

---

---

# Diagnostic Performance of the Visual Reading of $^{123}\text{I}$ -Ioflupane SPECT Images With or Without Quantification in Patients With Movement Disorders or Dementia

Jan Booij<sup>1</sup>, Jacob Dubroff<sup>2</sup>, Daniel Pryma<sup>2</sup>, Jian Yu<sup>3</sup>, Rajan Agarwal<sup>4</sup>, Paras Lakhani<sup>5</sup>, and Phillip H. Kuo<sup>6</sup>

<sup>1</sup>Department of Nuclear Medicine, Academic Medical Center, University of Amsterdam, Amsterdam, The Netherlands; <sup>2</sup>Department of Radiology, Hospital of the University of Pennsylvania, Philadelphia, Pennsylvania; <sup>3</sup>Diagnostic Imaging, Fox Chase Cancer Center, Philadelphia, Pennsylvania; <sup>4</sup>Radiology Group of Abington, Abington, Pennsylvania; <sup>5</sup>Department of Radiology, Thomas Jefferson University Hospital, Philadelphia, Pennsylvania; and <sup>6</sup>Departments of Medical Imaging, Medicine, and Biomedical Engineering, University of Arizona, Tucson, Arizona

Visual interpretation of  $^{123}\text{I}$ -ioflupane SPECT images has high diagnostic accuracy for differentiating parkinsonian syndromes (PS), from essential tremor and probable dementia with Lewy bodies (DLB) from Alzheimer disease. In this study, we investigated the impact on accuracy and reader confidence offered by the addition of image quantification in comparison with visual interpretation alone. **Methods:** We collected 304  $^{123}\text{I}$ -ioflupane images from 3 trials that included subjects with a clinical diagnosis of PS, non-PS (mainly essential tremor), probable DLB, and non-DLB (mainly Alzheimer disease). Images were reconstructed with standardized parameters before striatal binding ratios were quantified against a normal database. Images were assessed by 5 nuclear medicine physicians who had limited prior experience with  $^{123}\text{I}$ -ioflupane interpretation. In 2 readings at least 1 mo apart, readers performed either a visual interpretation alone or a combined reading (i.e., visual plus quantitative data were available). Readers were asked to rate their confidence of image interpretation and judge scans as easy or difficult to read. Diagnostic accuracy was assessed by comparing image results with the standard of truth (i.e., diagnosis at follow-up) by measuring the positive percentage of agreement (equivalent to sensitivity) and the negative percentage of agreement (equivalent to specificity). The hypothesis that the results of the combined reading were not inferior to the results of the visual reading analysis was tested. **Results:** A comparison of the combined reading and the visual reading revealed a small, insignificant increase in the mean negative percentage of agreement (89.9% vs. 87.9%) and equivalent positive percentages of agreement (80.2% vs. 80.1%). Readers who initially performed a combined analysis had significantly greater accuracy (85.8% vs. 79.2%;  $P = 0.018$ ), and their accuracy was close to that of the expert readers in the original studies (range, 83.3%–87.2%). Mean reader confidence in the interpretation of images showed a significant improvement when combined analysis was used ( $P < 0.0001$ ). **Conclusion:** The addition of quantification allowed readers with limited experience in the interpretation of  $^{123}\text{I}$ -ioflupane SPECT scans to have diagnostic accuracy equivalent to that of the experienced readers in the initial studies. Also, the results of the combined reading were not inferior to the results of the visual reading analysis and offered an increase in reader confidence.

**Key Words:**  $^{123}\text{I}$ -ioflupane SPECT; visual interpretation; quantification; dementia; movement disorders

**J Nucl Med 2017; 58:1821–1826**

DOI: 10.2967/jnumed.116.189266

**S**PECT with  $^{123}\text{I}$ -ioflupane (also known as  $^{123}\text{I}$ -FP-CIT SPECT and marketed by GE Healthcare as DaTSCAN in Europe or DaTscan in the United States) is a well-validated imaging tool for assessing dopamine transporter (DAT) binding in the striatum in vivo (1,2). Degenerative parkinsonian syndromes (PS) such as Parkinson disease (PD) and progressive supranuclear palsy, as well as dementia with Lewy bodies (DLB), are characterized neuropathologically by a pronounced loss of the DAT in the striatum, particularly in the putamen (3–5). Importantly, autopsy studies validated that the antemortem striatal DAT binding in patients with degenerative PS or DLB is positively associated with nigral dopaminergic neuronal density (6,7). In line with these reports, many clinical  $^{123}\text{I}$ -ioflupane SPECT studies showed a loss of striatal DAT binding in DLB and degenerative movement disorders such as PD already at early disease stages (for reviews, see Suwijn et al. (8) and Tatsch and Poepperl (9)).

