically evaluated yearly and underwent post-mortem neuropathological examination. The reporting for the longitudinal component focused

on the dementia patients, DLB and Alzheimer's disease. At present, there are 22 such patients with neuropathological results among the dementia patients. As in the case of the other studies, the standard of truth was not directly related to the proposed indication but to disease assessment.

2.2 Data Sources

The data sources for this review included the protocols, clinical study reports and amendments, individual study analysis datasets, and data tabulations from the Sponsor and Walker studies submitted in NDA 22-454 Sequence 0000, the response to FDA request for information submitted in NDA 22-454 Sequence 0005, and the response to the FDA questions concerning the "mismatch" analysis in study DP008-003 submitted in NDA 22-454 Sequence 0008.

For the US submission, the Sponsor converted the original study data from the supportive and confirmatory studies to CDSIC SDTM v. 3.1.1. In the conversion, adverse events were coded in MedDRA 11.0 and medications were coded in WHO Drug v. 20080301. From the CDISC SDTM study data, analysis files for the principal studies were created using CDISC ADaM methodology.

Study DP008-003
Basic subject data:
\\Cdsesub1\EVSPROD\NDA022454\0000\m5\datasets\dp008-003-us\analysis\adsl.xpt
Standard of truth and image evaluation data:
\\Cdsesub1\EVSPROD\NDA022454\0000\m5\datasets\dp008-003-us\analysis\addiag.xpt
Adverse event data:
\\Cdsesub1\EVSPROD\NDA022454\0000\m5\datasets\dp008-003-us\analysis\adae.xpt

Study PDT3004
Basic subject data:
\\Cdsesub1\EVSPROD\NDA022454\0000\m5\datasets\pdt304-us\analysis\adsl.xpt
Standard of truth and image evaluation data:
\\Cdsesub1\EVSPROD\NDA022454\0000\m5\datasets\pdt304-us\analysis\addiag.xpt
Adverse event data:
\\Cdsesub1\EVSPROD\NDA022454\0000\m5\datasets\pdt304-us\analysis\adae.xpt
Detailed diagnosis data:
\\Cdsesub1\EVSPROD\NDA022454\0000\m5\datasets\pdt304-us\tabulations\xc.xpt
\\Cdsesub1\EVSPROD\NDA022454\0000\m5\datasets\pdt304-us\tabulations\xd.xpt

Study PPDT301
Basic subject data:
\\Cdsesub1\EVSPROD\NDA022454\0000\m5\datasets\pdt301-us\analysis\adsl.xpt
Standard of truth and image evaluation data:
\\Cdsesub1\EVSPROD\NDA022454\0000\m5\datasets\pdt301-us\analysis\addiag.xpt
Adverse events:
\\Cdsesub1\EVSPROD\NDA022454\0000\m5\datasets\pdt301-us\analysis\adae.xpt

Walker Study
Basic subject, clinical diagnosis, image evaluation, and neuropathology results:

\\Cdsesub1\EVSPROD\NDA022454\0000\m5\datasets\ge-001-walker-us\tabulations\walker-autopsy-decoded.xls

3. STATISTICAL EVALUATION

This section presents a detailed review of the studies DP008-003, PDT3004, PDT301, and a discussion of the Walker study. The sponsor studies, which had specific protocols, are discussed first. The design, endpoints, methodology, and conduct are discussed for each study. After this discussion, the analyses including disposition, demographics, efficacy, and adverse event summaries are discussed together for all three studies. The review used common analyses for all three studies, which will be discribed. The discussion of the Walker study is more exploratory. Unless otherwise noted, all statements and findings refer to the original European clinical study reports, protocols, and amendments. Changes for the US clinical study report are summarized.

3.1 Study PD008-003

The study was conducted at 6 sites in Europe from 25 August 1997 to 24 February 1998. The primary objective of the study was

> To determine the sensitivity and specificity of striatal uptake of ^{123}I-FP- CIT in patients with a clinical diagnosis of Parkinson's disease, multiple system atrophy, progressive or definite essential tremor.

The study included subjects with PD, MSA, PSA and ET, based on previous and de novo clinical diagnosis at enrollment. The study also included healthy volunteers (HV) for "calibration purposes" at the site.

Patients with concomitant medication known or suspected of interacting with striatal uptake through direct competition with ^{123}I-FP- CIT were excluded. A non-exhaustive list of such medication was provided. This list did not contain levodopa. The protocol did state that Parkinson's disease therapy should be withdrawn for assessment in screening.

The study consisted of an initial screening visit, an imaging visit, a 24 – 72 hour post-injection follow-up visit, and a 7-day post-injection follow-up telephone contact. Adverse event information was collected at the imaging visit and the two follow-up visits.

The standard of truth was based on the baseline clinical determination of PS (PD, MSA, PSA) or ET.

The Datscan evaluations included an institutional read and reads by 5 reviewers. The institutional read was not known to be blinded. The 5 reviewer reads were conducted blindly and independently. However, the 5 reviewers were chosen among the study investigators.

The primary efficacy analysis was the sensitivity and specificity of distinguishing PS from ET. This analysis did not included HVs. The primary image read was the institutional read. The standard of truth was the baseline clinical determination. The primary analysis set was the per-protocol set. The primary statistical summary was the lower limit of a one-sided 95% confidence interval. There were no pre-specified statistical tests or thresholds.

