Movement Disorders
CLINICAL PRACTICE
REVIEW

# Diagnostic Performance of Putaminal Hypointensity on Susceptibility MRI in Distinguishing Parkinson Disease from Progressive Supranuclear Palsy: A Meta-Analysis

*Zhijuan Mao, MD, PhD and Ying Yu, MD, PhD**

**Abstract:** Background: Idiopathic Parkinson's disease (IPD) and progressive supranuclear palsy (PSP) have similar clinical signs and symptoms, making accurate clinical diagnosis difficult. T2* gradient echo (T2* GRE), susceptibility-weighted imaging (SWI), and quantitative susceptibility mapping (QSM) are susceptibility MR imaging sequences that provide more information about brain iron levels than other conventional MR imaging.
Objective: This study aimed to evaluate the diagnostic power of putaminal hypointensity on T2* GRE, SWI, and QSM in distinguishing PSP from IPD.
Methods: Eligible studies were identified via systematic searches of PubMed and Clarivate Analytics® Web of Science® Core Collection. Studies that satisfied the inclusion and exclusion criteria were reviewed. A meta-analysis was conducted using the hierarchical summary receiver operating characteristic curve approach.
Results: Our literature search of the two databases yielded 562 primary articles, 10 of which were deemed relevant and only six were eligible for further analyses. We performed a meta-analysis of putaminal hypointensity measurements: 438 patients with IPD and 109 patients with PSP were enrolled in the quantitative synthesis. The meta-analysis of six studies with 547 patients revealed a sensitivity of 69% (95% confidence interval (CI): 33%–90%) and specificity of 91% (95% CI: 80%–96%) for putaminal hypointensity on T2* GRE, SWI, or QSM distinguishing PSP from IPD.
Conclusions: Putaminal hypointensity on T2* GRE, SWI, or QSM is able to distinguish patients with PSP from those with IPD with high specificity. Further multicenter prospective studies on patients are needed to verify our results.

The principal clinical features of idiopathic Parkinson's disease (IPD) are bradykinesia, resting tremors, rigidity, and postural instability. Several other parkinsonisms, such as progressive supranuclear palsy (PSP), present with similar clinical manifestations. Because of the overlapping symptoms, it is sometimes challenging to obtain a correct diagnosis. Thus, a sensitive and reliable diagnostic marker is urgently required. Thus, the development of a potential MRI biomarker to differentiate between IPD and PSP is necessary.

IPD and PSP are neurodegenerative diseases that share similar pathogenic mechanisms, including mitochondrial dysfunction, oxidative stress,[1] impairment of protein clearance,[2] and neuroinflammation.[3] In addition, dysregulation of metabolism and increased mineral concentrations (especially of iron) are also involved in the occurrence and development of

*Department of Neurology, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China*

***Correspondence to:** Dr. Ying Yu, Department of Neurology Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology Wuhan, China. E-mail: yingyu@hust.edu.cn
**Keywords:** idiopathic Parkinson's disease (IPD), progressive supranuclear palsy (PSP), susceptibility MRI, meta-analysis.

Received 5 June 2022; revised 21 August 2022; accepted 31 August 2022.
Published online 19 October 2022 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/mdc3.13573

neurodegenerative diseases.[4] Accumulated evidence has shown that the degree of metabolism dysregulation and increase in mineral concentrations in different diseases are inconsistent, and that brain iron deposition patterns differ between neurodegenerative diseases, such as IPD and PSP.[5,6] These differences can be used as promising imaging features to support clinical diagnosis.

Postmortem studies have revealed that T2 hypointensity in the putamen is associated with ferritin deposition.[7,8] To distinguish between radiological findings related to pathological mineral deposition and those caused by physiological age-related accumulation in the putamen, selecting an appropriate MRI sequence is crucial. Among the routinely performed MR imaging sequences, T2* gradient echo (T2* GRE) sequences are more sensitive to mineral deposition than T2-weighted sequences. The deposition of paramagnetic substances (such as iron) is related to a decrease in T2* relaxation time, which results in a hypointense signal on T2* GRE imaging.[7,8] Susceptibility-weighted imaging (SWI) is extremely sensitive to mineralization and substances with magnetic susceptibility and can help evaluate the pattern of mineralization in deep gray matter.[9] Quantitative susceptibility mapping (QSM) can process phase MRI data to yield susceptibility maps and eliminate artifacts by deconvolving the field generated by the magnetic susceptibility.[10] Postmortem brain studies demonstrated a strong correlation between volumetric susceptibility and iron concentration in the deep gray matter measured using QSM.[5,11,12] Previous studies of T2* GRE, SWI and QSM in parkinsonism have supported that they may serve as new biomarkers for clinical use. Thus, we conducted a systematic review and meta-analysis to evaluate the diagnostic accuracy of putaminal hypointensity on T2* GRE, SWI, and QSM in distinguishing PSP from IPD.

