



The diagnostic value of diffusion kurtosis imaging in Parkinson's disease: a systematic review and meta-analysis

Qilin Li^{1#}, Jinfeng Cao^{1#}, Xinjiang Liu², Xin Luo¹, Ge Su³, Dejian Wang⁴, Bo Lin^{3,5}

¹Department of Radiology, Zibo Central Hospital, Zibo, China; ²Department of Radiology, Shanghai Pudong Hospital (Pudong Hospital Affiliated to Fudan University), Shanghai, China; ³College of Computer Science and Technology, Zhejiang University, Hangzhou, China; ⁴Department of R&D, Hangzhou Healink Technology, Hangzhou, China; ⁵Innovation Centre for Information, Binjiang Institute of Zhejiang University, Hangzhou, China

Contributions: (I) Conception and design: Q Li; (II) Administrative support: X Luo, X Liu; (III) Provision of study materials or patients: J Cao; (IV) Collection and assembly of data: J Cao, G Su; (V) Data analysis and interpretation: D Wang, B Lin; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work.

Correspondence to: Xin Luo. Department of Radiology, Zibo Central Hospital, 54, Gongqingtuan West Road, Zibo 255020, China.
Email: asd0601@126.com.

Background: Under the background that diffusion kurtosis imaging (DKI) has become a research hotspot of central nervous system diseases, there are no studies with large sample size evaluating the value of DKI in diagnosing Parkinson's disease (PD). Moreover, the diagnostic efficacy of DKI in PD is not consistent. Therefore, the main purpose of this study is to use the method of meta-analysis, to summarize and evaluate the diagnostic efficacy of DKI in the identification of PD, and to explore the value of its clinical application.

Methods: We use PICOS principles for project design. The included patients were PD patients, and the control group were healthy volunteers. We hope to use DKI to make a differential diagnosis between the two, and this study is a diagnostic test. We performed a literature search of English (PubMed, Embase, Cochrane Library, etc.) and Chinese (China knowledge Network, Wanfang Data Knowledge Service platform, China Science and Technology Journal Database, China Biomedical Literature Service system) databases for related literatures on the efficacy of DKI in the differential diagnosis of PD published before March 29, 2022. We used Revman 5.3 software to assess the quality of the literature, Meta-Disc 1.4 software for summarizing sensitivity (Sen), specificity (Spe), diagnostic odds ratios, and heterogeneity tests, and for subgrouping, and Stata 16.0 software for publication bias analysis.

Results: Fourteen articles were included through the literature search. The 14 studies included 535 patients with PD and 486 patients without PD. Most of the included literature had good clinical applicability and relatively low risk. By merging statistics, the results obtained were as follows: Sen =0.78 [95% confidence interval (CI): 0.74–0.81], Spe =0.83 (95% CI: 0.79–0.86), and the area under the summary receiver operating characteristic (SROC) curve was 0.8870.

Discussion: The results of the meta-analysis showed that magnetic resonance DKI has comparable diagnostic accuracy in the diagnosis of PD. However, this study also has limitations, and the use of different diagnostic gold standards in the included studies may have some impact on the case selection in the study.

Keywords: Magnetic resonance imaging (MRI); diffusion kurtosis imaging (DKI); Parkinson's disease (PD); meta-analysis

Submitted Mar 01, 2022. Accepted for publication Apr 20, 2022.

doi: 10.21037/atm-22-1461

View this article at: <https://dx.doi.org/10.21037/atm-22-1461>

Introduction

Parkinson's disease (PD) also known as tremor paralysis, is a progressive neurodegenerative disease (1) that mainly occurs in elderly patients, and its incidence is second only to Alzheimer's disease (2). According to the standards of the British Brain Bank Parkinson's Disease Research Center, the diagnostic accuracy rate of PD is 82.7%. Furthermore, based on neuropathological results, the clinical diagnosis accuracy rate of PD is only 26% for patients who are untreated or have no obvious drug response. Therefore, at present, relying on the existing clinical symptoms, early diagnosis of PD is very difficult, and there is an urgent need to identify more sensitive biomarkers (3-5).

In recent years, imaging technology has played a greater role in the diagnosis of PD. A study has found that, on transcranial ultrasound assessment, hyperechoic substantia nigra can be detected in more than 90% of PD patients in the experimental population at an early stage (6). However, its sensitivity (Sen) for diagnosing PD by substantia nigra echo is only 0.40, and its specificity (Spe) is 0.61, so it is considered that the diagnostic accuracy of transcranial ultrasonography for PD is low, which is not enough to meet the routine clinical application (7). A study has used positron emission tomography (PET) technology to use the characteristics of ^{18}F -FDG in the brain to reflect the changes in energy metabolism in different brain regions, so as to distinguish PD patients from other atypical PD syndromes patient (8). However, due to the characteristics of high ionizing radiation and high inspection cost, this technology has not been used as a routine imaging examination for PD patients, so it cannot be widely used in clinical practice. Magnetic resonance imaging (MRI) technology has the advantages of being non-invasive, multi-directional, multi-parameter, and highly reproducible, and has been widely used in clinical applications and subject research of the nervous system. Diffusion-weighted imaging (DWI) is another widely used clinical imaging technique, which is mostly used in the diagnosis of acute ischemic stroke, traumatic brain injury, and inflammatory lesions. However, there are many complex biological tissue structural barriers in the central nervous system. Due to the existence of these barriers, the diffusion and directional movement of water molecules are restricted.