The study reviewed mismatches between the clinical diagnoses and Datscan image results. In some cases, the clinical diagnosis was revised. However, according to the Sponsor, all data and analyses were based on the unrevised clinical diagnoses.

There were four protocol amendments, none of which addressed notable design or statistical issues.

The major modifications for the US version of the CSR were

- Change of the primary objective to detecting or excluding SDD
- Including the blinded reads in the primary analysis
- Change of primary analysis method to exact two-sided 95% confidence intervals for sensitivity and specificity.
- Inclusion of HVs in primary efficacy set

3.2 Study PDT3004

The study was conducted at 10 sites in Europe from 18 January 1999 to 28 June 2005. The primary objective of the study was

> To determine the predictive value of Datscan SPECT to differentiate between subjects with early features of Parkinsonism, and other causes of tremor (mainly essential tremor), and healthy volunteers.

The study included subjects with early features of Parkinsonism and healthy volunteers selected from movement disorder and general neurology clinics. The inclusion criteria for Parkinsonism stressed early Parkinsonism without clear diagnosis. Subjects with known cause of tremors were excluded. Subjects with features suggestive of multiple system atrophy or progressive supranuclear palsy were excluded. Subjects with both (1) history and response to drug therapy suggestive of idiopathic Parkinson's disease and (2) clinical history exceeding five years were excluded.

Subjects with concomitant medication known or suspected of interacting with striatal uptake through direct competition with ^{123}I-FP- CIT were excluded. A non-exhaustive list of such medication was provided. The protocols stated that levodopa is not such a medication.

The study was a three-year study with multiple imaging and diagnostic sessions. The primary visits occurred at 0, 18, and 36 months. At each of these time points, there was a sequence of three sessions: a screening and video-recorded assessment session, an imaging session, and a follow-up telephone call. At each screening session, the exclusion criteria including concomitant medicine usage were applied. The imaging session included assessment for adverse events. The follow-up telephone was to occur approximately seven days post-injection and recorded adverse events since the imaging session. The screening sessions at 18 and 36 months included assessment of adverse events since the previous telephone contact. The sessions also included local diagnosis and diagnostic confidence recording before and after Datscan images. There were additional follow-up visits at 24-72 hours after the time 0 image sessions and at month 3 for clinical assessment. The month 3 assessment was for local diagnosis only.

The standards of truth were clinical diagnosis at 18 and 36 months based on the video assessments by two central, independent movement specialists. The specialists were blinded to the Datscan images and the local clinical diagnoses, but had access to clinical information. The specialists classified subjects into: probable PD, possible PD, ET, and other. The assessment at 36 months was based on the video assessments at 0, 18, and 36 months. Differences between the two specialists for the 36 month assessments were resolved by consensus.

The Datscan evaluations included an institutional read and reads by three independent, blinded, central readers. The central readers had no access to clinical diagnosis or information, except age. The readers classified the images into normal, abnormal, or other (non-evaluable).

The primary efficacy analysis was the sensitivity, specificity, positive predictive value, negative predictive value, and accuracy of distinguishing PD from non-PD. PD included probable and possible PD and non-PD includes ET and other. The primary image reads were the institutional and central reads at time 0. The primary standard of truth was the clinical diagnosis by the movement specialists at 36 months. The primary analysis set was patients with the 36 month standard of truth evaluation and baseline evaluable images. The primary statistical summary was the lower limit of a 95% confidence interval. It is known whether the confidence interval was one- or two-sided. There were no pre-specified statistical tests or thresholds.

There were 10 protocol amendments. The above description of the study incorporates the 10 amendments. Amendment 4 (15 August 2000) greatly modified the original study to meet post-authorization commitments to the Committee for Medicinal Products for Human Use. This amendment extended the duration of the study from 3 to 36 months. This amendment added the 18 and 36 month central diagnostic assessment, the 18 and 36 month imaging sessions, the central image reads, and the diagnostic confidence assessments and increased the sample size from 60 to 180 subjects. Amendment 6 (7 December 2001) added the video assessment at time 0. Amendment 9 (29 April 2004) created the diagnostic categories and their classification into PD and non-PD. Amendment 9 (29 April 2004) added a preliminary analysis based on data up to 18 months. Amendment 10 (25 August 2005) added the second central movement specialist and the resolution procedure for differing diagnoses at 36 months. Amendment 10 also added sensitivity and specificity to the primary endpoints.

The major modifications for the US version of the CSR were

- Change of the primary objective to detecting or excluding SDD
- Focus on sensitivity and specificity
- Change of primary analysis method to exact two-sided 95% confidence intervals for sensitivity and specificity
- Combining of European 18-month preliminary clinical study report and 36-month clinical study report

3.3 Study PDT301

The study was conducted at 40 sites in Europe from 21 November 2003 to 28 June 2006. The primary objective of the study was

> To determine the diagnostic efficacy (i.e., sensitivity and specificity) of the visual assessment of Datscan single-photon emission computed tomography (SPECT) images in

differentiating between probable dementia with Lewy bodies (DLB) and non-DLB subjects as determined by the clinical diagnosis of an independent CP used as the standard of truth. The clinical diagnosis was established using internationally accepted diagnostic criteria (including the International Consensus Criteria [ICC]) based on a standardised and comprehensive

The study population consisted of dementia subjects with features of probable or possible DLB and subjects with features of non-DLB (e.g., AD or VaD), from movement disorder, dementia, memory, and general neurology clinics. The chief element of the inclusion criteria was an assessment of dementia with a further diagnosis of probable or possible DLB or AD, or VaD. Subjects with a diagnosis of idiopathic PD or symptoms of multiple system atrophy, cortico-basal degeneration, or Huntington's Chorea disease were excluded. Subjects with concomitant medication known or suspected of interacting with striatal uptake through direct competition with ^{123}I-FP- CIT were excluded. A list of six such medications was given, which does not include levodopa.