# Methods

## Study design and search strategy

Our meta-analysis complied with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses statement. Database searches were independently performed by two researchers (YY and ZJM) in PubMed and the Clarivate Analytics® Web of Science® Core Collection. We used the following search keywords: ("swi" OR "magnitude" OR "susceptibility-weighted image" OR "eswan" OR "Enhanced T2 Star Weighted Angiography" OR "t2*" OR "r2*" OR "phase" OR "gradient echo" OR "GRE") AND ("Progressive supranuclear palsy" OR "PSP") AND ("Parkinson's disease" OR "IPD"). Only papers published between January 1, 2000, and March 18, 2022, were considered.

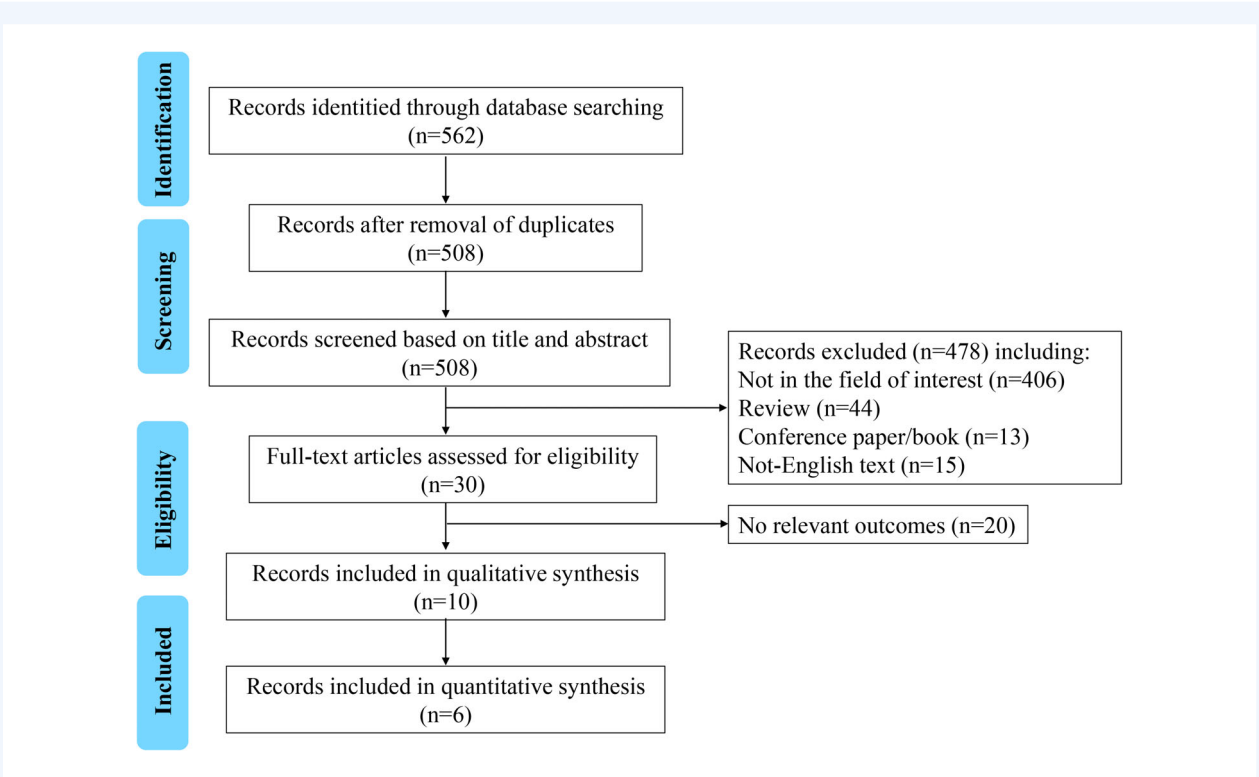


FIG 1. Flowchart of the studies included in the meta-analysis.

## Study selection

Two independent researchers screened the papers considering the following eligibility criteria: (1) written in English. (2) including PSP and IPD patients. (3) T2* GRE, SWI, or QSM of the brain had been performed; and (4) reporting the numbers of true positives, true negatives, false positives, and false negatives or total samples as well as sensitivity and specificity values.

## Data extraction and analysis

Quality Assessment of Diagnostic Accuracy Studies-2 (QUADAS-2) tool was used to evaluate the risk of bias in each study. Two researchers conducted the assessments independently. Any divergence was resolved through discussion and consensus. Data from the qualifying studies were extracted as follows: (1) patient numbers in each group; (2) sensitivity and specificity values; (3) clinical features (age, sex, disease duration, Hoehn and Yahr scale); and (4) MRI technical parameters (sequences, positive radiologic definition, magnetic field strength, slice thickness, echo time [TE], and spatial resolution). Hierarchical summary receiver operating characteristic curve (HSROC) modeling was performed to calculate the general sensitivity and specificity, based on the bivariate normal model for the logit transforms of sensitivity and specificity values reported in each study, and a 95% confidence region and 95% prediction region were provided.[13] Forest plots of sensitivity and specificity were created for the visual assessment. Potential heterogeneity between the datasets was assessed using the inconsistency index ($I^2$). $I^2$ = 25%–50% was regarded as low heterogeneity, $I^2$ = 50%–75% was considered moderate heterogeneity, and $I^2$ > 75% represented high heterogeneity.[14] Statistical analyses were conducted using STATA software (STATA 13.1).

