OPEN

Visual and Semiquantitative Accuracy in Clinical Baseline ^{123}I-Ioflupane SPECT/CT Imaging

Rudolf A. Werner, MD,†‡ Charles Marcus, MD,*§ Sara Sheikhbahaei, MD,* Lilja B. Solnes, MD, MPH,* Jeffrey P. Leal, BA,* Yong Du, PhD,* Steven P. Rowe, MD, PhD,* Takahiro Higuchi, MD, PhD,†‡|| Andreas K. Buck, MD,†‡ Constantin Lapa, MD,† and Mehrbod S. Javadi, MD**

Purpose: We aimed to (a) elucidate the concordance of visual assessment of an initial ^{123}I-ioflupane scan by a human interpreter with comparison to results using a fully automatic semiquantitative method and (b) to assess the accuracy compared to follow-up (f/u) diagnosis established by movement disorder specialists.

Methods: An initial ^{123}I-ioflupane scan was performed in 382 patients with clinically uncertain Parkinsonian syndrome. An experienced reader performed a visual evaluation of all scans independently. The findings of the visual read were compared with semiquantitative evaluation. In addition, available f/u clinical diagnosis (serving as a reference standard) was compared with results of the human read and the software.

Results: When comparing the semiquantitative method with the visual assessment, discordance could be found in 25 (6.5%) of 382 of the cases for the experienced reader ($\kappa = 0.868$). The human observer indicated region of interest misalignment as the main reason for discordance. With neurology f/u serving as reference, the results of the reader revealed a slightly higher accuracy rate (87.7%, $\kappa = 0.75$) compared to semiquantification (86.2%, $\kappa = 0.719$, $P < 0.001$, respectively). No significant difference in the diagnostic performance of the visual read versus software-based assessment was found.

Conclusions: In comparison with a fully automatic semiquantitative method in ^{123}I-ioflupane interpretation, human assessment obtained an almost perfect agreement rate. However, compared to clinical established diagnosis serving as a reference, visual read seemed to be slightly more accurate as a solely software-based quantitative assessment.

Key Words: Parkinson disease, parkinsonism, DaTscan, ^{123}I-ioflupane, SPECT, SPECT/CT

(*Clin Nucl Med* 2019;44: 1–3)

Received for publication June 26, 2018; revision accepted September 8, 2018.

From the *Division of Nuclear Medicine and Molecular Imaging, The Russell H Morgan Department of Radiology and Radiological Science, Johns Hopkins University School of Medicine, Baltimore, MD; †Department of Nuclear Medicine, University of Wuerzburg, Wuerzburg, Germany; ‡Comprehensive Heart Failure Center, University of Wuerzburg, Wuerzburg, Germany; §Department of Radiology, West Virginia University School of Medicine, Morgantown, WV; and ||Department of Biomedical Imaging, National Cardiovascular and Cerebral Research Center, Suita, Japan.

Rudolf A. Werner and Charles Marcus contributed equally to the manuscript.

Conflicts of interest and sources of funding: This work was supported by the Competence Network of Heart Failure funded by the Integrated Research and Treatment Center (IFB) of the Federal Ministry of Education and Research (BMBF) and German Research Council (DFG grant HI 1789/3-3). This project has received funding from the European Union's Horizon 2020 research and innovation program under the Marie Sklodowska-Curie grant agreement No 701983. None declared to all authors.

Correspondence to: Rudolf A. Werner, MD, The Russell H Morgan Department of Radiology and Radiological Science, Johns Hopkins University School of Medicine, 601 N Caroline St, JHOC 3230, 21287 Baltimore, MD, USA. E-mail: rwerner3@jhmi.edu.