On the one hand, visual interpretation of  $^{123}\text{I}$ -ioflupane SPECT images has a high diagnostic accuracy for the differentiation of PS from essential tremor and probable DLB from Alzheimer disease (10,11). On the other hand, quantitative analyses are frequently used in routine practice and in scientific studies, and quantification is especially mandatory for detecting small changes in DAT binding (e.g., for measuring the progression of DAT loss in PD) (12). Recent studies suggested that the addition of quantitative analyses of  $^{123}\text{I}$ -ioflupane SPECT images may improve diagnostic accuracy and reproducibility (inter-rater agreement) in routine practice (13–15). Therefore, in this study, we investigated the impact on accuracy, reader confidence, and intra- and interobserver agreement offered by the addition of image quantification in comparison with visual interpretation alone in a large sample of  $^{123}\text{I}$ -ioflupane SPECT images obtained from patients who had movement disorders or dementia and who participated in 3 different phase 3 or 4 studies (10,11,16).

Received Jan. 9, 2017; revision accepted Apr. 19, 2017.

For correspondence or reprints contact: Jan Booij, Department of Nuclear Medicine, Academic Medical Center, University of Amsterdam, Meibergdreef 9, 1105 AZ Amsterdam, The Netherlands.

E-mail: j.booij@amc.uva.nl

Published online May 4, 2017.

COPYRIGHT © 2017 by the Society of Nuclear Medicine and Molecular Imaging.

## MATERIALS AND METHODS

### Subjects and Reconstruction of Images

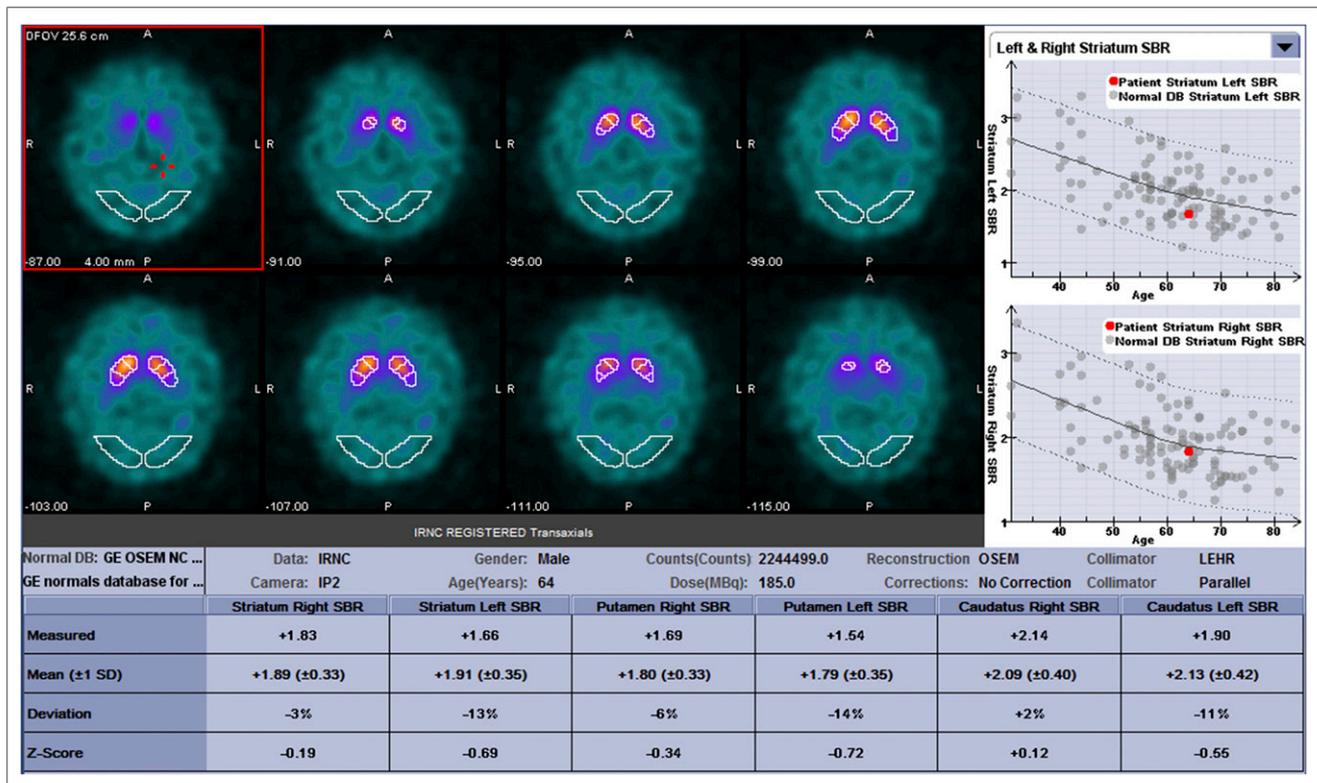
In this study, we collected 304  $^{123}\text{I}$ -ioflupane SPECT images from 3 multicenter phase 3 or 4 trials (10,11,16). All 3 trials were formally approved by the institutional review boards at all participating centers; all subjects in every trial or their legal guardians (according to local regulations) signed an informed consent form. Two of the published studies included subjects with a clinical diagnosis of PS or non-PS (in total, 118 images) (10,16), and the other one included subjects with a clinical diagnosis of probable DLB or non-DLB (mostly Alzheimer disease; 186 images) (11). Images were included only if the subjects had a definitive diagnosis at the 1- to 3-y follow-up. Subjects with a diagnosis of vascular dementia, parkinsonism, or possible DLB were excluded. Also, cases that did not include the whole brain (2 of the original 306 cases; see the Discussion section for more information) were excluded because the images did not register properly with the template in DaTQUANT software (GE Healthcare).