# Results

A systematic search of PubMed and Clarivate Analytics® Web of Science® Core Collection revealed 562 publications. After excluding 54 duplicates, two independent reviewers screened the titles and abstracts of the remaining 508 publications and identified 30 potentially eligible original articles. After full-text screening, 20 publications were excluded because they had no relevant outcomes. A total of 10 publications were included in the qualitative analyses. However, four publications were evaluated using continuous variables and could not be quantified together. Finally, six articles that met the inclusion criteria were included in our meta-analysis. A detailed flowchart of the selection process is shown in Figure 1. The characteristics of the selected studies are presented in Table 1.

As shown in Figure S1, of the studies included, 50% (3 of 6) for patient selection, 83.3% (5 of 6) for index test, 33.3% (2 of 6) for reference standard and 16.6% (1 of 6) for flow and timing showed a low risk of bias. The rest assessment indicated unclear risk of bias.

**TABLE 1** *Study and patient characteristics*

| First author | Examined region | Inspect procedure | ROI measurement | n | | PSP vs. IPD | | Age (y) | | Gender (M/F) | | Disease duration | | Hoehn and Yahr | | MRI | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | PSP | IPD | Sensitivity | Specificity | PSP | IPD | PSP | IPD | PSP | IPD | PSP | IPD | Magnetic field (T) | Slice thickness | Interslice gap | Protocol |
| Arabia 2010 | putamen | visual | NA | 41 | 189 | 60.98% | 97.35% | 69.93 ± 6.01 | 66.50 ± 7.98 | 26/15 | 114/75 | 3.22 ± 2.91 (y) | 5.46 ± 4.29 (y) | NA | NA | 1.5 | 4 mm | NA | T2* GRE |
| Gupta 2010 | posterolateral putamen | visual | NA | 12 | 11 | 50% | 81.82% | 63.3 ± 5.1 | 61.5 ± 5.9 | 9/3 | 7/4 | 3.5 ± 3.9 (y) | 8.1 ± 3.9 (y) | NA | NA | 1.5 | 2 mm | 1.5 mm | SWI |
| Wadia 2013 | putamen | visual | NA | 21 | 60 | 23.80% | 90% | 68 ± 8.4 | 65.8 ± 10.6 | 17/21 | 41/20 | 55.4 ± 50.2 (m) | 84.3 ± 63.8 (m) | 3.7 ± 0.9 | 2.6 ± 0.9 | 1.5 or 3 | 5 mm | 2 mm | T2* GRE |
| Ito 2017 | anterior/ posterior putamen | quantitative | automated | 16 | 26 | 100% | 69.20% | 62–82 (68.5) | 46–80 (64) | 10/4 | 17/9 | 1.0–3.0 (1.5) (y) | 0.5–3.0 (1.5) (y) | 2.5–3 (3) | 1–3 (2.25) | 3 | 2 mm | 0 | QSM |
| Azuma 2019 | putamen | quantitative | manual | 8 | 18 | 37.50% | 100% | 69.5 ± 7.7 | 69.6 ± 6.2 | 5/3 | 8/10 | 40.5 ± 40.3 (m) | 74.6 ± 54.0 (m) | NA | NA | 3 | NA | NA | QSM |
| Sjöström 2019 | putamen | quantitative | manual | 11 | 134 | 90.90% | 81.30% | 72.2 ± 5.5 | 66.9 ± 9.6 | 86/48 | 5/6 | 5.5 ± 2.8 (y) | 6.0 ± 5.0 (y) | 4 (3–5) | 2 (1–2.5) | 3 | NA | NA | SWI |

Abbreviations: ROI, region of interest; n, number; PSP, progressive supranuclear palsy; IPD, idiopathic Parkinson's disease; y, years; m, months; MRI, magnetic resonance imaging; T, tesla; GRE, gradient echo; SWI, susceptibility-weighted imaging; QSM, quantitative susceptibility mapping; NA, not applicable.