ISSN: 0363-9762/19/4401–0001
DOI: 10.1097/RLU.0000000000002333

^{123}I-N-ω-fluoropropyl-2β-carbomethoxy-3β-(4-iodophenyl) nortropane [^{123}I-FP-CIT, ^{123}I-ioflupane (GE Healthcare, Waukesha, WI)] has been utilized in numerous studies and clinical trials to differentiate between patients with nigrostriatal degeneration versus patients without neurodegeneration.[1–3] Most commonly, the result of an ^{123}I-ioflupane scan is based on visual assessment and binary reporting, but several concerns have been raised with regard to solely using visual analysis to interpret those scans, e.g. its intrinsically subjective nature[4–6] or the impact of ageing on dopamine transporter loss.[7–9] Therefore, novel semiquantification approaches including a region of interest (ROI) setting on relevant anatomical structures (i.e., striatum, consisting of caudate nucleus and putamen) are increasingly embedded in clinical routine.[5,10,11] However, these semiquantitative methods also present several challenges, including ROI misalignment in borderline cases, qualitative or artificial asymmetry, or arbitrary setting of predefined ROI.[10,12] In this monocentric setting, we aimed to elucidate the agreement rate between visual interpretation of an initial ^{123}I-ioflupane scan by an experienced reader with a fully automatic semiquantitative method and to assess the reasons that might hamper its diagnostic accuracy. Moreover, we aimed to assess the accuracy of those methods with comparison to criteria-based neurology follow-up (f/u) established by Movement Disorder Specialists.

MATERIALS AND METHODS

In total, 382 patients with clinically uncertain parkinsonian syndrome who underwent a baseline ^{123}I-ioflupane single photon emission computed tomography/computed tomography (SPECT/CT) were included. All patients had been clinically referred from our institutional Movement Disorders Center to assist in diagnosis and were analyzed as part of an institutional review board–approved protocol (Johns Hopkins University, Baltimore, MD, USA).

Imaging Procedure and Fully Automatic Semiquantitative Analysis

Integrated SPECT/CT using a Symbia T2 (Siemens, Erlangen, Germany) equipped with a low-energy, high-resolution collimator was performed in all patients. The obtained data were analyzed using a Xeleris Workstation (DaTQUANT 4.0, GE Healthcare, Waukesha, WI). Regions of interest were automatically defined over the caudate nucleus, putamen, and striatum bilaterally, as well as over the occipital cortex serving as a nonspecific reference region. The mean counts measured by separate ROI over the specific-striatal regions were then divided by the mean counts measured in the occipital region.[13] The derived ratios define the specific uptake in each of those investigated basal ganglion regions. As each center needs to establish its own cutoff for interpretation,[10] the cutoff used within our institution is a value of 1.6 within any of the investigated striatal-specific regions.

Concordance Assessment and Imaging Interpretation

An experienced reader, blinded to the clinical status and neurological diagnosis, performed a visual evaluation of all scans independently (binary reporting). If the findings were discordant with the semiquantitative analysis, the predefined ROIs were displayed and the human observer indicated a reason for the inconsistent findings. These were (1) ROI misalignment, (2) qualitative asymmetry between different specific-striatal regions, or (3) increased background activity by visual assessment of the reference region.

Accuracy Assessment—Comparison With Neurology Follow-up

The clinical diagnosis was assessed based on clinical criteria by Movement Disorder Specialists.[14] To assess the accuracy of the human assessment and the semiquantification, the available criteria-based follow-up (f/u) clinical diagnosis served as a criterion standard.

Statistical Analysis

The degree of agreement between visual, semiquantitative assessment, and neurology f/u as a reference standard were assessed using the Cohen κ coefficient (κ).[15] A κ value of less than 0.2, 0.2–0.4, 0.4–0.6, 0.6–0.8, and 0.8–1.0 indicate slight, fair, moderate, substantial and almost perfect agreement, respectively.[15] McNemar test was performed to compare the performance of visual and semiquantitative assessment. The relation between the semiquantitative uptake in putamen and visual assessment of ^{123}I-ioflupane scan is shown in a scatter plot. Statistical analysis was performed using IBM SPSS Statistics (version 22; Chicago, IL). The statistical significance level was set at $P < 0.05$.

RESULTS

One hundred sixty-eight (44%) of 382 patients were female; the median age of the entire cohort was 66 years (range, 17–93 years). The fully automatic semiquantitative evaluation method revealed the following ratios [median (range)]: striatum right, 1.74 (0.2–4.38); striatum left, 1.79 (0.21–4.32); putamen right, 1.47 (0.1–4.34); putamen left, 1.47 (0.13–4.16); caudate right, 2.17 (0.38–4.46); caudate left, 2.24 (0.28–4.87). The software rated 163 (42.7%) of 382 normal [visual assessment, 178/382 (46.6%)]. When comparing the software-derived findings with the visual read, a discordance could be found in 25 of 382 (6.5%, κ = 0.868, $P < 0.001$, i.e., almost-perfect reproducibility) cases. The reasons for the discordances were as follows: ROI misalignment (16/25, 64%), qualitative asymmetry (5/25, 20%), and increased background activity (4/25, 16%). Figure 1 displays the semiquantitative values for the putamen versus the visual read.