In 2 trials (11,16), a strict acquisition protocol was used to minimize, as much as possible, the variation in raw data acquired per center. In the third trial (10), the imaging was done according to clinical practice at each site. To create uniformity of the reconstructed  $^{123}\text{I}$ -ioflupane SPECT scans, all images were reconstructed with standardized parameters at 1 core laboratory (ordered-subset expectation maximization; 10 subsets, 2 iterations; a Butterworth filter with a cutoff of 0.5 and an order of 10 was applied with the Chang attenuation correction and an attenuation coefficient of  $0.11\text{ cm}^{-1}$ ). For the purpose of a standardized visual analysis, all reconstructed images were displayed in a “cool” color map, as shown in Figure 1.

Regarding quantification, all images were compared against a normal database that was derived from 118 subjects who participated as healthy controls and underwent an  $^{123}\text{I}$ -ioflupane SPECT scan in the Parkinson Progression Markers Initiative (PPMI; <http://www.ppmi-info.org>) study; DaTQUANT software was used. In brief, images were loaded in the DaTQUANT application, which performed a volume-of-interest determination of radiotracer binding in the caudate nucleus and anterior and posterior putamen bilaterally. For the determination of nonspecific binding, 2 additional volumes were determined, in the occipital cortex. The specific striatal binding ratios were calculated by determining the ratio of specific striatal binding to nonspecific binding with the following formula: (mean counts in striatal area – mean counts in occipital cortex)/mean counts in occipital cortex. Only 2 of the 306 image sets from the 3 original studies (10,11,16) were excluded from the masked readings because they could not be quantified by the software.

### Image Reading Sessions

All  $^{123}\text{I}$ -ioflupane SPECT images were assessed by 5 board-certified nuclear medicine physicians in the United States; these physicians had very limited experience with  $^{123}\text{I}$ -ioflupane SPECT image interpretation (i.e., 5–50 prior assessments) and were designated as readers 1–5. Two reading sessions, separated by at least 1 mo, consisted of either a visual interpretation only or a combined reading (both visual and quantitative data were available to the reader). Readers were randomized to the session with visual reading alone or to the session with combined reading first. They were unaware of any clinical data except for sex and age (because aging is associated with a natural loss of striatal DAT binding (17)).



**FIGURE 1.** Example of how DaTQUANT results were displayed to individual readers (along with central image slices with regions of interest overlaid). Readers viewed images on GE Healthcare Xeleris workstation, and quantitative values were available on printed page. DB = distribution; DFOV = dual field of view; LEHR = low-energy high-resolution; NC = normal controls; OSEM = ordered-subset expectation maximization; SBR = striatal binding ratio.

For the combined reading, the readers also received the DaTQUANT data output for each individual scan, including the binding ratios of the caudate nucleus and posterior and anterior putamen (compared with those of age-matched controls), the asymmetry of binding, and the ratio of binding in the putamen to that in the caudate nucleus. The available DaTQUANT output also included images of the reoriented central slices in the striatal region. Before the readings, the masked readers received brief training on the visual differentiation of normal images from various patterns on abnormal images (11) and were familiarized with the quantitative output of the software (Fig. 1).

For each subject's image set, the readers were asked to score the images as normal or abnormal and to rate their confidence about image interpretation on a 5-point scale (from 1 [very challenging] to 5 [very easy]); they were also asked to judge scans as either easy or difficult to read. To help avoid the potential for reader bias in the difficulty assessment, we used a 9-point difficulty rating system (where 0 represented no difficulty and 8 represented maximum difficulty) that incorporated the binary difficulty assessment, confidence about interpretation, and interreader agreement (Table 1).

### Statistics

For this study, the final clinical diagnosis from the original study databases was used as the standard of truth; this diagnosis was established after a minimum follow-up of 12 mo.

The image results were compared with the standard of truth by calculating the positive percentage of agreement (PPA, equivalent to sensitivity), the negative percentage of agreement (NPA, equivalent to specificity), the positive predictive value, the negative predictive value, and accuracy. We assumed that the weighted-average combined-assessment PPA and NPA would be 1.5% and 2.5% higher, respectively, than visual assessment values. Because there were not enough subjects to statistically power superiority assessments, noninferiority comparisons of the combined reading and the visual reading alone were chosen, with a noninferiority margin of 10%, a power of 80%, and a 1-sided  $\alpha$  value of 0.025.