Diffusion tensor imaging (DTI) can explain anisotropic diffusion; it can evaluate the degree of diffusion of water molecules in three-dimensional space, and describe the directionality of water molecules in the fiber tissue (9).

However, DTI assumes that water molecules move in a Gaussian distribution pattern in an ideal environment, but this is overly ideal. Owing to the existence of biological tissue structural barriers, the diffusion method of water molecules is altered, and exhibits non-Gaussian diffusion (10). In contrast, DKI can more accurately quantify the degree of non-Gaussian diffusion of water molecules, reflecting the true water molecular motion and complex microstructure of the organization. DKI is an emerging sequence developed in recent years; it is a pulse sequence that fits the imaging advantages of each sequence and extends its development based on the characteristics of DWI and DTI. The biggest advantage of DKI over DTI is that it not only provides diffusion tensor measurement parameters [fractional anisotropy (FA), mean diffusivity (MD), etc.], but also provides diffusion kurtosis measurement parameters: mean kurtosis (MK), axial kurtosis (AK), and radial kurtosis (RK) (11), which are more sensitive to monitoring the pathological changes of gray matter nuclei and white matter cross fibers (12,13). At present, DKI has been widely used in research regarding both central nervous system diseases (such as acute ischemic stroke, glioma, meningioma, and brain injury, etc.), and neurodegenerative diseases (such as PD and Alzheimer's disease). Applied research is also becoming increasingly extensive and in-depth. Therefore, DKI may be used as a biomarker to monitor the pathological changes of PD.

To our knowledge, under the background that DKI has become a research hotspot of central nervous system diseases, there are no studies with large sample size evaluating the value of DKI in diagnosing PD. Moreover, the diagnostic efficacy of DKI in PD is not consistent, and its clinical application is also controversial. Most studies (5,14-23) believe that DKI has the highest accuracy in the diagnosis of PD; Sun *et al.* (24), Yao *et al.* (25), and Zhang *et al.* (26) believe that DKI has a certain value in the diagnosis of PD, but the accuracy is not high. Among them, Sun *et al.* (24) included 32 PD patients and 20 healthy controls. The results showed that the PD group had an AUC =0.69, a Sen of 62.5%, and a Spe of 80.0%. It shows that the MK value of substantia nigra can be used for the diagnosis of PD, but the diagnostic efficiency is low. Zhang *et al.* (26) included 45 PD patients and 39 healthy controls. Their results showed that the substantia nigra MK value increased in the PD group, with an AUC =0.654, a Sen of 42.22%, and a Spe of 92.31%. The substantia nigra MK value has low Sen and low diagnostic efficiency for PD diagnosis, and its clinical application needs to be further

studied. Since the sample size of most studies is small, it is impossible to draw relatively accurate conclusions about the performance of DKI. Therefore, the main purpose of this study is to use the method of meta-analysis, by increasing the sample size and integrating the data of multiple studies, to summarize and evaluate the diagnostic efficacy of DKI in the identification of PD, and to explore the value of its clinical application. We present the following article in accordance with the PRISMA-DTA reporting checklist (available at <https://atm.amegroups.com/article/view/10.21037/atm-22-1461/rc>).

Methods

Literature sources and retrieval strategies

In this study, we performed a literature search of English (PubMed, Embase, Cochrane Library, etc.) and Chinese (China knowledge Network, Wanfang Data Knowledge Service platform, China Science and Technology Journal Database, China Biomedical Literature Service system) databases for related literatures on the efficacy of DKI in the differential diagnosis of PD published before March 29, 2022. A relatively comprehensive literature search was carried out using subject terms, free words, and compiled search formulas to reduce the rate of document misdetection and missed detection. When designing the search formula, we collected all documents related to the subject in the searched databases as comprehensively as possible. In order to expand the scope of the literature search, we also screened the references of high-quality literature, screened the articles related to this research, and obtained as many studies related to all topics as possible.

Literature inclusion criteria

- (I) Studies evaluating the diagnostic efficacy of DKI in PD;
- (II) Patients with PD diagnosed by clinical diagnostic criteria;
- (III) Retrospective or prospective studies;
- (IV) This research is diagnostic research, and it is necessary to directly or indirectly extract the four tables of data used in the diagnostic test, namely true positive (TP), false positive (FP), false negative (FN), true negative (TN), etc.;
- (V) Studies with sample sizes >30 cases;
- (VI) Studies involving only human research subjects.

Literature exclusion criteria

- (I) Studies with no diagnostic criteria for the study cases;
- (II) Summaries, case reports, editorials, and non-retrospective or prospective studies;
- (III) Studies with incomplete data and those where it was impossible to extract the documents of the four-table data;
- (IV) For repeated studies by the same author, we excluded the low-quality literature and those with small sample sizes;
- (V) Basic experiments, such as animal experiments and genetic research experiments, etc.;
- (VI) Relevant documents without informed consent or ethics committee approval.