Of the 382 investigated cases, our institutional Movement Disorders Center established a final clinical diagnosis in 138 (36.1%) of 382 after a median f/u of 29 ± 16.1 months, and a neurodegenerative disease was ruled out in 63 (45.7%) of 138. The experienced reader rated 58 (42.0%) of 138 of the scans normal [semiquantification, 52/138 (37.7%)]. With neurology f/u serving as criterion standard, the discordance rate of the experienced reader was 17 of 138 (12.3%; κ = 0.75, $P < 0.001$). In 11 (64.7%) of 17 cases, the diagnosis by the experienced reader was false positive. For the semiquantification, a discordance rate of 19 (13.8%) of 138 was observed (κ = 0.719, $P < 0.001$) with false-positive findings in 15 (78.9%) of 19. There was no significant difference in the sensitivity or specificity of visual assessment versus software-based analysis. Table 1 displays the diagnostic performance of both methods.

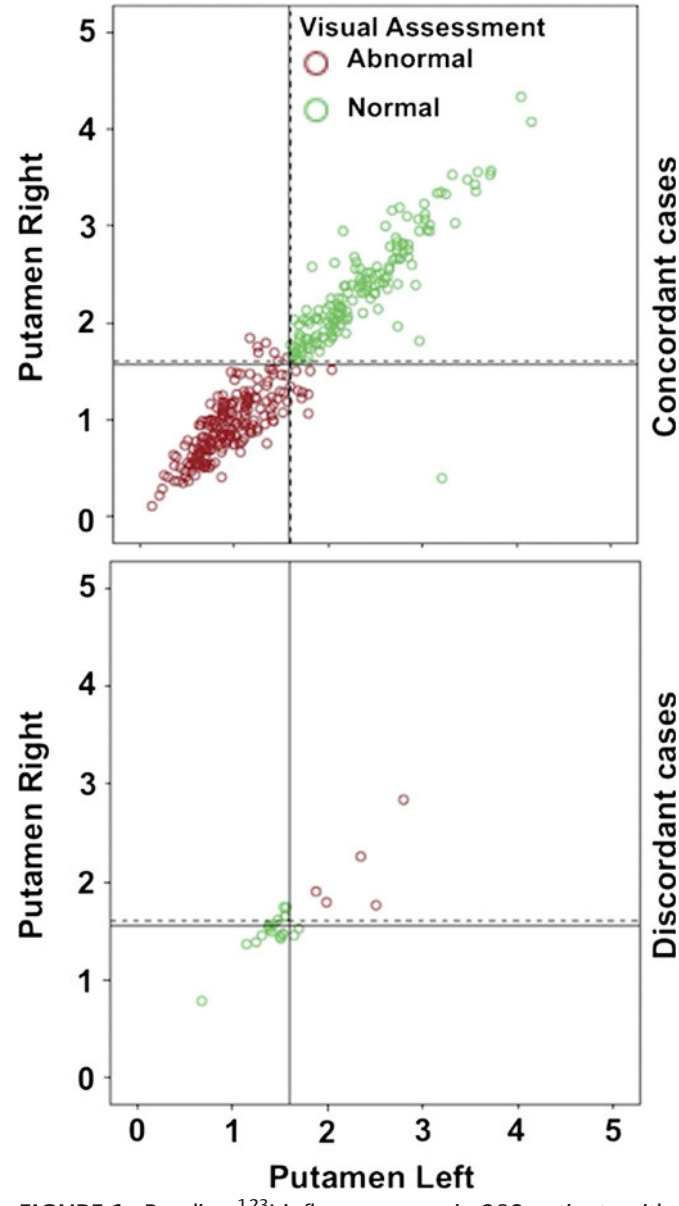

FIGURE 1. Baseline ^{123}I-ioflupane scans in 382 patients with clinically uncertain parkinsonian syndrome with each dot representing an individual subject. Semiquantitative ratios for the right and left putamen as well as the results for an experienced reader are given. Upper panel indicates the concordant (i.e., similar results to software) and lower panel displays the discordant cases (i.e., inconsistent results compared to software) for the human observer (κ = 0.868, $P < 0.001$). Dashed lines indicate the institutional reference cutoff ratio for the software,[10] solid lines indicate the mean.