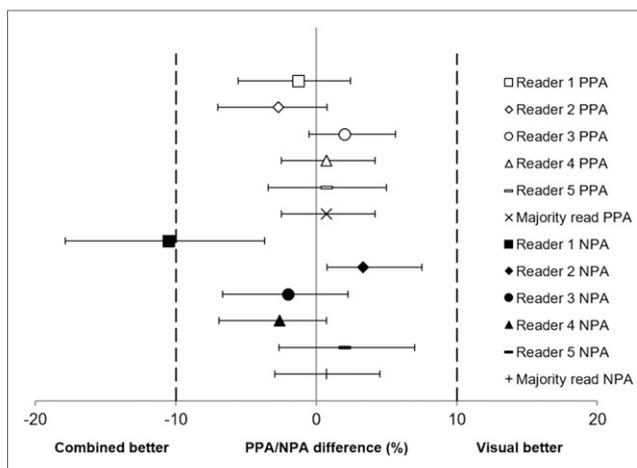
For the assessment of intraobserver agreement, 10% of the images were randomly selected to be read twice by the readers. The intra- and interobserver agreements were calculated and reported as  $\kappa$ -coefficients (18).

### RESULTS

To minimize bias from  $^{123}\text{I}$ -ioflupane assessment experience that would be gained during the course of the reading, readers 1 and 2 started with the visual interpretation alone, and readers 3, 4, and 5 started with the combined assessment.

**TABLE 1**  
Summary of Difficult-to-Read Criteria

Criterion	Definition
i	Confidence of $\leq 2$ by at least 2 readers
ii	Confidence of $\leq 2$ by at least 3 readers
iii	Confidence of $\leq 3$ by at least 2 readers
iv	Confidence of $\leq 3$ by at least 3 readers
v	"Difficult to read" by at least 2 readers
vi	"Difficult to read" by at least 3 readers
vii	Maximum disagreement between readers (2 from 5 readers)
viii	Any disagreement between readers (1 from 5 readers)



**FIGURE 2.** Combined reading was statistically not inferior to visual reading alone. Individual scores of 5 readers for NPA (specificity) and PPA (sensitivity) were all close to line of 0% difference in diagnostic performance between 2 different reads. Error bars represent 95% CI.

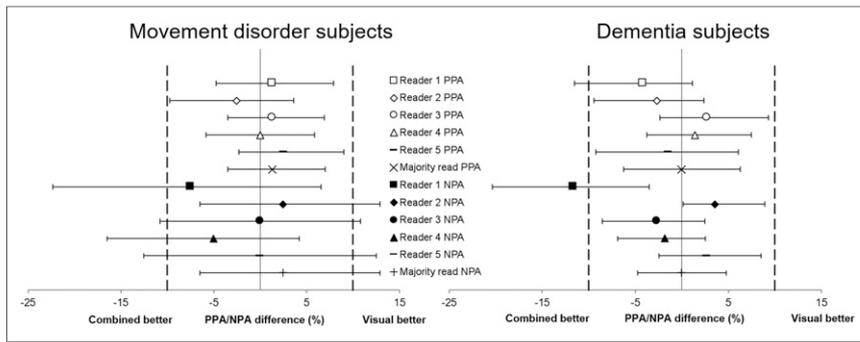
Regarding diagnostic performance, the combined reading was statistically not inferior to the visual reading alone. A small, statistically insignificant increase in the mean NPA (87.9% vs. 89.9%) and equivalent PPAs (80.1% vs. 80.2%) were observed when the results of the visual reading alone and the combined reading were compared. Figure 2 shows that the individual scores for the NPA and the PPA were all close to the line of 0% difference between the 2 readings. This was also true when the subsets of images obtained from patients with movement disorders and patients with dementia were analyzed separately (Fig. 3). The positive predictive values were 90.5% and 89.6% and the negative predictive values were 83.2% and 82.6% for the visual reading alone and the combined reading, respectively (combined sample of patients with movement disorders and patients with dementia).

Readers who initially read in the combined session showed a statistically significantly greater accuracy than readers who initially read in the visual-only session (85.8% vs. 79.2%;  $P = 0.0178$ ) (Fig. 4).

The mean reader confidence score for the interpretation of the images showed a statistically significant improvement when the combined analysis was used instead of the visual-only analysis for the total population (i.e., 4.25 for visual only vs. 4.37 for combined;  $P < 0.0001$ ). These results were observed in the subsets of images obtained from both patients with movement disorders and patients with dementia.

More images from the visual reading alone than from the combined reading met at least 3 of 8 criteria for difficult to read (12.2% vs. 7.9%). In difficult-to-read cases (2 examples are shown in Fig. 5), the PPA remained high in both the visual-only and the combined assessments, whereas the NPA tended to be higher in the combined assessment, with differences becoming greater with increasing levels of difficulty (Fig. 6).