Data extraction

First, we performed a preliminary search on multiple databases using subject terms, free words, and related search formulas. Next, we screened the references of the retrieved articles in order to avoid omissions, and used NoteExpress software (Beijing Aegean Software Company, Beijing China) to screen out duplicate documents. In order to reduce selection bias, two researchers independently extracted the data of each article. Disagreements were resolved through negotiation or objective analysis a third senior researcher. The data extracted from the included articles were as follows: literature information (first author, year, region, language), patient information (sample size, gender, age, diagnostic criteria), inspection information (machine model, inspection location, b-value), study design (prospective or retrospective), and diagnosis data in four tables (TP, FP, FN, TN), etc.

Risk of bias

Quality assessment of the included literature was also independently carried out by two researchers under the premise of mastering the quality assessment rules. We used the Quality Assessment of Diagnostic Accuracy Studies-2 (QUADAS-2) quality evaluation tool in the Revman 5.3 software (The Cochrane Collaboration, UK) to evaluate the quality of the original literature (27). The QUADAS-2 tool mainly includes four evaluation items: (I) patient selection; (II) index test; (III) reference standard; (IV) flow and timing (27). Each article answered “yes”, “no” or “uncertain” according to the relevant item, and the corresponding risk

of bias level could be judged as “low”, “high”, “uncertain”. If the quality of the literature was inconsistent, the two researchers resolved the inconsistency through negotiation or objective analysis by a third senior researcher.

Statistical analysis

Heterogeneity test

Heterogeneity testing is a key aspect of meta-analysis research, as it can help in identifying the relevant reasons affecting the accuracy assessment as well as evaluating whether the combined statistical model is appropriate. In the diagnostic test research, the threshold effect is one of the important factors causing heterogeneity. Threshold refers to the standard or cut-off point used to define positive (or negative) test results in different studies. Different thresholds in each study lead to Sen and Spe or positive (+LR) and negative (–LR) likelihood ratios, the difference between which is known as the threshold effect. When there is a threshold effect, Sen and Spe are negatively correlated (or Sen and 1–Spe are positively correlated), forming a typical “shoulder-arm” graph in the summary receiver operating characteristic (SROC) curve plan (28). The threshold effect can be evaluated by three Meta-Disc [the Unit of Clinical Biostatistics team of the Ramón y Cajal Hospital in Madrid (Spain)] methods (28): (I) observe the relationship between the paired accuracy estimates in the forest map. If there is a threshold effect, the forest plot will show that as Sen increases, Spe will decrease, and vice versa. There is also a negative correlation between the +LR and –LR; (II) the image formed by the estimated accuracy of each study in the SROC plan shows a typical “shoulder-arm-like” distribution, indicating the existence of a threshold effect; and (III) through the Meta-Disc quantitative test threshold effect: use software to calculate the Spearman correlation coefficient between the Sen logarithm and 1–Spe logarithm; a significant positive correlation indicates that the study has a threshold effect.

In the evaluation of diagnostic test systems, in addition to the heterogeneous changes caused by the threshold effect, some factors may also lead to changes in the accuracy assessment between different studies, which are called non-threshold effects. These include the condition of the study subjects, the parameters of the experimental sequence, the experience level of the experimental operators, the selection of reference standards, and the implementation process. For heterogeneity caused by non-threshold effects, two Meta-Disc software methods can be used to evaluate (28):

(I) visual inspection of accuracy assessment through forest maps; and (II) statistical methods for testing, including the Q test (Cochran-Q), Chi-square test (Chi-square), and other methods. Among them, the Q test (Cochran-Q) indicates statistically whether there is heterogeneity in the research. When the test result is $P > 0.10$, it is considered that there is no heterogeneity between the studies; however, when the test result $P \leq 0.10$, it is believed that the results of multiple studies are heterogeneous. I^2 is an index that is used to measure the degree of heterogeneity; the larger the value, the more obvious the heterogeneity of the research. Typically, 50% is the limit; when $I^2 > 50\%$, it is considered that there is a high heterogeneity between the research results (29).

Combined statistics

If there was heterogeneity caused by the threshold effect between the included studies, the SROC curve would be constructed or the receiver operating characteristic (ROC) curve could be used, and the area under the ROC curve (AUC) could be combined. If there was no threshold effect between studies and $I^2 < 50\%$, statistics could be directly combined. The Sen, Spe, LR, diagnostic odds ratio (DOR), and AUC were then calculated (30). It is generally considered that an AUC value above 80% is meaningful, that is, the closer the value is to 100%, the higher the diagnostic accuracy.

Publication bias

Publication bias refers to the publication process of research papers, as journals or review institutions are subjectively more likely to accept positive research results, making statistically significant studies easier to publish. Publication bias is easy to make the results of meta-analysis exaggerate the diagnostic performance of the test to be evaluated, resulting in biased evaluation of diagnostic accuracy. At present, the most commonly used method to identify publication bias in meta-analysis is funnel plot, which is highly subjective. Due to the large degree of dispersion of studies with small sample size, the distribution in the funnel diagram is more scattered, and most of them are arranged symmetrically at the bottom; while the studies with large sample size are more concentrated in the middle of the funnel diagram, and are mainly distributed at the top of the funnel diagram, showing as an inverted funnel shape, the application is relatively simple and practical, but because it cannot be quantitatively analyzed, it is greatly affected by subjectivity. In this article, we use Stata 16.0's (StataCorp