The study involved five scheduled visits: screening, baseline, imaging, 48-hour follow-up, and 12-month follow-up. The baseline visit included neuropsychological tests and on-site diagnosis. The image session included the Datscan imaging and the collection of adverse event information. The 48-hour follow-up included a revised on-site diagnosis, using all previous clinical information and the Datscan image and the collection of adverse event information. The 12-month follow-up included neuropsychological tests and on-site diagnosis.

The standards of truth was clinical diagnosis at 0 and 12 months by a panel of three central dementia experts. The experts had access to the clinical data and tests at baseline and at 12 months, and local-diagnosis and management decisions prior to the Datscan image. The panel did not have access to the Datscan image. The panel classified subjects into probable DLB, possible DLB, and non-DLB (AD, VaD), and other.

The Datscan evaluations included an institutional read and reads by three independent, blinded, central readers. The central readers did not have access to clinical information other than age. The central readers classified the images into normal, abnormal, and other categories.

The primary efficacy analysis was the sensitivity and specificity of distinguishing probable DLB from non-DLB. Note that possible DLB was not included. The primary image reads were three central reads. The primary standard of truth was the diagnosis by the central panel. It appears that the primary time point for the standard of truth was baseline. The 12-month standard of truth was discussed as a secondary analysis in the protocol. The primary analysis set was the patients with the baseline standard of truth evaluation and baseline evaluable images. The primary statistical analysis was exact binomial test of the sensitivity and specificity with a one-sided α=0.025. The threshold for sensitivity was 0.65 and the threshold for specificity was 0.73.

There were 4 protocol amendments. There were no major statistical issues in these amendments, except sample size revisions based on new assumptions.

The major modifications for the US version of the CSR were

- Change of the primary objective to detecting or excluding SDD
- Removal of some secondary objectives
- Combining of European original clinical study report and 12-month follow-up clinical study report.

3.4 Sponsor Studies: Disposition, Demographic, and Baseline Characteristics

Table 2 gives a summary of the patient disposition for the three studies. The efficacy set for the purpose of the review was as follows: dosed patients with the protocol-defined primary SOT evaluation and at least one evaluable blinded image read. Note in the case of study DP008-003 this definition differs from protocol definition. In this study, the efficacy set was based on the per-protocol set and the local unblinded image read. The primary efficacy set for the purpose of the review was as follows: patients in the efficacy set that were in the protocol-defined primary efficacy diagnostic distinction. For example, for study PDT301, the protocol-defined primary efficacy diagnostic distinction was probable DLB versus non-DLB; therefore, possible DLB patients were not part of the primary efficacy set.

For study DP008-003, 26 patients withdrew before dosing most from withdraw of consent. All dosed patients were SOT evaluated. 220 of the 224 dosed patients had evaluable images by at least one blinded reader. Of these 220 patients, 35 were healthy volunteers, which were not part of the primary efficacy evaluation.

For study PDT3004, 23 patients withdrew before dosing most from withdraw of consent or investigator decision. 174 of the 179 dosed patients had evaluable images by at least one blinded reader. Only 102 patients had the primary standard of truth, which occurred at 36 months. The majority of dosed patients without the 36-month standard of truth were lost to follow up or withdrew consent. However, since the study involved Datscan imaging at 18 months and 36 months, some withdrawals were because of safety or protocol reasons.

For study PDT301, 23 patients withdrew before dosing most from withdraw of consent. All dosed patients were SOT evaluated at baseline. 313 of the 326 dosed patients have evaluable images by at least one blinded reader. 82 patients were excluded from the primary efficacy analysis, because either they had no diagnosis (26 patients) or a possible DLB diagnosis (56 patients).

Table 2: Sponsor Studies: Patient Disposition.

	DP008-003	PDT3004	PDT301
Enrolled	250	202	351
Dosed	224	179	326
SOT Evaluated[1]	224	102	326
Image Evaluable[2]	220	174	313
Efficacy[3]	220	102	313
Primary Efficacy[4]	185	102	231

Notes:
1. Primary SOT evaluation among dosed patients.
2. Available and unambiguous image for at least one of central blinded readers.
3. Dosed, primary SOT evaluated, and at least one evaluable blinded read.
4. Patients in efficacy set in primary diagnostic distinction (see Table 1).

Table 3 gives the baseline demographics for the Sponsor studies for the dosed set and the efficacy set. For Study DP008-003, among the patients in the dosed set, 39% were female, 46% were aged 65 or over and 98% were Caucasian. The demographics were similar in the efficacy set. For study PDT3004, among the patients in the dosed set, 43% were female, 43% were aged 65 or over and 98% were Caucasian. For this study, because of the large loss-to-follow-up, it is especially important to note that the demographics were similar in the efficacy set. For study PDT301, among the patients in the dosed set, 43% were female, 52% were aged 75 or over and 100% were Caucasian. The demographics were similar in the efficacy set.