Overall, the intra- and interreader agreements were high. Interreader  $\kappa$ -coefficients for reader pairs were between 0.74 and 0.93 for the visual assessment and between 0.86 and 0.97 for the combined assessment. Intrareader  $\kappa$ -coefficients were between 0.86 and 1.00 for the visual assessment and between 0.92 and 1.00 for the combined assessment. Similar trends were observed in the subsets of images obtained from both patients with movement disorders and patients with dementia. All readers had an intrareader  $\kappa$ -value for the



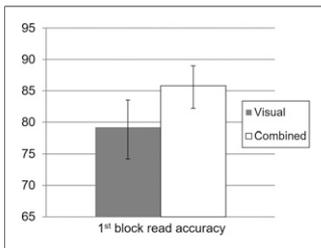
**FIGURE 3.** Combined reading was statistically not inferior to visual reading alone in subsets of both patients with movement disorders (left) and patients with dementia (right). Individual scores of 5 readers for NPA (specificity) and PPA (sensitivity) were all close to line of 0% difference in diagnostic performance between 2 different reads. Error bars represent 95% CI.

combined assessment that was equal to or greater than the  $\kappa$ -value for the visual assessment.

## DISCUSSION

The results obtained when quantification was provided to readers with limited experience in the interpretation of  $^{123}\text{I}$ -ioflupane SPECT scans were statistically not inferior to the results obtained in a session in which the images were analyzed only by visual interpretation. This finding was reflected by minor (statistically insignificant) improvements in diagnostic performance. Finally, the addition of quantification led to an increase in reader confidence about the interpretation of  $^{123}\text{I}$ -ioflupane SPECT scans.

Interestingly, readers who initially read in the combined session showed statistically significantly greater accuracy than readers who initially read in the visual-only session (an increase from 79.2% to 85.8%;  $P = 0.0178$ ). More specifically, the addition of quantification allowed readers with limited experience in the interpretation of  $^{123}\text{I}$ -ioflupane SPECT scans to perform as well as the more experienced readers in the initial clinical studies, as their accuracy score of 85.8% was very similar to the scores of the highly experienced readers (range, 83.3%–87.2%) (10,11,16). Thus, quantification may improve the accuracy of readers with limited experience in the interpretation of these scans. The design of the present study, however, did not allow us to evaluate whether experienced readers may also increase their accuracy using quantitative data. Previous studies suggested that this might be the case, although in those studies, other software was used to analyze the images (13–15).



**FIGURE 4.** Readers who initially read in combined session (white bar; mean  $\pm$  SD) showed statistically significantly greater accuracy (85.8% vs. 79.2%;  $P = 0.0178$ ) than readers who initially read in visual-only session (gray bar).

The sensitivity and specificity of the combined reading were statistically not inferior to those of the visual reading alone in both patients with movement disorders and patients with dementia. This finding is important because it has been suggested that visual analysis may be more difficult in DLB than in PD. In PD, degeneration is commonly much more pronounced in the putamen than in the caudate

nucleus, resulting in images that are easy to interpret visually. However, the posterior-to-anterior gradient may be flatter in DLB than in PD (19) because of relatively early involvement of the caudate nucleus as well, possibly resulting in images with “weak commas” or “balanced loss.” In this regard, a recent Cochrane review concluded that semiquantitative analysis of DAT SPECT scans seemed to be more accurate than visual rating in DLB (20). Nevertheless, our present data do not support the idea that the accuracy of  $^{123}\text{I}$ -ioflupane interpretation in patients with dementia is different from that in patients with movement disorders, but we cannot exclude this possibility in a larger study.

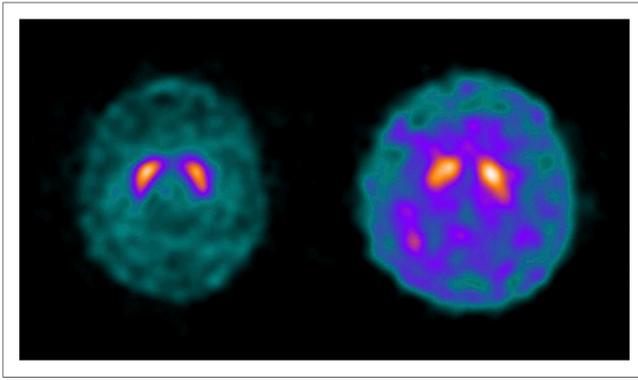
Compared with the diagnostic performance of visual interpretation alone, the diagnostic performance of combined interpretation of  $^{123}\text{I}$ -ioflupane SPECT scans demonstrated statistically insignificant improvement in specificity (the NPA increased from 87.9% to 89.9%) and equivalent sensitivities (PPA, 80.1% vs. 80.2%). Although there may be some slight trends in Figures 2 and 3 toward either combined or visual readings being better, there was no statistically significant difference, and almost all of the error bars overlapped the line of 0% difference between combined and visual readings.