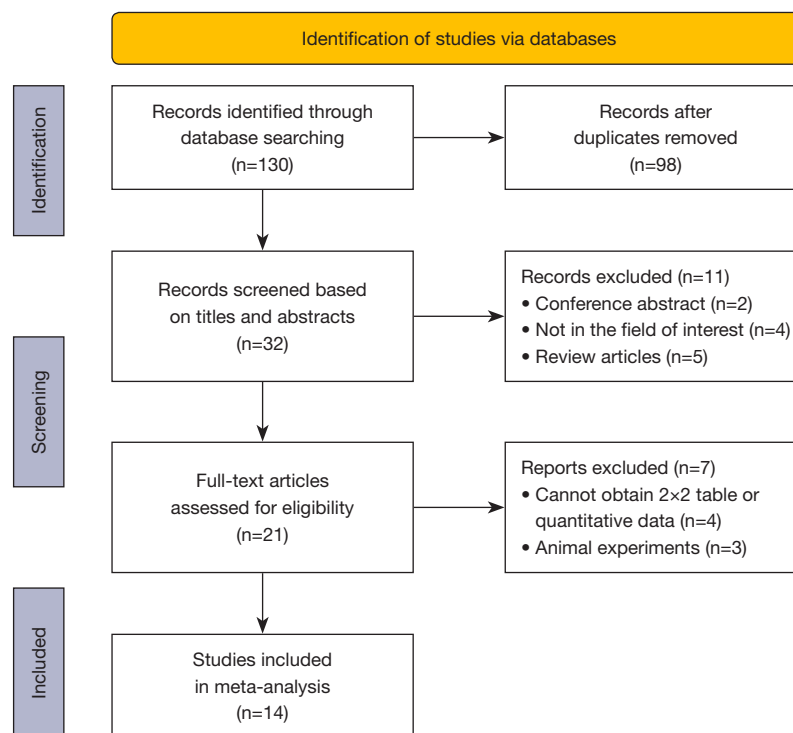


Figure 1 Literature retrieval flow chart.

LLC., USA) metabias command to construct a Deeks funnel chart to detect publication bias.

Results

Literature search results

Through a comprehensive search of the Chinese and English databases, 130 documents were initially retrieved, including 77 Chinese articles and 53 English documents. By using NoteExpress software to remove duplicate documents as well as reading the titles and abstracts, 98 articles (such as irrelevant documents) were excluded. We then read the full texts of the remaining 32 articles and strictly screened them according to the established standards. Finally, this meta-analysis included a total of 14 relevant studies to evaluate the diagnostic efficacy of DKI in the differential diagnosis of PD, including eight Chinese articles and six English articles. The literature search flow chart is shown in *Figure 1*.

Data extraction of included literature

After screening, 14 articles were included for meta-analysis,

with a total of 535 patients and 486 healthy volunteers. The general information of the included studies is shown in *Table 1*, and the four-grid table data of each parameter value of the included literature is shown in *Table 2*.

Included research quality evaluation

We evaluated the included literature using the QUADAS-2 diagnostic accuracy study quality evaluation tool in Revman 5.3 software. After the quality assessment of the included literature according to the QUADAS-2 standard, the quality assessment chart was produced. The quality assessment chart is shown in *Figures 2,3*.

In summary, most of the included literature had good clinical applicability and relatively low risk.

Heterogeneity test

First, Meta-Disc was used to assess the threshold effect of the relevant data of the included studies. Using the SROC plan, we found that the precise estimator of each study did not have a “shoulder-arm-like” distribution; thus, the Sen logarithm and 1-Spe Spearman correlation coefficient of

Table 1 Basic information of the included studies

Study, year	Region	Type of study	Equipment type	B value	Guideline
Liu Y, 2020	Mainland China	P	Philips 3.0 T	0, 1,000, 2,000	I
Li Y, 2019	Mainland China	P	SIEMENS 3.0 T	0, 1,000, 2,000	II
Bingbing G, 2020	Mainland China	P	GE HDXT 3.0 T	0, 1,000, 2,000	I
Sun QY, 2019	Mainland China	P	Philips 3.0 T	0, 500, 1,000, 1,500, 2,000	II
Sejnoha Minsterova A, 2020	Czech Republic	P	Siemens 3.0 T	500, 1,000, 2,000, 4000	II
Cai CC, 2019	Mainland China	P	GE Discovery 3.0 T	0, 1,000, 2,000	I
Sun YQ, 2017	Mainland China	P	Philips Achieva 3.0 T	0, 1,000, 2,000	I
Zhang ZW, 2020	Mainland China	P	GE Discovery 3.0 T	0, 1,000, 2,000	I
Zhang G, 2015	Mainland China	P	GE EXCITE 3.0 T	0, 1,000, 2,000	III
Kamagata K, 2013	Japan	P	Philips 3.0 T	0, 1,000, 2,000	III
Wang JJ, 2011	Taiwan	P	Siemens 3.0 T	0, 1,000, 4,000	IV
Kamagata K, 2017	Japan	P	Philips 3.0 T	1,000, 2,000	III
Yao JQ, 2020	Mainland China	P	Philips TX 3.0 T	0, 1,000, 2,000	I
Gao H, 2020	Mainland China	P	GE Discovery 3.0 T	0, 1,250, 2,500	III

P: prospective study. (I) The diagnostic criteria for Parkinson's disease formulated by the Chinese Medical Association in 2016; (II) the diagnostic criteria for Parkinson's disease formulated by the International Society for Movement Disorders in 2015; (III) the Parkinson's disease clinical diagnostic criteria for the Brain Bank of the Parkinson's Disease Association of the United Kingdom; (IV) National Research on Neurological Diseases and Stroke in the United States (NINDS) Parkinson's disease diagnostic criteria.