Table 3: Sponsor Studies: Baseline Demographics of Dosed and Efficacy Sets.

Study DP008-003	**Dosed N=224 n (%)**	**Efficacy N=220 n (%)**
Female n (%)	87 (39)	84 (38)
Age ≥ 65 n (%)	103 (46)	101 (46)
Age Median (Range)	64 (40 – 80)	64 (40 – 80)
Caucasian n (%)	220 (98)	216 (98)
PDT3004	**Dosed N=179 n (%)**	**Efficacy N=102 n (%)**
Female n (%)	77 (43)	45 (44)
Age ≥ 65 n (%)	77 (43)	42 (41)
Age Median (Range)	63 (33 – 86)	61 (33 – 79)
Caucasian n (%)	176 (98)	102 (100)
PDT301	**Dosed N=326 n (%)**	**Efficacy N=313 n (%)**
Female n (%)	139 (43)	136 (43)
Age ≥ 75 n (%)	168 (52)	163 (52)
Age Median (Range)	75 (54 – 90)	75 (54 – 90)
Caucasian n (%)	326 (100)	313 (100)

Table 4 gives the baseline local diagnosis for the Sponsor studies for the dosed set and the efficacy set. For Study DP008-003, among the patients in the dosed set, the majority (59%) of patients had a baseline local diagnosis of Parkinson's disease. Patients with Parkinson's syndrome diseases (PD, MSA, PSA) made up 71% of this set, patients with essential tremor made up 13%, and healthy volunteers made up 16%. The distribution of the diagnoses was similar in the efficacy set.

For Study PDT3004, among the patients in the dosed set, 44% had probable Parkinson's disease, 31% had possible Parkinson's disease, 4% had benign Parkinson's disease, 12% had possible essential tremor, and 8% had other diseases. Again, for this study, because of the large loss-to-follow-up, it is especially important to note that the baseline local diagnoses were similar in the efficacy set.

For Study PDT301, among the patients in the dosed set, 36% had probable dementia with Lewy bodies, 15% had possible dementia with Lewy bodies, 44% had Alzheimer's disease, 4% had vascular dementia, and 1% had other diseases. The distribution of the diagnoses was similar in the efficacy set.

Table 4: Sponsor Studies: Baseline Local Diagnosis of Dosed and Efficacy Sets.

Study DP008-003	**Dosed N=224 n (%)**	**Efficacy N=220 n (%)**
Healthy Volunteer	35 (16)	35 (16)
Essential Tremor	29 (13)	27 (12)
Parkinson's Disease	132 (59)	130 (59)
Multiple System Atrophy	18 (8)	18 (8)
Progressive Supranuclear Palsy	10 (5)	10 (5)
PDT3004	**Dosed N=179 n (%)**	**Efficacy N=102 n (%)**
Probable Parkinson's Disease	79 (44)	44 (43)
Possible Parkinson's Disease	55 (31)	31 (30)
Benign Parkinson's Disease	8 (4)	6 (6)
Possible Essential Tremor	22 (12)	14 (14)
Other	15 (8)	7 (7)
PDT301	**Dosed N=326 n (%)**	**Efficacy N=313 n (%)**
Probable DLB	118 (36)	111 (35)
Possible DLB	50 (15)	47 (15)
Alzheimer's Disease	142 (44)	140 (45)
Vascular Dementia	13 (4)	12 (4)
Other	3 (1)	3 (1)

3.5 Sponsor Studies: Efficacy Findings

Table 5 gives the standard of truth diagnoses for all dosed and SOT evaluated patients for the three studies. The diagnoses were classified by the sponsor by their association with striatal dopaminergic deficit (SDD). Study DP008-003 had 35 (16%) healthy volunteers. These patients were not part of the original protocol primary efficacy set but were included in the primary efficacy analysis for the US report. For this study, the majority of patients with SDD diagnoses had a Parkinson's disease diagnosis (132/160). For study PDT3004, the majority of patients with SDD diagnoses had a probable Parkinson's disease diagnosis (66/71). A small number of patients 5 (5%) had possible Parkinson's disease. For study PDT301, there were notable numbers of patients for both probable and possible DLB. For this study, only patients with probable DLB were part of the primary efficacy comparison.

Table 5: Sponsor Studies: Primary SOT Diagnosis for all Dosed and SOT Evaluated Patients.

	Diagnosis	DP008-003 N=224 n (%)	PDT3004 N=102 n (%)	PDT301 N=326 n (%)
SDD	Parkinson's Disease	132 (59)		
	Multiple System Atrophy	18 (8)		
	Progressive Supranuclear palsy	10 (4)		
	Probable Parkinson's Disease		66 (65)	
	Possible Parkinson's Disease		5 (5)	
	Probable Dementia with Lewy Bodies			94 (29)
Possible SDD	Possible Dementia with Lewy Bodies			57 (17)
Non-SDD	Healthy Volunteer	35 (16)	3 (3)	
	Essential Tremor	29 (13)	14 (14)	
	Probable Alzheimer's Disease			92 (28)
	Possible Alzheimer's Disease			35 (11)
	Possible Vascular Dementia			9 (3)
	Other		14 (14)	11 (3)
No Diagnosis				28 (9)

Notes: SDD: striatal dopaminergic deficit. Classification of diagnosis by SDD was Sponsor defined.