In the present study, the overall sensitivity (PPA) of the scan interpretation was approximately 80%, and the specificity was 90%, with the final clinical diagnosis as the standard of truth. However, in 2 of the 3 phase 3 or 4 studies, the results of the image interpretation were not available to the clinician determining the final clinical diagnosis (11,16). It is well known that the diagnostic accuracy of DAT imaging may be better than the accuracy of clinical diagnosis, particularly in early cases and especially in DLB (21,22). Therefore, the sensitivity and specificity found in the present study actually may be higher.

When combined reading was used, the mean reader confidence in the interpretation of images showed a statistically significant improvement over that obtained with visual reading alone in the total population (i.e., 4.25 for visual only vs. 4.37 for combined;  $P < 0.0001$ ). Although this finding is relevant, it is remarkable that the scores were already very high for visual interpretation alone, as a score of 5 means that the readers were already very confident in their visual interpretation. In addition, although the improvement was statistically significant, it may not necessarily indicate additional benefit from the semiquantitative analysis. We cannot exclude the possibility that confirmation of readers’ assessments by the semiquantitative analysis, even in clear cases, contributed to this finding.

More images met difficult-to-read criteria in the visual reading alone than in the combined reading. Interestingly, the PPA (sensitivity) remained high in both the visual-only and the combined assessments, whereas the NPA (specificity) tended to be higher in the combined assessment, with differences becoming greater with increasing levels of difficulty (Fig. 6). These findings can be explained as follows: a reader who judged a scan as difficult to read was more likely to score it as abnormal. Consequently, the sensitivity remained high, at the cost of specificity.

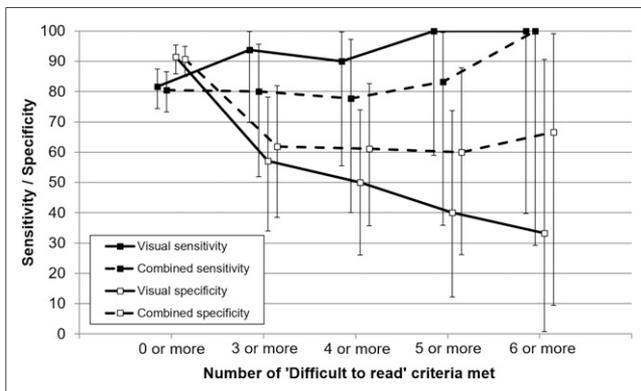
Overall, the intra- and interreader agreements were very high. The combined reading improved the percentage of agreement



**FIGURE 5.** Examples of 2  $^{123}\text{I}$ -ioflupane SPECT studies that were classified as difficult to read. (Left) Some readers classified this scan as difficult to read because dopamine transporter binding was slightly asymmetric and somewhat lower in posterior putamen than in caudate nucleus. Quantification was normal in this patient (i.e., no region of this patient's image had quantified binding ratios or left-right asymmetry more than 2 SDs below those for age-matched controls). (Right) Some readers classified this scan as difficult to read because some binding was still visible in putamen bilaterally (weak commas). Quantification was abnormal in this patient (i.e., this patient's image had multiple regions with quantified binding ratios more than 2 SDs below those for age-matched controls).

among all readers and thus the  $\kappa$ -values. These results are in agreement with those of other  $^{123}\text{I}$ -ioflupane SPECT and amyloid brain PET studies, which also showed that high  $\kappa$ -values for both intraobserver and interobserver agreements (14,23–26) improved as quantitative data were provided to readers (14,26).

Several semiquantitative and automated methods have been developed to analyze  $^{123}\text{I}$ -ioflupane SPECT images (27). In the present study, we used DaTQUANT software to analyze images quantitatively. Of the initial 306 scans that were available from the 3 phase 3 or 4 studies, all except 2 were quantified successfully. For these 2, the software did not correctly register the image to the template for placement of the regions of interest; therefore, these scans were excluded from the readings. The most likely explanation for the registration failure is that these acquisitions had very poor counts (overall, not only in the striatum); however, other, visually similar, count-poor images were registered successfully.



**FIGURE 6.** Effects on mean sensitivity and specificity for visual reading alone and combined reading (all readers) as number of difficult to read criteria met by  $^{123}\text{I}$ -ioflupane SPECT images increased.

In a recent study, the results of automated semiquantitative analyses and independent visual analyses of  $^{123}\text{I}$ -ioflupane SPECT scans were also compared in 120 patients with clinically uncertain parkinsonism. In 12 patients (10% of the total sample), discrepant findings were observed between the 2 analysis methods. More specifically, 9 of the 12 cases were categorized as normal by the automated analyses but as abnormal by the visual analyses, and the opposite was true for the other 3 cases. Also, discrepant cases occurred in relatively old subjects. Interestingly, the authors investigated the clinical characteristics of this subgroup of 12 patients and found that after a minimum of 4.5 y of clinical follow-up, none of the patients developed neurodegenerative parkinsonism (25). In the present study, we also observed discrepant findings, particularly in the difficult-to-read cases. It would be interesting to monitor the clinical and imaging characteristics of the patients to optimize the interpretation of such difficult-to-read cases.