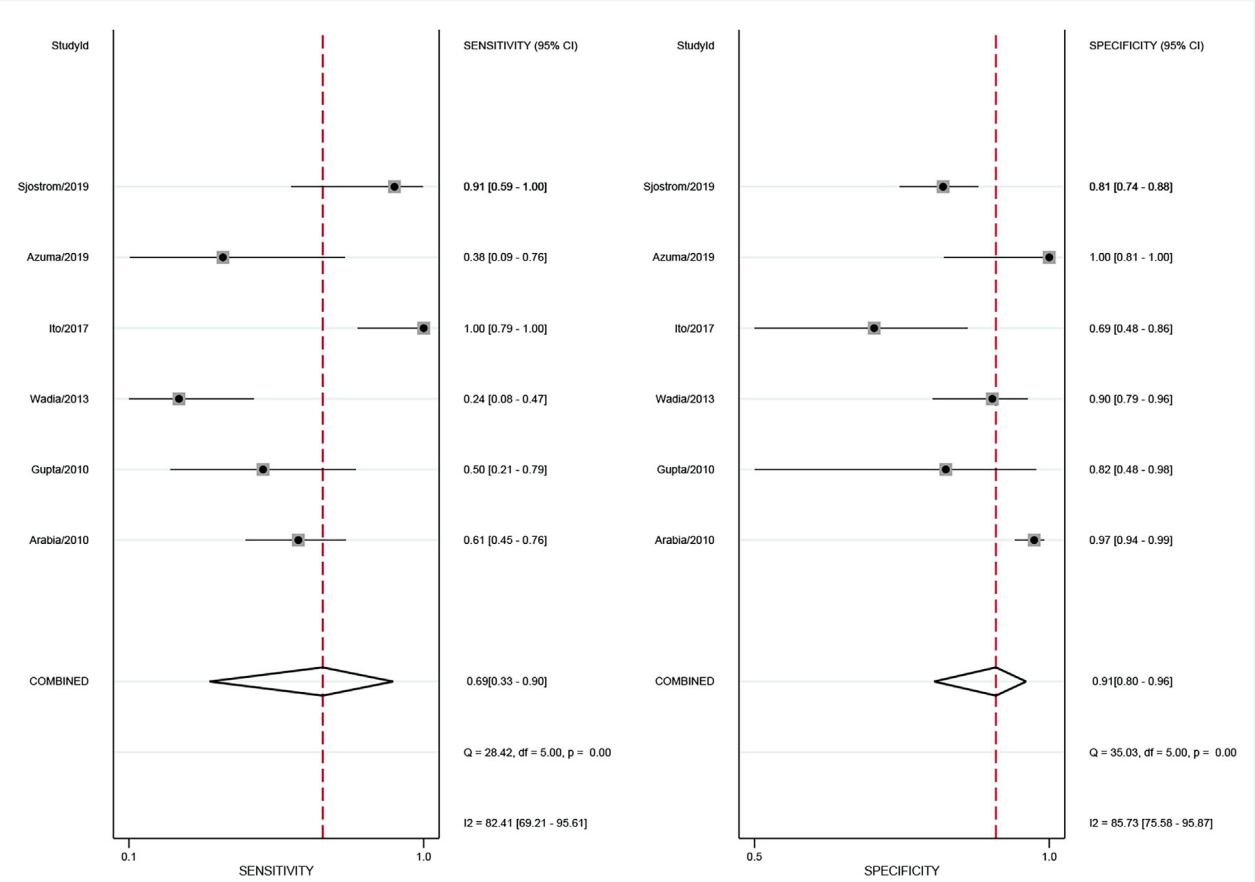


**FIG 2.** Coupled forest plot of pooled sensitivity and specificity. Pooled estimates with 95% CIs are provided. Corresponding heterogeneity statistics are shown at the bottom right corners. CI, confidence intervals.

The six studies included 109 PSP patients and 438 IPD patients. The overall sensitivity and specificity for discriminating PSP from IPD patients were 0.69 (95% confidence interval [CI]: 0.33–0.90) and 0.91 (95% CI: 0.80–0.96), respectively (Fig. 2). Considerable heterogeneity between studies was indicated by $I^2$ scores of 82.41 and 85.73 for sensitivity and specificity, respectively. The area under the HSROC curve was 0.91 (95% CI: 0.88–0.93) (Fig. 3).

In addition to the putamen, other brain regions were investigated. Gupta et al.[15] reported that hypointensity scores of the red nucleus ($P = 0.001$) and substantia nigra ($P = 0.006$) were higher in the PSP group than in the IPD group. Azuma et al.[16] demonstrated that the mean susceptibility values of the substantia nigra and globus pallidus were significantly higher in PSP patients than those in IPD patients ($P < 0.05$). Sjöström et al.[17] also found an apparent increase in the susceptibility of the red nucleus ($P < 0.0001$), globus pallidus ($P < 0.0001$), substantia nigra ($P < 0.0001$), and dentate nucleus ($P < 0.0001$) in PSP patients compared to those in IPD patients.