Table 2 Data of four grids of parameters' values included in the study

Index	Study	Sensitivity	Specificity	AUC	Sample	TP	FP	FN	TN
MK	Liu Y, 2020	0.750	0.883	0.845	124	48	7	16	53
FA	Li Y, 2019	0.762	0.775	0.799	41	16	4	5	16
MD	Sun QY, 2019	0.700	0.850	0.823	40	14	3	6	17
MK	Cai CC, 2019	0.667	0.837	0.782	117	41	9	20	47
MK	Sun YQ, 2017	0.625	0.800	0.689	52	20	4	12	16
FA	Zhang ZW, 2020	0.800	0.640	0.754	84	36	14	9	25
MK	Zhang G, 2015	0.944	0.917	0.976	144	68	6	4	66
MK	Kamagata K, 2013	0.867	0.941	0.912	32	15	1	2	14
MK	Wang JJ, 2011	0.920	0.870	0.950	60	28	4	2	26
MK	Kamagata K, 2017	0.710	0.820	0.770	58	21	5	9	23
MK	Yao JQ, 2020	0.630	0.750	0.700	55	22	5	13	15
FA	Sejnoha Minsterova A	0.920	0.710	0.810	71	21	14	2	34
MD	Bingbing G	0.675	0.870	0.811	58	23	3	12	20
MK	Gao H, 2020	0.860	0.830	0.914	85	43	6	7	29

AUC, area under the receiver operating characteristic curve; FA, fractional anisotropy; FN, false negative; FP, false positive; MD, mean diffusivity; MK, mean kurtosis; TN, true negative; TP, true positive.