In the present study, we showed that the results of the combined reading were not inferior to the results of the visual reading. Although a large number of cases were included in the present study, superiority assessments could not be performed because there were not enough cases to statistically power such assessments. The sample size required for a superiority assessment of combined reading and visual reading with a superiority margin of 1%, a power of 80%, and a 1-sided  $\alpha$ -value of 0.025 was estimated to be at least 3,000 cases.

A limitation of the present study is that the images were acquired at different centers, potentially inducing heterogeneity in the data input (17). Importantly, however, only images acquired on multihead cameras with low-energy high-resolution collimators were included in the image readings, and all images were reconstructed at 1 core laboratory. Also, in 2 trials (11,16), a strict acquisition protocol was used to minimize, as much as possible, the variation in raw data acquired per center. In the third trial (10), however, the DAT imaging was done according to clinical practice at each site. The inclusion of a wide range of images without limitations may be more similar to clinical practice and demonstrates the value of quantification under more “real-world” conditions. Finally, no images were excluded on the basis of visual quality, and we did not note that any center that provided images that were of lower general visual quality resulted in more images classified as nondiagnostic by the readers.

## CONCLUSION

Combined reading (i.e., visual interpretation and automated software analysis) was statistically not inferior to visual interpretation alone, as reflected by a minor (insignificant) improvement in diagnostic performance. Also, the addition of quantification allowed readers with limited experience in the interpretation of  $^{123}\text{I}$ -ioflupane SPECT scans to perform as well as the more experienced readers in the initial clinical studies. Finally, the addition of semiquantification and comparison with age-matched normal values led to an increase in reader confidence in the interpretation of  $^{123}\text{I}$ -ioflupane SPECT scans and therefore may have resulted in fewer scans being considered difficult to read.

## DISCLOSURE

This study was sponsored by GE Healthcare. Jan Booij is a consultant at GE Healthcare and received research grants from GE Healthcare (paid to the institution). Phillip H. Kuo is a consultant at

GE Healthcare and received research and education grants from GE Healthcare. Phillip H. Kuo is a consultant for MD Training at Home and receives royalties from authorship. No other potential conflict of interest relevant to this article was reported.

## ACKNOWLEDGMENT

This study was presented orally at the SNMMI Congress in San Diego, CA, June 2016, and the EANM Congress in Barcelona, Spain, October 2016.