The review analyzed the Sponsor studies using a common definition of the efficacy set and a common analysis method. The efficacy set, as defined in the disposition Table 2, was all dosed patients with the protocol defined primary standard of truth evaluation and at least one evaluable blinded image read. The primary efficacy set was those patients in the efficacy set who had a diagnosis that was part of the protocol defined primary efficacy diagnostic distinction. All confidence intervals were two-sided 95% intervals calculated using the exact method for a binomial distribution.

Table 6 gives the sensitivity and specificity estimates for the blinded readers using the protocol defined primary efficacy standard of truth and diagnostic distinction.

For study PD008-003, the sensitivity point estimates for five central blinded readers ranged from 92.4% to 96.8% and the lower limits of the two-sided 95% confidence intervals ranged from 87.1% to 92.8%. The specificity point estimates for five readers ranged from 74.1% to 96.3% and the lower limits of the two-sided 95% confidence intervals ranged from 53.7% to 81.0%.

For study PDT3004, the sensitivity point estimates for three central blinded readers ranged from 77.5% to 78.6% and the lower limit of the two-sided 95% confidence intervals ranged from 66.0% to 67.1%. The specificity point estimates for three readers were all 96.8% and the lower limits of the two-sided 95% confidence intervals were all 83.3%.

Both studies DP008-003 and PDT3004 involved movement disorders. However, study DP008-003 involved patients with existing diagnoses, whereas study PDT3004 involved patients with early signs of Parkinson's disease. It is noteworthy that the sensitivities for study DP008-003 were higher than those for study PDT3004. Neither study had pre-specified thresholds or tests for the sensitivity and specificity.

Study PDT301 involved dementia patients. The sensitivity point estimates for three central blinded readers ranged from 75.3% to 80.3% and the lower limits of the two-sided 95% confidence intervals ranged from 64.2% to 69.9%. The specificity point estimates for three readers ranged from 88.5% to 91.2% and the lower limits of the two-sided 95% confidence intervals ranged from 82.0% to 85.2%. This study was the only study with predefined thresholds for sensitivity and specificity, 0.65 and 0.73 respectively. Referring to the lower confidence limits, the study met the sensitivity threshold for 2 of the 3 readers and met the specificity threshold all 3 readers.

For study PDT301, the protocol defined primary efficacy analysis excluded patients with possible DLB. Table 7 gives the results of a secondary analysis with the diagnostic distinction of probable or possible DLB versus non-DLB. For this analysis the sensitivity was notably lower than in the primary analysis.

Table 6: Sponsor Studies: Sensitivity and Specificity for Primary Efficacy.

	Sensitivity (95 % CI)	Specificity (95 % CI)
DP008-003		
Reader A N=185	93.0 (87.9, 96.5)	96.3 (81.0, 99.9)
Reader B N=185	96.8 (92.8, 99.0)	74.1 (53.7, 88.9)
Reader C N=185	96.2 (91.9, 98.6)	85.2 (66.3, 95.8)
Reader D N=185	92.4 (87.1, 96.0)	92.6 (75.7, 99.1)
Reader E N=185	94.3 (89.5, 97.4)	92.6 (75.7, 99.1)
PDT3004		
Reader A N=102	77.5 (66.0, 86.5)	96.8 (83.3, 99.9.)
Reader B N=99	77.9 (66.2, 87.1)	96.8 (83.3, 99.9)
Reader C N=101	78.6 (67.1, 87.5)	96.8 (83.3, 99.9)
PDT301		
Reader A N=216	79.8 (69.2, 88.0)	91.2 (85.2, 95.4)
Reader B N=216	75.3 (64.2, 84.4)	88.5 (82.0, 93.3)
Reader C N=218	80.3 (69.9, 88.3)	90.5 (84.3, 94.9)

Notes: The definition of primary efficacy analysis differs from that used by the Sponsor. See text for details.

Table 7: Study PDT301, Sensitivity and Specificity for Probable or Possible Dementia with Lewy bodies v. Non- Dementia with Lewy Bodies.

	Sensitivity % (95% CI)	**Specificity % (95% CI)**
Reader A N=269	64.4 (55.6, 72.5)	91.2 (85.2, 95.4)
Reader B N=268	60.5 (51.5, 69.0)	88.5 (82.0, 93.3)
Reader C N=273	61.8 (53.1, 70.0)	90.5 (84.3, 94.9)

In addition to the primary SOT at baseline, study PDT301 had a SOT evaluation at 12 months. Table 8 cross-tabulates the two SOTs for the 218 patients with both a baseline and 12 month

standard of truth evaluation. Overall, 218 (83%) of these patients had the same diagnosis at two time points. Among the 46 patients with a diagnosis of possible DLB at baseline, 19 had a diagnosis of probable DLB, 19 had a diagnosis of possible DLB, and 7 had a diagnosis of non-DLB at 12 months.

Table 8: Study PDT301, Baseline SOT Diagnosis v. 12 Month SOT Diagnosis.

		12 Month SOT				
		Probable DLB	**Possible DLB**	**Non-DLB**	**No Diagnosis**	**Total**
Baseline SOT	**Probable DLB**	66	6	0	0	72
	Possible DLB	19	19	7	1	46
	Non-DLB	3	2	118	6	129
	No Diagnosis	1	0	1	15	17
Total		89	27	126	22	264

Table 9 gives the sensitivity and specificity for study PDT301 based on the 12 month SOT. The results based on the 12 month standard of truth were similar those based on the baseline standard of truth. It is possible that 12 months of follow-up is not sufficient for DLB diagnosis.