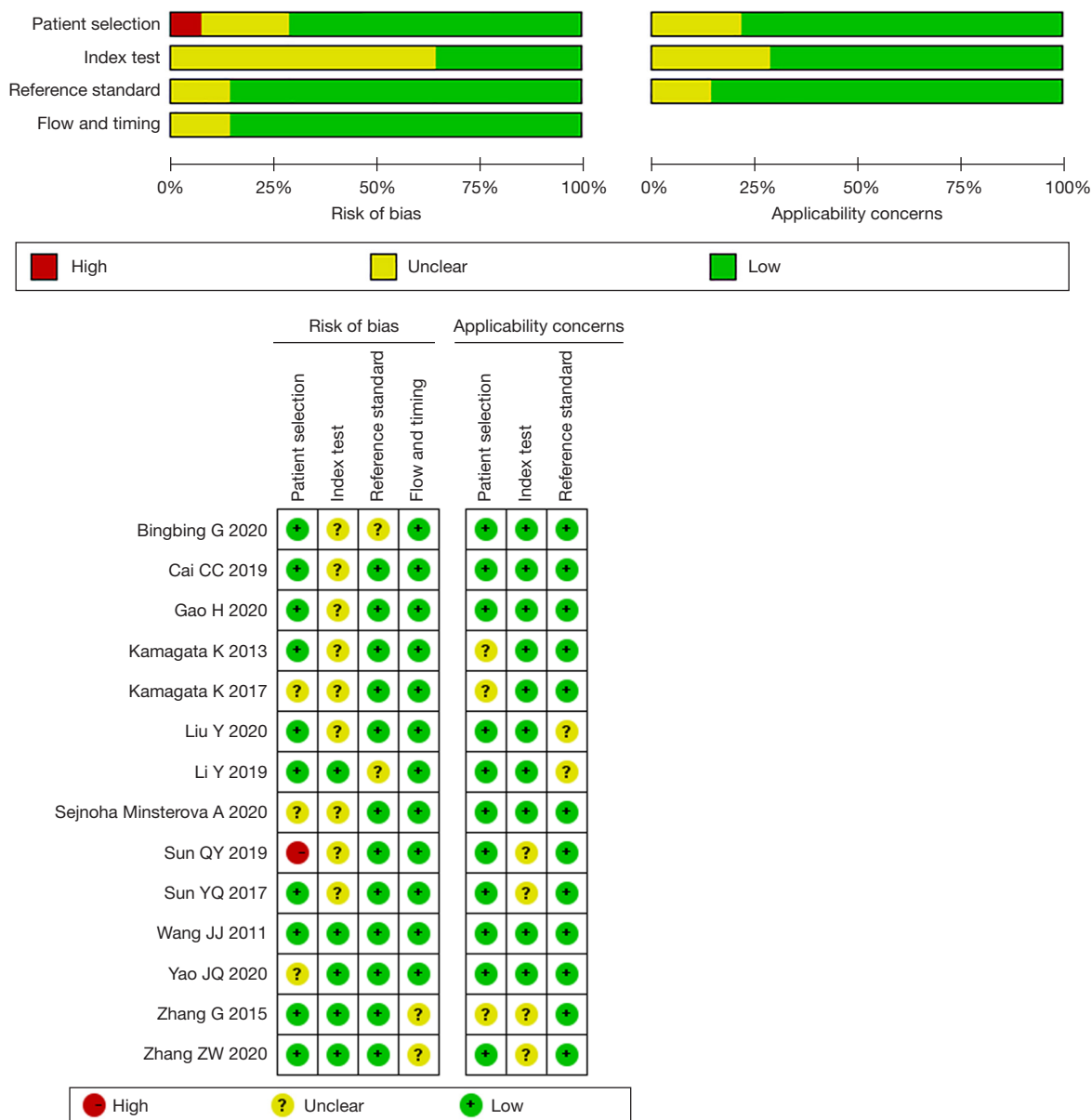


Figure 2 Quality evaluation bar chart and summary.

the logarithm was -0.262 , $P=0.365>0.05$. The above results suggested that the included studies have no threshold effect, as shown in Figure 3.

To test the heterogeneity caused by non-threshold effects, the DOR was used as the effect size. The Q test showed that the χ^2 value of DOR was 30.43, degree of freedom (df) value =13, $P=0.0041$, $I^2=57.3\%$ (see Figure 4 for details), suggesting that the included studies had high heterogeneity ($P<0.05$, $I^2>50\%$), so the random effects model was selected for the combined effect size.

Consolidated statistics and analysis

The random effects model results obtained by Meta-Disc 1.4 software based on the data extracted from the literature and the heterogeneity test results showed the following: Sen =0.78 [95% confidence interval (CI): 0.74–0.81], Spe =0.83 (95% CI: 0.79–0.86), +LR =4.26 (95% CI: 3.22–5.65), -LR =0.28 (95% CI: 0.21–0.37), DOR =17.82 (95% CI: 10.46–30.36), SROC AUC =0.8870, and Q^* value =0.8176 [standard error (SE)(Q^*) =0.0210]. The AUC value showed

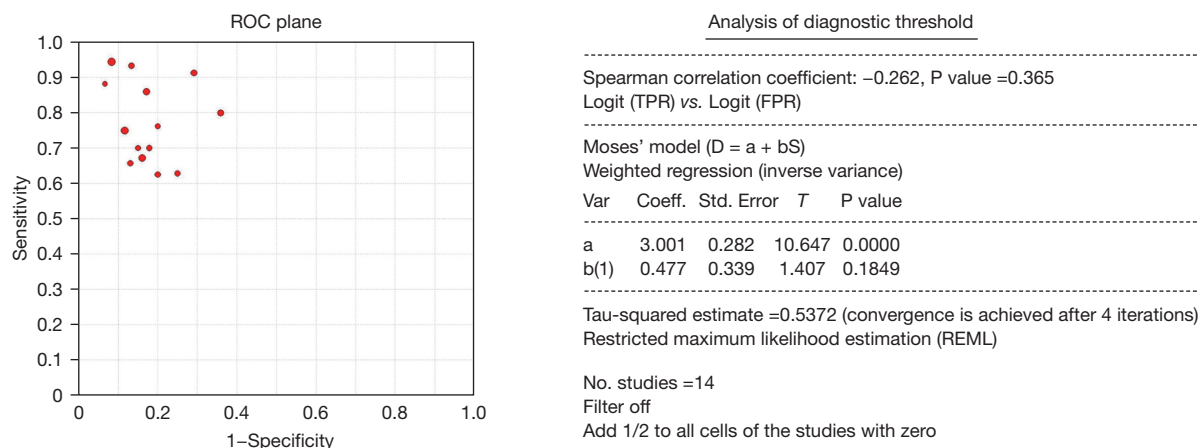


Figure 3 ROC plan and threshold effect test results. a, true positives; b, false positives; S, a measure of threshold; Coeff., coefficient; D, the log of DOR; DOR, diagnostic odds ratio; ROC, receiver operating characteristic; Std. Error, standard error; FPR, false positive rate; TPR, true positive rate; Var., variable.

that DKI has a high diagnostic accuracy for patients with PD, and Sen and Spe were high, which further indicates a high diagnostic value. Sen, Spe, +LR, and -LR forest plots, as well as SROC curves are shown in *Figure 4*.

Analysis and treatment of heterogeneity sources

We used the meta-regression application of Meta-Disc 1.4 software to explore the source of heterogeneity of the included literature, and the choice of compound variables were as follows: sample size (take 60 cases as the boundary; marked “1” if ≥ 60 and “0” if < 60), reference standards (marked as “1, 2, 3, 4” according to four different standards), the delineation features of the area of interest (position, size, calculation, and measurement method of the region of interest (ROI), etc.; when all aspects are described in detail, marked as “1”, and marked as “0” when one of them is not specifically described or explained), the detection part (marked as “1” when the detection part is substantia nigra; otherwise, marked as “0”), and b-value (documents with ≤ 3 b-values are recorded as “0”, and documents with > 3 b-values are recorded as “1”). The regression model results are shown in *Figure 5*. From these results, it can be seen that the source of heterogeneity between the included studies may be related to the reference standard for PD diagnosis in the literature ($P < 0.05$).