Two studies used the SWI sequence,[15,17] two used T2* GRE,[18,19] and two used QSM to obtain images.[16,20] For field strength, three studies used 3.0 T,[16,17,20] two used 1.5 T,[15,18] and one used 1.5 T or 3.0 T.[19] The magnetic susceptibility of the bilateral putamina was measured in three studies,[15–17] and the visual presence of putaminal hypointensity was evaluated in three.[18–20] Out of three visually evaluated studies, Arabia et al. demonstrated that the intraclass correlation coefficients (ICCs) for intrarater and interrater reliability in the evaluation of the putaminal intensity were 1.000 and 0.939 in the IPD group, and 0.886 and 1.000 in the PSP group ($P < 0.001$).[18] Gupta et al. reported that ICCs of intrarater and interrater reliability were 0.7 and 0.9 for the putamen.[15] Wadia et al. showed that the kappa value for intrarater and interrater reliability analysis involving GRE hypointensity of the putamen were 1.0 and 0.4576, respectively.[19] The TE was 40 in two studies,[15,18] 20 in one,[17] and 15 in one.[20] Two studies used more than one kind of TE.[16,19] The slice thickness used ranged from 2 to 5 mm, and the interslice gap varied from 0 to 2 mm. Four studies did not mention the years of experience of the image raters[15,17,19,20]; two studies included the data (more than 10 years,[18] 8, and 11 years, respectively[16]).

Although four articles were not included in our quantitative analysis, Sjöström et al.[6] and Mazzucchi et al.[21] found that putamen susceptibility was higher in PSP patients than in PD

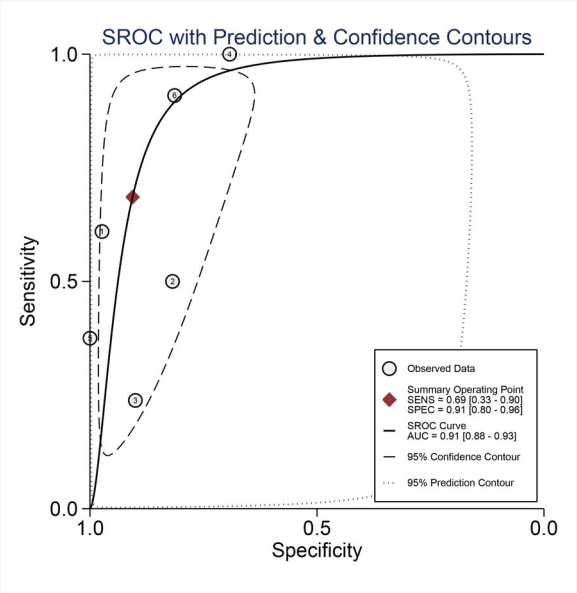


**FIG 3.** Hierarchical summary receiver operating characteristic curve for the diagnostic performance of SWI, T2* GRE, or QSM in differentiating PSP from PD. HSROC, hierarchical summary receiver operating characteristic.

patients. However, Fedeli et al.[22] and Sakurai et al.[23] did not observe any differences.

# Discussion

Our meta-analysis evaluated the diagnostic performance of putaminal hypointensity on T2* GRE, SWI, or QSM to discriminate PSP from IPD. The pooled sensitivity was 69% (95% CI: 3%–90%), the specificity was 91% (95% CI: 80%–96%), and the area under the HSROC curve was 0.91 (95% CI: 0.88–0.93). This indicates that putaminal hypointensity on SWI, T2* GRE, and QSM help distinguish PSP from IPD.

Sometimes, it is difficult to distinguish PSP from IPD using only the clinical diagnostic criteria, and neuroimaging plays a vital role in further differentiation. Iron accumulation leads to considerable local magnetic field inhomogeneity, causing a decline in the local signal (susceptibility gradient). Some studies have demonstrated that low putaminal signal intensity on T2-weighted images in atypical and poorly levodopa-responsive parkinsonism.[24,25] MRI and postmortem neuropathological studies have shown that MRI putaminal signal changes are associated with mineral deposition, cell damage, and gliosis within the putamen,[26–28] implicating putaminal mineral deposition as a neuropathological marker of parkinsonism. The paramagnetic effect of mineral deposition in the cerebrum leads to a hypointense signal on T2* GRE imaging.[18,29] SWI is a comparatively new technique for identifying iron deposition in the brain. By using magnitude and phase images, SWI can reveal both quantified and visible brain mineralization with high sensitivity.[9,30] QSM is a new post-processing technique that can evaluate iron concentrations in the cerebrum by mapping phase-MRI data.[31] Currently, T2* GRE, SWI, and QSM are all neuroimaging techniques used in clinical practice. Therefore, in our study, they were all considered index tests. The overall diagnostic yield of T2* GRE, SWI, or QSM was 69% (95% CI: 33%–90%), and the specificity (91% CI: 80%–96%) was high, indicating that brain MRI including T2* GRE, SWI, or QSM would be a useful tool for distinguishing PSP from IPD.