Table 9: Study PDT301, Sensitivity and Specificity for 12 Month SOT Probable Dementia with Lewy bodies v. Non- Dementia with Lewy Bodies.

	Sensitivity % (95% CI)	**Specificity % (95% CI)**
Reader A N=198	82.3 (72.1, 90.0)	95.0 (89.4, 98.1)
Reader B N=199	74.4 (63.2, 83.6)	91.7 (85.3, 96.0)
Reader C N=199	78.8 (68.2, 87.1)	91.6 (85.1, 95.9)

3.6 Sponsor Studies: Safety Findings

Table 10 summarizes the adverse event experience from the three sponsor studies. Study DP008-003 was a single dose study with the maximum follow-up at 7 days via telephone contact. For this study there were 69 adverse events of which 32 were possibly or probably related to the drug

according to the investigator. The most common system organ class was nervous system disorders. Among the dosed patients, 36 (16%) experienced an adverse event. One subject experienced a serious adverse event, which was "Exacerbation of Parkinson's disease due to change in medicine," which was unrelated to the drug according to the investigator,

Patients in Study PDT3004 received Datscan doses at 0, 18, and 36 months. There were follow-up contacts via telephone 7 days after each dose via telephone. At the 18 and 36 month screening visits, AEs since the last contact were collected. For this study there were 403 adverse events of which 23 were possibly or probably related to the drug according to the investigator. The most common system organ class was nervous system disorders. Among the dosed patients, 125 (70%) experienced an adverse event. 32 subjects experienced a serious adverse event, none of which related to the drug according to the investigator. There were 4 deaths in the study, none of which related to the drug according to the investigator. 3 of the deaths were noted at the 18 month screening and had occurred in the period after the follow-up from the baseline dose to the screening visit for the 18 month dose. One death was recorded at the telephone follow-up for the 36 month dose. The death came from a pneumonia event.

Study PDT301 was a single dose study with the maximum follow-up at 48 – 96 hours via telephone contact. For this study there were 53 adverse events of which 9 were possibly or probably related to the drug according to the investigator. The most common system organ class was general disorders and administrative site conditions. Among the dosed patients, 42 (13%) experienced an adverse event, which were unrelated to the drug according to the investigator. There was one death, which occured 3 days after the dose. The patient fell and died of complications from resulting surgery. The non-fatal serious adverse event was pyrexia and malaise, which occurred 2 days after the dose and was not related to drug according to investigator.

Table 10: Sponsor Studies: Adverse Events.

	DP008-003 N=224	**PDT3004 N=179**	**PDT301 N=326**
AEs	69	403	53
AEs Possibly or Probably Related to Drug	32	23	9
With AEs n (%)	36 (16)	125 (70%)	42 (13)
With SAEs n (%)	1 (0)	32 (18%)	2 (1)
With AEs Leading to Discontinuation n (%)	0 (0)	10 (6)	0 (0)
With Deaths n (%)	0 (0)	4 (2)	1 (0)

3.7 Walker Study

The Walker study consisted of two components: the initial cross-sectional component and a follow-up longitudinal component. Statistical analysis plans have not been obtained for the study

and it is not clear how much of the analyses were prespecified. The initial cross-sectional component compared ioflupane image ratios among clinically diagnosed patient groups. The patient groups included patients with dementia with Lewy bodies, Alzheimer's disease, Parkinson's disease, cortico-basal degeneration, and healthy volunteers. Patients were imaged and clinically evaluated at baseline. The study initiated in June 1996 and the cross-sectional component, including all patient recruitment and baseline diagnosis, was completed by November 1999.

In the follow-up component, some patients were clinically evaluated yearly and underwent post-mortem neuropathological examination. The neuropathological examination was used for disease diagnosis based on standardized criteria for the. The neuropathological examination did not evaluate presence or absence of the DaT protein. The reporting for the longitudinal component focused on the dementia patients, DLB and Alzheimer's disease. At present, there are 22 such patients with neuropathological results. The mean age at the time of the Datscan imaging was 78 years with a range of 58 to 95 years. The mean time from Datscan imaging to death was 39 months with a range of 6 to 106 months. In the analysis of the longitudinal component, patients satisfying clinical criteria for DLB and Alzheimer's disease were assigned a DLB diagnosis.

Table 11 gives the breakdown of the 22 patients for the baseline clinical diagnosis and ioflupane image results by neuropathological results. For the neuropathological defined DLB group, the baseline clinical diagnosis and the ioflupane images each resulted in 7 DLB and 2 non-DLB diagnoses. For the neuropathological defined non-DLB group, the baseline clinical diagnosis incorrectly classified more patients as DLB than the ioflupane images did. The rule of assigning patients satisfying clinical criteria for DLB and Alzheimer's disease to DLB diagnosis may have accounted for some of this misclassification. Table 12 gives the corresponding sensitivities and specificities. The clinical diagnosis misclassification of the neuropathological-determined non-DLB patients was reflected in the lower specificity compared to ioflupane.

Table 11: Walker Study Baseline Clinical Diagnosis and Ioflupane Images by Neuropathological Results.