The 14 included studies were divided into two subgroups (A and B) according to the different reference standards. Group A included studies with reference standard of 1, while group B included studies with reference standard of

not 1, and the DOR was combined again. Subgroup analysis showed that the heterogeneity of the DOR in group A was $I^2 = 0.0\%$ and $P = 0.4296$, and that of group B was $I^2 = 53.1\%$ and $P = 0.0369$. The heterogeneity of the DOR was significantly reduced after the subgroup analysis. On the other hand, these findings demonstrate that the compound variable of the subgroup analysis was related to the source of the heterogeneity; that is, the reference standard for the included cases in the literature (see *Figure 6*).

Publication bias

The publication bias results of DKI in the diagnosis of PD were as follows: the analysis results obtained through the Deeks funnel chart and quantitative analysis of the P value (*Figure 7, Table 3*) showed that $P = 0.25 > 0.05$, and the scattered points on both sides of the figure are roughly symmetrical, indicating that there was no obvious publication bias.

Discussion

In traditional magnetic resonance (MR) inspection, DWI considers the diffusion of water molecules to be an ideal Gaussian dispersion model. The coefficient detected by DWI of the change in biological tissue signal is known as the apparent diffusion coefficient (ADC), and the calculation of the ADC value usually needs to fit more than two different b-values. ADC value measurement usually assumes that the imaging voxel has a uniform diffusion coefficient. However, in biological tissues, cell membranes,

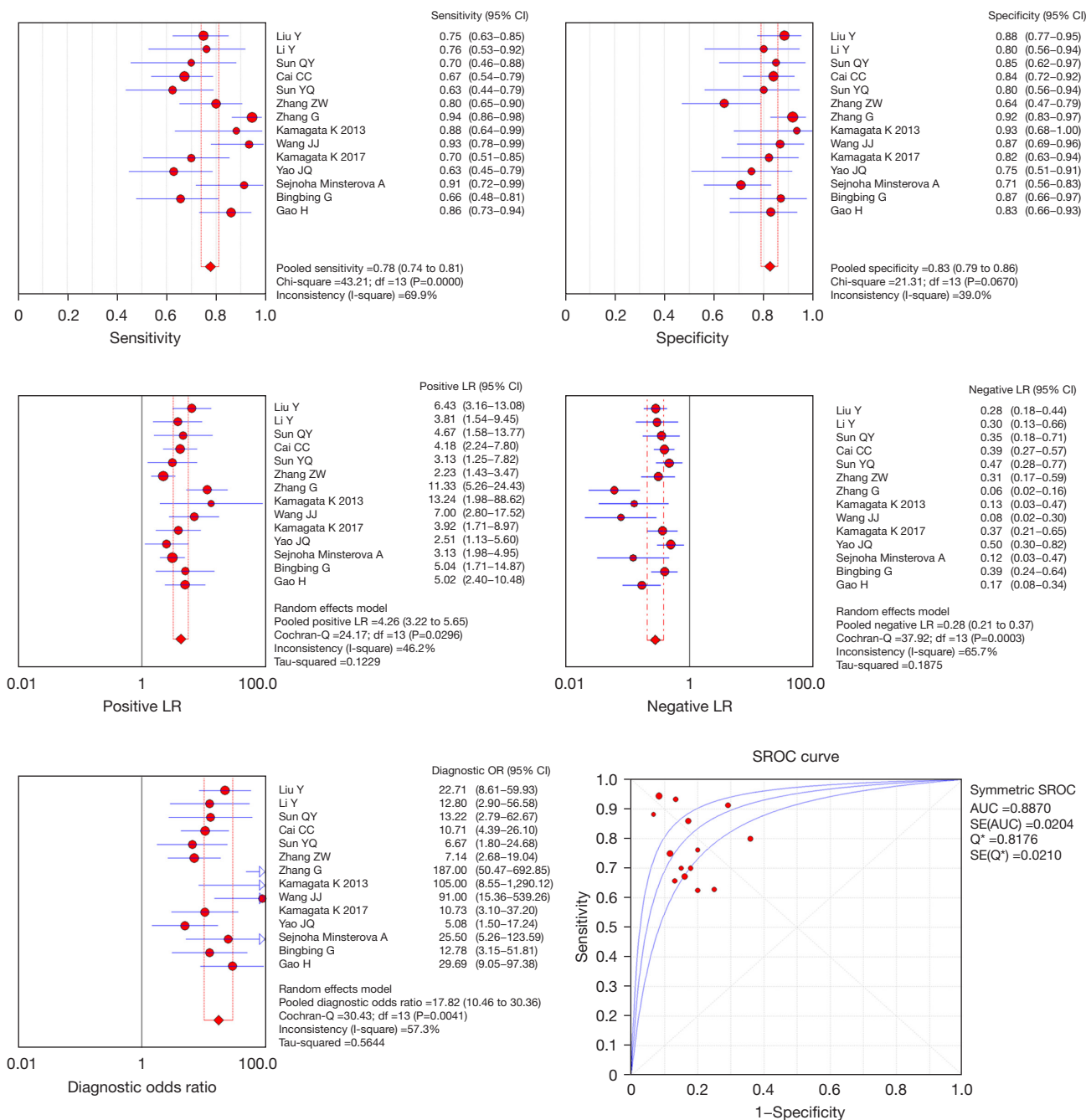


Figure 4 Combined statistics. AUC, area under the curve; CI, confidence interval; LR, likelihood ratio; OR, odds ratio; SE, standard error; SROC, summary receiver operating characteristic.

organelles, cell compartments, and macromolecular structures may change the free diffusion of water molecules. Due to the existence of biological tissue structural barriers, the displacement of water molecules deviates from the Gaussian distribution and manifests as non-Gaussian

diffusion (31). Therefore, traditional DWI cannot truly evaluate the diffusion of water molecules affected by the tissue structure barrier.