In addition to the putamen, some studies indicated that the susceptibility of the red nucleus, substantia nigra, and globus pallidus could also be used to distinguish PSP from PD.[15–17] The latest study even showed that diagnostic accuracy comparing PD and PSP ranged from good (Putamen, Subthalamic nucleus) to optimal (red nucleus).[32] However, there have been only a few related studies, which used different imaging techniques, so no unified conclusion can be drawn yet.

Putaminal hypointensity on susceptibility MRI is a nonspecific imaging finding which can also be seen in multiple system atrophy (MSA).[19] Several studies have found that putaminal hypointensity can also distinguish MSA from PD.[6,19,23] These suggest that putaminal hypointensity on susceptibility MRI can help to differentiate patients with PSP or MSA from IPD, but cannot support a PSP or MSA diagnosis in patients presenting with parkinsonism. Whether this could have a role in distinguishing between each parkinsonism requires more and further research.

In addition to susceptibility MRI, other neuroimaging techniques also contribute to distinguishing PSP from IPD, each with its advantages and shortcomings. Though midbrain atrophy and characteristic "hummingbird" shape on midsagittal T1-weighted images are highly specific (99.5%), those signs are less sensitive (51.6%).[33] MRPI and MRPI 2.0 have been used to distinguish between these two diseases in their early stages with high accuracy[34,35]; however, they require expertise in manual measurement and MR image reconstruction, which is complex and limits their widespread use. The automated version of MRPI 2.0, which is currently under development, also requires certain technical skills and is not applicable to routine MRI procedures. A new simple manual MRI measurement of the third ventricle (3 rd V) width showed the 3 rd V/ID ratio demonstrated high diagnostic performance (accuracy >87% and AUC >91%) in distinguishing IPD from PSP patients.[36] As the MRI images were from different 1.5- and 3.0-T scanners, the variability might have increased. A recent study demonstrated great potential for diagnosing suspected PSP using $^{18}$F-PI-2620 PET. When using at least one positive target region, the sensitivity and specificity for detecting PSP-RS are 85% and 77%, respectively.[37] However, $^{18}$F-PI-2620 PET imaging is extremely expensive and difficult to perform in clinical practice. Compared with those above-mentioned neuroimaging techniques, susceptibility MRI is cheaper and easier to perform and showed a good diagnostic performance in discriminating PSP from IPD.

There was substantial heterogeneity in our meta-analysis, and several factors might have caused this variability: (1) magnetic

resonance field strength, slice thickness, and spatial resolution were different among the studies included in this meta-analysis: two studies used 1.5 T,[15,18] three used 3.0 T[16,17,20] and one used both field strengths[19]; thinner slice thickness and smaller inter-slice gaps displayed better detection results; (2) The evaluation methods for positive imaging results were different: some studies used visual evaluation to compare the signal of the thalamus,[18] cerebrospinal fluid[15] and globus pallidus[19]; however, quantitative studies used magnetic susceptibility values measured by in-house software programs[17,20] or ImageJ.[16]; (3) Differences in the placement and size of the region of interest (ROI) might also have affected the results: although most studies focused on the putamen, some divided it into frontal putamen and posterior putamen,[20] while others focused on the outer putamen.[15] Standardized placement of the ROI may help coordinate the results among different study sites; and (4) The drawing methods for the ROI were different: most studies used manually drawn ROI and only one study used automatic segmentation.[19] It remains unclear which method can provide better test–retest reliability.

The difficulty in the clinical diagnosis of IPD and PSP may contribute to heterogeneity. Although the studies used similar clinical diagnostic criteria, none of the patients in these studies were confirmed by pathological examination. Misdiagnosis in some cases cannot be excluded, especially in those in early disease stages. Since the accuracy of clinical diagnosis increases with disease progression, most studies involved patients in the advanced stages, making clinical diagnosis more reliable. When clinical identification becomes more difficult, it is necessary to test the diagnostic performance of T2⋆ GRE, SWI, and QSM for these diseases.

Our study had several limitations. First, only a few studies published thus far have focused on this aspect, and several related studies could not be included in the pooled analysis because of the unavailability of their original data. Consequently, only six studies were included, which might have affected the accuracy of the results. However, the sensitivity and specificity have certain confidence intervals, and excluded one or two studies would not have affected the overall results. Second, one study was retrospective; therefore, the possibility of selection bias could not be excluded. Third, because of the limited number of studies, we did not assess the heterogeneity caused by the three different sequences as indicator tests nor were we able to determine the relationship between severity and T2⋆ GRE, SWI, or QSM abnormalities. Fourth, our pooled analysis revealed significant heterogeneity, which affected the applicability of our results. Nevertheless, in view of the clinical practicability of T2⋆ GRE, SWI, and QSM, it was highly necessary to conduct this meta-analysis, and its results still have a certain practicability to help people make clinical decisions or perform further analyses.

In conclusion, although the sensitivity of T2⋆GRE, SWI, and QSM is still far from ideal, their good specificity prove that they might be a promising diagnostic tool for distinguishing PSP from IPD. Large-scale multicenter imaging studies and long-term clinical follow-up for parkinsonism are needed to verify their diagnostic efficiency.