	Baseline Clinical Diagnosis		**Ioflupane**	
Neuropathological	DLB	Non-DLB	Abnormal	Normal
DLB	7	2	7	2
Non-DLB	7	6	2	11

Table 12: Walker Study Baseline Clinical Diagnosis and Ioflupane Images Sensitivity and Specificity.

	Sensitivity % (95% CI)	**Specificity % (95% CI)**
Baseline Clinical N=22	77.8 (40.0, 97.2)	46.2 (19.2, 74.9)
Ioflupane N=22	77.8 (40.0, 97.2)	84.6 (54.5, 98.1)

4. FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

Table 13 gives the sensitivity and specificity for the three Sponsor studies by demographic subgroups. The analysis sets and methods were the same as that described in Section 3. The protocol-defined primary efficacy comparison was used. For each study, the results from the multiple central readers were reduced to a single majority result for each patient. For studies DP008-003 and PDT3004, the two age subgroups were < 65 and ≥ 65 years. For study PDT3004, the two age subgroups were < 75 and ≥ 75 years. For studies DP008-003 and PDT3004, the negative category for the standard of truth had relatively small samples for the subgroups. Therefore, the specificity estimates for these studies were very variables. Overall, because of the variability in the estimates, it was difficult to determine differences among the subgroups.

Table 13: Sponsor Studies: Sensitivity and Specificity by Demographic Subgroups.

	Sensitivity % (95% CI)	Specificity % (95% CI)
	DP008-003	
Females	91 (81, 97)	100 (59, 100)
Males	97 (92, 99)	90 (68, 99)
Age < 65	95 (88, 99)	100 (74, 100)
Age ≥ 65	95 (87, 98)	87 (60, 98)
	PDT3004	
Females	77 (60, 90)	100 (69, 100)
Males	81 (64, 92)	95 (76, 100)
Age < 65	83 (67, 93)	95 (75, 100)
Age ≥ 65	74 (55, 88)	100 (72, 100)
	PDT3001	
Females	75 (55, 89)	92 (84, 98)
Males	78 (64, 88)	88 (78, 95)
Age < 75	85 (71, 94)	88 (77, 95)
Age ≥ 75	68 (50, 82)	92 (84, 97)

5. SUMMARY AND CONCLUSIONS

5.1 Statistical Issues and Collective Evidence

The sponsor has submitted three Phase III studies (DP008-003, PDT3004, and PDT301) and an independent investigator exploratory study (Walker Study) in support of the NDA for Datscan. The objective of each of these four studies concerned the diagnostic potential of Datscan images for various classes of diseases. Studies DP008-003 and PDT3004 addressed patients with movement disorders. The primary diagnostic distinction for these studies was between Parkinson's syndrome diseases and other movement disorders, principally essential tremor. Studies PDT301 and Walker addressed patients with dementia. The primary diagnostic

distinction for these studies was between dementia with Lewy Bodies and other forms of dementia, such as Alzheimer's disease and vascular dementia. For all four studies, the diagnostic distinction separated diseases associated with degeneration of nigrostriatal dopaminergic neurons and disease without such degeneration. Through the association and lack of association of the degeneration with the various diseases, the Sponsor proposed to use the diagnostic studies to support the DaT protein visualization indication.

Study PDT301 was the only study with prespecified statistical hypotheses. For this study, the primary endpoints were sensitivity and specificity for distinguishing probable dementia with Lewy bodies from non-dementia with Lewy bodies, chiefly Alzheimer's disease and vascular dementia. The prespecified thresholds for sensitivity and specificity were 0.65 and 0.75 respectively. The study statistically demonstrated that the endpoints exceeded these thresholds for 2 out of the 3 central blinded readers and missed the threshold for sensitivity for the third reader by a very small amount. The sensitivity for distinguishing probable or possible dementia with Lewy bodies from non-dementia with Lewy bodies was notably lower. The standard of truth was central blinded clinical diagnosis based on information as baseline. Both the standard of truth determination and the image reads appeared adequately blinded. However, the clinical diagnosis for these dementia diseases at baseline may not have validity. Using a standard of truth at 12 month did not improve the performance measures.

Study DP008-003 had a prespecified statistical analysis plan but no prespecified statistical hypotheses. The original plan was modified for the US submission in several important ways, in some cases to make use of more rigorous data. However, the post-hoc change in plan does diminish the statistical rigor of the analysis. The original primary statistical analysis was based on unblinded reads, whereas the revised analysis included the use of five blinded reads. These blinded readers were, chosen among the site investigators. The original primary analysis excluded healthy volunteers, whereas the revised analysis included healthy volunteers. Overall, the point estimates of sensitivity ranged from 92.4% to 96.8% and for specificity from 74.1% to 96.3%, based on the blinded readers and the exclusion of the healthy volunteers. These sensitivities may have been high because the patient population consisted of patients with existing diagnoses.

Study PDT3004 had a prespecified statistical analysis plan but no prespecified statistical hypotheses. The study concerned movement disorder patients but with early features of Parkinson's syndrome and not existing diagnosis as in study DP008-003. The standard of truth was central blinded clinical diagnosis based on information as clinical information and video assessment up to 36 months. The image review was based on three central blinded readers. Both the standard of truth determination and the image reads appeared adequately blinded. Overall, the point estimates of sensitivity were all approximately 78% and the specificities were all approximately 97%. These sensitivity estimates were notably lower than those for study DP008-003, perhaps because the patient population had less clear disease state.