DISCUSSION

In this largest single-center study published to date, the concordance of ^{123}I-ioflupane scan interpretation assessed by an experienced reader in comparison to a fully automatic semiquantitative evaluation method revealed an almost-perfect reproducibility

TABLE 1. Accuracy of Visual Read and Fully Automatic Semiquantitative Method Compared to Clinically Established Diagnosis Serving as Reference Standard (n = 138)

		Visual Read		Semiquantitative Method	
Sensitivity, %		92		94.7	
Specificity, %		82.5		76.2	
Accuracy, %		87.7		86.2	
True positive	False positive	69	11	71	15
False negative	True negative	6	52	4	48

There was no significant difference in the diagnostic performance of the visual read (performed by an experienced reader) and the semiquantitative method.

(κ = 0.868). In addition, the agreement with available neurology f/u can also be described as substantial for both human read and semiquantification (κ ≤ 0.75).[15]

In previous interagreement studies, κ in the range of 0.87 to 0.97 for experienced readers had been obtained.[16–18] However, patients were enrolled from multiple sites (between 10 and 40 different centers[16,17]) and one might assume that the SPECT cameras utilized, their settings, as well as other procedures might not be standardized among different institutions. In the present study, neither the standard procedures nor the SPECT camera used changed throughout the assessment. Thus, the herein presented high software-reader agreement rates in a single-center setting on a larger scale may further underscore the considerable reliability of ^{123}I-ioflupane imaging. Moreover, the human observer also indicated that ROI misalignment was the main reason for discordance compared to semiquantification. Disease progression in Parkinson disease is first reflected by a decrease in putaminal uptake and followed subsequently by a reduction in the caudate.[19,20] Consequently, the well-known comma shape of the striatal-specific regions on a ^{123}I-ioflupane scan might be altered and a predefined ROI might miss putaminal loss of dopamine transporter binding at an early stage of disease.[12] As displayed in Figure 1, discordant putamen ratio scores tended to be low for the normal visual assessment and as a possible explanation, those cases may be formally beyond the institutional cutoff, but appear to be normal in a visual human read. In addition, putaminal binding ratios tended to be higher for the abnormal visual assessments. One might hypothesize that the value in the caudate can have potential impact on the uptake ratio of the ROI defining the adjacent putamen: Although a dopaminergic deficit in the putamen can be obviously present in a visual assessment, its semiquantitative value may be still considerably high but primarily reflecting the performance of the caudate ("spillover"). Hence, novel approaches considering potential dysmorphism of the comma shape with regard to progressive disease, or improvements in ROI alignment to striatal-specific regions might be intensively sought.[12]

Although not statistically significant, the visual assessment seems to be more accurate when compared to the semiquantitative method, as the human evaluation matched slightly better with available neurology f/u (accuracy of the experienced observer, 87.7% versus software, 86.2%; Table 1). Human observers gather information from the entire scan, which allows for an evaluation of the pattern of uptake instead of an analysis exclusively obtained by a predefined ROI.[21] As demonstrated in the present study, a reader is able to identify even subtle differences of dopamine transporter binding, e.g., qualitative asymmetry between different anatomic regions. Obviously, current software-based methodologies in place cannot overcome these hurdles and human observers still seem to be more sensitive to slight correlative variations in uptake patterns.[21]

The present investigation suffers from several limitations. A larger prospective study would confirm our preliminary findings. Clinical follow-up could not be acquired in all investigated cases; however, this reflects clinical reality, as we relied exclusively on clinical consensus of experts at our institutional Movement Disorders Center.

CONCLUSION

In the present largest single-center assessment, visual read of ^{123}I-ioflupane scan interpretation was compared to a semiquantitative method. With neurology f/u serving as reference standard, the visual observer's interpretation seemed to be more accurate to a solely software-based quantitative assessment.