# Author Roles

(1) Research project: A. Conception, B. Organization, C. Execution; (2) Statistical Analysis: A. Design, B. Execution, C. Review and Critique; (3) Manuscript: A. Writing of the first draft, B. Review and Critique; (4) Literature Search: A. Design, B. Execution.

Z.M.: 1C, 2A, 2B, 2C, 3A, 3B, 4A, 4B.

Y.Y.: 1A, 1B, 1C, 2A, 2B, 2C, 3A, 3B, 4A, 4B.

# Disclosures

**Funding Sources and Conflicts of Interest:** This study was supported by grants from the National Natural Science Foundation of China (No. 82101404 and No. 81901303). The authors declare that there are no conflicts of interest relevant to this work.

**Financial Disclosures for the Previous 12 Months:** The authors declare that there are no additional disclosures to report.

**Ethical Compliance Statement:** The authors confirm that the approval of an institutional review board and patient consent was not required for this work. We confirm that we have read the Journal's position on issues involved in ethical publication and affirm that this work is consistent with those guidelines. ■

# References

1. Jurcau A. Insights into the pathogenesis of neurodegenerative diseases: focus on mitochondrial dysfunction and oxidative stress. *Int J Mol Sci* 2021;22(21). DOI: 10.3390/ijms222111847.
2. Guo F, Liu X, Cai H, Le W. Autophagy in neurodegenerative diseases: pathogenesis and therapy. *Brain Pathol* 2018;28(1):3–13.
3. Kwon HS, Koh SH. Neuroinflammation in neurodegenerative disorders: the roles of microglia and astrocytes. *Transl Neurodegener* 2020;9(1):42.
4. Qian ZM, Ke Y. Brain iron transport. *Biol Rev Camb Philos Soc* 2019; 94(5):1672–1684.
5. Langkammer C, Schweser F, Krebs N, et al. Quantitative susceptibility mapping (QSM) as a means to measure brain iron? A post mortem validation study. *Neuroimage* 2012;62(3):1593–1599.
6. Sjöström H, Granberg T, Westman E, Svenningsson P. Quantitative susceptibility mapping differentiates between parkinsonian disorders. *Parkinsonism Relat Disord* 2017;44:51–57.
7. Kraft E, Trenkwalder C, Auer DP. T2⋆-weighted MRI differentiates multiple system atrophy from Parkinson's disease. *Neurology* 2002;59(8): 1265–1267.
8. von Lewinski F, Werner C, Jörn T, Mohr A, Sixel-Döring F, Trenkwalder C. T2⋆-weighted MRI in diagnosis of multiple system atrophy. A practical approach for clinicians. *J Neurol* 2007;254(9):1184–1188.
9. Haller S, Haacke EM, Thurnher MM, Barkhof F. Susceptibility-weighted imaging: technical essentials and clinical neurologic applications. *Radiology* 2021;299(1):3–26.
10. Liu T, Surapaneni K, Lou M, Cheng L, Spincemaille P, Wang Y. Cerebral microbleeds: burden assessment by using quantitative susceptibility mapping. *Radiology* 2012;262(1):269–278.
11. Sun H, Walsh AJ, Lebel RM, et al. Validation of quantitative susceptibility mapping with Perls' iron staining for subcortical gray matter. *Neuroimage* 2015;105:486–492.
12. Zheng W, Nichol H, Liu S, Cheng YC, Haacke EM. Measuring iron in the brain using quantitative susceptibility mapping and X-ray fluorescence imaging. *Neuroimage* 2013;78:68–74.

13. Rutter CM, Gatsonis CA. A hierarchical regression approach to meta-analysis of diagnostic test accuracy evaluations. *Stat Med* 2001;20(19):2865–2884.

14. Cumpston M, Li T, Page MJ, et al. Updated guidance for trusted systematic reviews: a new edition of the Cochrane handbook for systematic reviews of interventions. *Cochrane Database Syst Rev* 2019;10:Ed000142.

15. Gupta D, Saini J, Kesavadas C, Sarma PS, Kishore A. Utility of susceptibility-weighted MRI in differentiating Parkinson's disease and atypical parkinsonism. *Neuroradiology* 2010;52(12):1087–1094.

16. Azuma M, Hirai T, Nakaura T, et al. Combining quantitative susceptibility mapping to the morphometric index in differentiating between progressive supranuclear palsy and Parkinson's disease. *J Neurol Sci* 2019;406:116443.

17. Sjöström H, Surova Y, Nilsson M, et al. Mapping of apparent susceptibility yields promising diagnostic separation of progressive supranuclear palsy from other causes of parkinsonism. *Sci Rep* 2019;9(1):6079.