Likely because all studies were conducted in Europe, the patients in the three Sponsor studies were overwhelmingly Caucasian and may not be representative of the US patient population.

Because of the small sample size, difference among age and gender subgroups could not be determined.

A review of the adverse events from the three Sponsor studies did not reveal any safety issue. There were 35 serious adverse event and 5 deaths, none of these events were related to the drug according to the investigator. Four of the 5 deaths occurred in a long-term follow-up period and the remaining death came from a pneumonia event.

On a population-level the inter-reader variability of the Datscan image reads appeared small. This determination was based on comparing the sensitivity and specificity values across the readers within each study. Additionally, in the three sponsor studies, the number of non-evaluable images was small.

The Walker study was a follow-up study with a small sample of 22 patients. No statistical analysis plan was available for the study. Because of the small sample size, the estimates were very variable. The standard of truth was neuropathology disease diagnosis. For the Datscan images, the estimate for sensitivity was 77.8% (95% CI: 40.0%, 97.2%) and the estimate for specificity was 84.6% (95% CI: 54.5%, 98.1%) for distinguishing between DLB and non-DLB. For the baseline clinical diagnosis, the estimate for sensitivity was same as for the Datscan image, 78.8% (95% CI: 40.0%, 97.2%). The estimate of the specificity of baseline clinical diagnosis was notably lower, 46.2% (95% CI: 19.2%, 74.9%). The lower specificity of the clinical diagnosis may in part be due to patients with multiple clinical diagnoses being classified as DLB. However, the low specificity of the baseline clinical diagnosis calls into question the use of clinical diagnosis of dementias as a standard of truth in study PDT301.

5.1 Conclusions and Recommendations

There are three important issues in evaluating Datscan for its proposed indication based on the Phrase III program: (1) the use of the clinical diagnosis as a standard of truth for visualization of the DaT protein, (2) the statistical rigor of the Phase III program, and (3) the clinical usefulness of the Datscan. Issue (1) is generally outside the scope of the statistical review.

The statistical rigor of the Phase III program was not strong. According to the FDA guidance,— Providing Clinical Evidence of the Effectiveness for Human Drug and Biological Products (1998)— generally at least two adequate and well-controlled studies are needed establish effectiveness. Furthermore, the FDA Guidance—ICH E9 Statistical Principals of Clinical Trials (1998)—"A confirmatory trial is an adequately controlled trial in which the hypotheses are stated in advance and evaluated." All three studies had prespecified analysis plans. However, only one study PDT301 had prespecified statistical hypothesis. Clinical expertise and results from the Walker study showed that the standard of truth for study PDT301 may not have had high validity. In addition to the lack of prespecified hypotheses, the statistical plan for study PD008-003 had notable deficiencies.

Comparisons within the studies and between the studies reveal a balance between the validity of the standard of truth and the clinical utility of Datscan. In study PDT301, the inclusion of the possible-DLB patients in the positive diagnosis category notably lowered the sensitivity. This may be explained by (1) the decrease validity of the clinical diagnosis as a standard of truth or

(2) the lower effectiveness as a diagnostic agent of Datscan for this class of patients. Similarly the sensitivities in study PD008-003 was notably higher than those in study PDT3004, possibly because the patients in study PD008-003 had established diagnoses and the patients in study PDT3004 had early signs of disease. Finally the clinical utility of the product should be judged in with respect to the studied patient population, by comparing the sensitivities and specificities estimates from the study to those of clinical alternatives.

The label should make clear the context of the use of the drug. The proposed indication is not a diagnostic indication. However, the assessment that the drug provides is only useful in the context of making clinical diagnosis. Since the loss of functional nigrostriatal dopaminergic neurons is associated with multiple diseases, it is important to state differentiation between certain diseases, such as dementia with Lewy Bodies and Parkinson's with dementia, is not possible.

Overall, the Phase III program does not support for the approval of Datscan. The Phase III program did not meet the standard of providing two adequate and well controlled trials. Additionally, there are questions concerning the validity of the standards of truths in two studies and the image reading in one study.

The clinical utility of Datscan is questionable. The performance measures of Datscan were notably lower in patients with less certain disease states. It is likely that these patients would have the greatest need for such a product.

SIGNATURES/DISTRIBUTION LIST

Primary Statistical Reviewer: Mark Levenson
Date: 26 August 2009

Concurring Reviewer(s): Jyoti Zalkikar, Aloka Chakravarty

Statistical Team Leader: Jyoti Zalkikar

Biometrics V Division Director: Aloka Chakravarty

cc:
James Moore (Project Manager)
Phillip Davis (Medical Officer)
Louis Marzella (Medical Team Leader)
Dwaine Rieves (Medical Division Chief)
Jyoti Zalkikar (Statistical Team Leader)
Rajeshwari Sridhara (Biometrics V Deputy Division Director)
Aloka Chakravarty (Biometrics V Division Director)
Lillian Patrician (Office of Biostatistics)

Linked Applications	Submission Type/Number	Sponsor Name	Drug Name / Subject
NDA 22454	ORIG 1	GE HEALTHCARE INC	DA TSCAN

This is a representation of an electronic record that was signed electronically and this page is the manifestation of the electronic signature.

/s/

MARK S LEVENSON
08/26/2009

JYOTI ZALKIKAR
08/26/2009

ALOKA G CHAKRAVARTY
08/26/2009