## REFERENCES

1. Booij J, Andringa G, Rijks LJ, et al. [<sup>123</sup>I]FP-CIT binds to the dopamine transporter as assessed by biodistribution studies in rats and SPECT studies in MPTP-lesioned monkeys. *Synapse*. 1997;27:183–190.
2. Geisler S, Beindorff N, Cremer M, et al. Characterization of [<sup>123</sup>I]FP-CIT binding to the dopamine transporter in the striatum of tree shrews by quantitative in vitro autoradiography. *Synapse*. 2015;69:497–504.
3. Kish SJ, Shannak K, Hornykiewicz O. Uneven pattern of dopamine loss in the striatum of patients with idiopathic Parkinson's disease: pathophysiologic and clinical implications. *N Engl J Med*. 1988;318:876–880.
4. Piggott MA, Marshall EF, Thomas N, et al. Striatal dopaminergic markers in dementia with Lewy bodies, Alzheimer's and Parkinson's diseases: rostrocaudal distribution. *Brain*. 1999;122:1449–1468.
5. Maloteaux JM, Vanisberg MA, Laterre C, Javoy-Agid F, Agid Y, Laduron PM. [<sup>3</sup>H]GBR 12935 binding to dopamine uptake sites: subcellular localization and reduction in Parkinson's disease and progressive supranuclear palsy. *Eur J Pharmacol*. 1988;156:331–340.
6. Colloby SJ, McParland S, O'Brien JT, Attems J. Neuropathological correlates of dopaminergic imaging in Alzheimer's disease and Lewy body dementias. *Brain*. 2012;135:2798–2808.
7. Kraemmer J, Kovacs GG, Perju-Dumbrava L, Pirker S, Traub-Weidinger T, Pirker W. Correlation of striatal dopamine transporter imaging with post mortem substantia nigra cell counts. *Mov Disord*. 2014;29:1767–1773.
8. Suwijn SR, van Boheemen CJ, de Haan RJ, Tissingh G, Booij J, de Bie RM. The diagnostic accuracy of dopamine transporter SPECT imaging to detect nigrostriatal cell loss in patients with Parkinson's disease or clinically uncertain parkinsonism: a systematic review. *EJNMMI Res*. 2015;5:12.
9. Tatsch K, Poeppel G. Nigrostriatal dopamine terminal imaging with dopamine transporter SPECT: an update. *J Nucl Med*. 2013;54:1331–1338.
10. Kupsch AR, Bajaj N, Weiland F, et al. Impact of DaTSCAN SPECT imaging in clinical management, diagnosis, confidence of diagnosis, quality of life, health resource use and safety in patients with clinically uncertain parkinsonian syndromes: a prospective 1-year follow-up of an open-label controlled study. *J Neurol Neurosurg Psychiatry*. 2012;83:620–628.
11. McKeith I, O'Brien J, Walker Z, et al. Sensitivity and specificity of dopamine transporter imaging with [<sup>123</sup>I]-FP-CIT SPECT in dementia with Lewy bodies: a phase III, multicentre study. *Lancet Neurol*. 2007;6:305–313.
12. Marek K, Innis R, van Dyck C, et al. [<sup>123</sup>I]β-CIT SPECT imaging assessment of the rate of Parkinson's disease progression. *Neurology*. 2001;57:2089–2094.
13. Albert NL, Unterrainer M, Diemling M, et al. Implementation of the European multicentre database of healthy controls for [<sup>123</sup>I]FP-CIT SPECT increases diagnostic accuracy in patients with clinically uncertain parkinsonian syndromes. *Eur J Nucl Med Mol Imaging*. 2016;43:1315–1322.
14. Söderlund TA, Dickson JC, Prvulovich E, et al. Value of semiquantitative analysis for clinical reporting of [<sup>123</sup>I]-2-β-carbomethoxy-3β-(4-iodophenyl)-N-(3-fluoropropyl)nortropane SPECT studies. *J Nucl Med*. 2013;54:714–722.
15. Sixel-Döring F, Liepe K, Mollenhauer B, Trautmann E, Trenkwalder C. The role of [<sup>123</sup>I]-FP-CIT-SPECT in the differential diagnosis of Parkinson and tremor syndromes: a critical assessment of 125 cases. *J Neurol*. 2011;258:2147–2154.
16. Marshall VL, Reiningner CB, Marquardt M, et al. Parkinson's disease is over-diagnosed clinically at baseline in diagnostically uncertain cases: a 3-year European multicenter study with repeat [<sup>123</sup>I]FP-CIT SPECT. *Mov Disord*. 2009;24:500–508.
17. Varrone A, Dickson JC, Tossici-Bolt L, et al. European multicentre database of healthy controls for [<sup>123</sup>I]FP-CIT SPECT (ENC-DAT): age-related effects, gender differences and evaluation of different methods of analysis. *Eur J Nucl Med Mol Imaging*. 2013;40:213–227.
18. Vanbelle S, Albert A. A bootstrap method for comparing correlated kappa coefficients. *J Stat Comput Simul*. 2008;78:1009–1015.
19. O'Brien JT, Colloby S, Fenwick J, et al. Dopamine transporter loss visualized with FP-CIT SPECT in the differential diagnosis of dementia with Lewy bodies. *Arch Neurol*. 2004;61:919–925.
20. McCleery J, Morgan S, Bradley KM, Noel-Storr AH, Ansong O, Hyde C. Dopamine transporter imaging for the diagnosis of dementia with Lewy bodies. *Cochrane Database Syst Rev*. 2015;(1):CD010633.
21. Walker Z, Jaros E, Walker RW, et al. Dementia with Lewy bodies: a comparison of clinical diagnosis, FP-CIT single photon emission computed tomography imaging and autopsy. *J Neurol Neurosurg Psychiatry*. 2007;78:1176–1181.
22. Marek K, Seibyl J, Eberly S, et al. Longitudinal follow-up of SWEDD subjects in the PRECEPT Study. *Neurology*. 2014;82:1791–1797.
23. Papathanasiou N, Rondogianni P, Chroni P, et al. Interobserver variability, and visual and quantitative parameters of [<sup>123</sup>I]-FP-CIT SPECT (DaTSCAN) studies. *Ann Nucl Med*. 2012;26:234–240.
24. Seibyl JP, Kupsch A, Booij J, et al. Individual-reader diagnostic performance and between-reader agreement in assessment of subjects with Parkinsonian syndrome or dementia using [<sup>123</sup>I]-ioflupane injection (DaTscan) imaging. *J Nucl Med*. 2014;55:1288–1296.
25. Mäkinen E, Joutsa J, Johansson J, Mäki M, Seppänen M, Kaasinen V. Visual versus automated analysis of [I-123]FP-CIT SPECT scans in parkinsonism. *J Neural Transm (Vienna)*. 2016;123:1309–1318.
26. Nayate AP, Dubroff JG, Schmitt JE, et al. Use of standardized uptake value ratios decreases interreader variability of [<sup>18</sup>F] florbetapir PET brain scan interpretation. *AJNR*. 2015;36:1237–1244.
27. Badiavas K, Molyvda E, Iakovou I, Tsolaki M, Psarrakos K, Karatzas N. SPECT imaging evaluation in movement disorders: far beyond visual assessment. *Eur J Nucl Med Mol Imaging*. 2011;38:764–773.