18. Arabia G, Morelli M, Paglionico S, et al. An magnetic resonance imaging T2*-weighted sequence at short echo time to detect putaminal hypointensity in Parkinsonisms. *Mov Disord* 2010;25(16):2728–2734.

19. Wadia PM, Howard P, Ribeirro MQ, Robblee J, Asante A, Mikulis DJ, Lang AE. The value of GRE, ADC and routine MRI in distinguishing parkinsonian disorders. *Can J Neurol Sci* 2013;40(3):389–402.

20. Ito K, Ohtsuka C, Yoshioka K, et al. Differential diagnosis of parkinsonism by a combined use of diffusion kurtosis imaging and quantitative susceptibility mapping. *Neuroradiology* 2017;59(8):759–769.

21. Mazzucchi S, Frosini D, Costagli M, et al. Quantitative susceptibility mapping in atypical Parkinsonisms. *Neuroimage Clin* 2019;24:101999.

22. Fedeli MP, Contarino VE, Siggillino S, et al. Iron deposition in Parkinsonisms: a quantitative susceptibility mapping study in the deep grey matter. *Eur J Radiol* 2020;133:109394.

23. Sakurai K, Imabayashi E, Tokumaru AM, et al. Volume of interest analysis of spatially normalized PRESTO imaging to differentiate between Parkinson disease and atypical parkinsonian syndrome. *Magn Reson Med Sci* 2017;16(1):16–22.

24. Boelmans K, Holst B, Hackius M, Finsterbusch J, Gerloff C, Fiehler J, Münchau A. Brain iron deposition fingerprints in Parkinson's disease and progressive supranuclear palsy. *Mov Disord* 2012;27(3):421–427.

25. Foroutan P, Murray ME, Fujioka S, Schweitzer KJ, Dickson DW, Wszolek ZK, Grant SC. Progressive supranuclear palsy: high-field-strength MR microscopy in the human substantia nigra and globus pallidus. *Radiology* 2013;266(1):280–288.

26. Lee JH, Lee MS. Brain iron accumulation in atypical parkinsonian syndromes: in vivo MRI evidences for distinctive patterns. *Front Neurol* 2019;10:74.

27. Macia F, Yekhlef F, Ballan G, Delmer O, Tison F. T2-hyperintense lateral rim and hypointense putamen are typical but not exclusive of multiple system atrophy. *Arch Neurol* 2001;58(6):1024–1026.

28. Schwarz J, Weis S, Kraft E, et al. Signal changes on MRI and increases in reactive microgliosis, astrogliosis, and iron in the putamen of two patients with multiple system atrophy. *J Neurol Neurosurg Psychiatry* 1996;60(1):98–101.

29. De Reuck JL, Deramecourt V, Auger F, et al. Iron deposits in postmortem brains of patients with neurodegenerative and cerebrovascular diseases: a semi-quantitative 7.0 T magnetic resonance imaging study. *Eur J Neurol* 2014;21(7):1026–1031.

30. Murakami Y, Kakeda S, Watanabe K, et al. Usefulness of quantitative susceptibility mapping for the diagnosis of Parkinson disease. *AJNR Am J Neuroradiol* 2015;36(6):1102–1108.

31. Miyata M, Kakeda S, Toyoshima Y, et al. Potential usefulness of signal intensity of cerebral gyri on quantitative susceptibility mapping for discriminating corticobasal degeneration from progressive supranuclear palsy and Parkinson's disease. *Neuroradiology* 2019;61(11):1251–1259.

32. Mazzucchi S, Del Prete E, Costagli M, et al. Morphometric imaging and quantitative susceptibility mapping as complementary tools in the diagnosis of parkinsonisms. *Eur J Neurol* 2022;29:2944–2955.

33. Mueller C, Hussl A, Krismer F, et al. The diagnostic accuracy of the hummingbird and morning glory sign in patients with neurodegenerative parkinsonism. *Parkinsonism Relat Disord* 2018;54:90–94.

34. Nigro S, Antonini A, Vaillancourt DE, et al. Automated MRI classification in progressive supranuclear palsy: a large international cohort study. *Mov Disord* 2020;35(6):976–983.

35. Quattrone A, Morelli M, Nigro S, et al. A new MR imaging index for differentiation of progressive Supranuclear palsy-parkinsonism from Parkinson's disease. *Parkinsonism Relat Disord* 2018;54:3–8.

36. Quattrone A, Antonini A, Vaillancourt DE, et al. A new MRI measure to early differentiate progressive Supranuclear palsy from De novo Parkinson's disease in clinical practice: an international study. *Mov Disord* 2021;36(3):681–689.

37. Brendel M, Barthel H, van Eimeren T, et al. Assessment of 18F-PI-2620 as a biomarker in progressive Supranuclear palsy. *JAMA Neurol* 2020;77(11):1408–1419.

# Supporting Information

Supporting information may be found in the online version of this article.

**Figure S1.** Quality assessment of diagnostic accuracy studies-2.