REFERENCES

1. Tatsch K, Poepperl G. Nigrostriatal dopamine terminal imaging with dopamine transporter SPECT: an update. *J Nucl Med*. 2013;54:1331–1338.
2. Lapa C, Spehl TS, Brumberg J, et al. Influence of CT-based attenuation correction on dopamine transporter SPECT with [(123)I]FP-CIT. *Am J Nucl Med Mol Imaging*. 2015;5:278–286.
3. Meyer PT, Hellwig S. Update on SPECT and PET in parkinsonism—part 1: imaging for differential diagnosis. *Curr Opin Neurol*. 2014;27:390–397.
4. Benamer TS, Patterson J, Grosset DG, et al. Accurate differentiation of parkinsonism and essential tremor using visual assessment of [123I]-FP-CIT SPECT imaging: the [123I]-FP-CIT study group. *Mov Disord*. 2000;15:503–510.
5. Söderlund TA, Dickson JC, Prvulovich E, et al. Value of semiquantitative analysis for clinical reporting of ^{123}I-2-β-carbomethoxy-3β-(4-iodophenyl)-*N*-(3-fluoropropyl)nortropane SPECT studies. *J Nucl Med*. 2013;54:714–722.
6. Booij J, Dubroff J, Pryma D, et al. Diagnostic performance of the visual reading of ^{123}I-ioflupane SPECT images with or without quantification in patients with movement disorders or dementia. *J Nucl Med*. 2017;58:1821–1826.
7. Rinne JO, Hietala J, Ruotsalainen U, et al. Decrease in human striatal dopamine D2 receptor density with age: a PET study with [11C]raclopride. *J Cereb Blood Flow Metab*. 1993;13:310–314.
8. Volkow ND, Ding YS, Fowler JS, et al. Dopamine transporters decrease with age. *J Nucl Med*. 1996;37:554–559.
9. Lavalaye J, Booij J, Reneman L, et al. Effect of age and gender on dopamine transporter imaging with [123I]FP-CIT SPET in healthy volunteers. *Eur J Nucl Med*. 2000;27:867–869.
10. Djang DS, Janssen MJ, Bohnen N, et al. SNM practice guideline for dopamine transporter imaging with ^{123}I-ioflupane SPECT 1.0. *J Nucl Med*. 2012;53:154–163.
11. Tossici-Bolt L, Hoffmann SM, Kemp PM, et al. Quantification of [123I]FP-CIT SPECT brain images: an accurate technique for measurement of the specific binding ratio. *Eur J Nucl Med Mol Imaging*. 2006;33:1491–1499.
12. Augimeri A, Cherubini A, Cascini GL, et al. CADA-computer-aided DaTSCAN analysis. *EJNMMI Phys*. 2016;3:4.
13. Van Laere K, Varrone A, Booij J, et al. EANM procedure guidelines for brain neurotransmission SPECT/PET using dopamine D2 receptor ligands, version 2. *Eur J Nucl Med Mol Imaging*. 2010;37:434–442.
14. Postuma RB, Berg D, Stern M, et al. MDS clinical diagnostic criteria for Parkinson's disease. *Mov Disord*. 2015;30:1591–1601.
15. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33:159–174.
16. Marshall VL, Reininger CB, Marquardt M, et al. Parkinson's disease is overdiagnosed clinically at baseline in diagnostically uncertain cases: a 3-year European multicenter study with repeat [123I]FP-CIT SPECT. *Mov Disord*. 2009;24:500–508.
17. McKeith I, O'Brien J, Walker Z, et al. Sensitivity and specificity of dopamine transporter imaging with 123I-FP-CIT SPECT in dementia with Lewy bodies: a phase III, multicentre study. *Lancet Neurol*. 2007;6:305–313.
18. Papathanasiou N, Rondogianni P, Chroni P, et al. Interobserver variability, and visual and quantitative parameters of (123)I-FP-CIT SPECT (DaTSCAN) studies. *Ann Nucl Med*. 2012;26:234–240.
19. Nurmi E, Ruottinen HM, Bergman J, et al. Rate of progression in Parkinson's disease: a 6-[18F]fluoro-L-dopa PET study. *Mov Disord*. 2001;16:608–615.
20. Winogrodzka A, Bergmans P, Booij J, et al. [123I]FP-CIT SPECT is a useful method to monitor the rate of dopaminergic degeneration in early-stage Parkinson's disease. *J Neural Transm (Vienna)*. 2001;108:1011–1019.
21. Acton PD, Newberg A, Plossl K, et al. Comparison of region-of-interest analysis and human observers in the diagnosis of Parkinson's disease using [99mTc]TRODAT-1 and SPECT. *Phys Med Biol*. 2006;51:575–585.