With the improvement of MRI equipment, the application of ultra-high b-values has revealed the existence

of non-Gaussian effects, so a pulse sequence capable of sensitively analyzing this complex diffusion mode is required (32). Jensen *et al.* (10,33) often used polynomial models to quantify the non-Gaussian diffusion rate of water molecules, and this new imaging sequence is known as DKI. It is a pulse sequence developed according to the characteristics of diffusion imaging, and applies a polynomial model to fit the imaging advantages of each

sequence, and extends the developed pulse sequence. By quantifying the non-Gaussian diffusion rate of tissue water molecules, DKI reflects the degree of mixing caused by the diffusion of water molecules due to the tissue structure barrier. Therefore, it can more accurately reflect the free diffusion of water molecules in the tissue, explain the heterogeneity in the tissue, and evaluate the corresponding pathogenic area more comprehensively.

A previous study (34) conducted routine DKI examinations on 60 normal people and found that the MK, FA, and MD values of different structures in the brain were significantly associated with changes in age, and the correlation of each parameter value corresponding to the same part is also different. Indicates that DKI has the potential to assess structural changes in the brain caused by age differences. Lätt *et al.* (35) conducted DKI examinations of normal people between 19–64 years of age and found that the MK value in the brain varies depending on the anatomical area, and negatively correlated with age changes. This demonstrates that DKI can better reflect the changes in the microstructure of the brain with age, and these characteristics make it applicable to the diagnosis of PD patients. Wang *et al.* (15) first applied DKI to the clinical diagnosis of PD patients. Through ROI quantitative and ROC curve analysis, they found that the MK values of the substantia nigra and basal ganglia of PD patients were significantly higher than those of healthy volunteers. Compared with traditional DTI parameters, the MK value of substantia nigra showed better diagnostic performance.

The MK value is the most widely used and valuable parameter in DKI. Most studies have shown that the MK value of PD patients is higher than that of healthy controls (15,18,19,21,24,26). This may be related to the increased complexity of the tissue structure, the summary includes the following reasons: (I) due to the nerves of the basal ganglia,

Meta-regression (inverse variance weights)					
Var.	Coeff.	Std. Err.	P value	RDOR	(95% CI)
S	0.259	0.4057	0.5359	–	–
Sample	0.622	0.6432	0.3543	1.86	(0.45; 7.67)
S	0.476	0.3581	0.2107	–	–
Region	–0.077	0.6570	0.9086	0.93	(0.22; 3.93)
S	0.529	0.3597	0.1694	–	–
Part	–0.406	0.6492	0.5440	0.67	(0.16; 2.78)
S	0.103	0.3041	0.7411	–	–
Standard	0.685	0.2467	0.0179	1.98	(1.15; 3.42)
S	0.512	0.3663	0.1898	–	–
b	–0.287	0.8786	0.7503	0.75	(0.11; 5.19)

No. studies =14
 Filter off
 Add 1/2 to all cells of the studies with zero

Figure 5 Regression model results. Coeff., coefficient; CI, confidence interval; S, threshold effect; Std. Err., standard error; RDOR, relative diagnostic odds ratios; Var., variable.

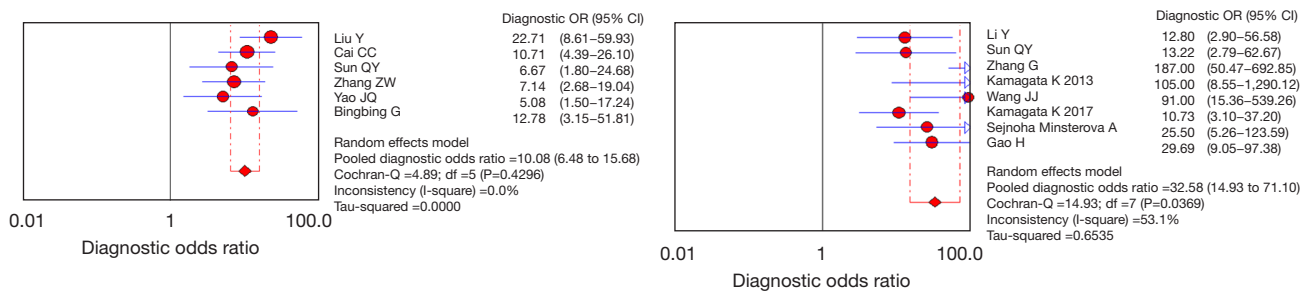


Figure 6 DOR forest map of groups A and B. CI, confidence interval; DOR, diagnostic odds ratio; OR, odds ratio.

loss of cells leads to secondary gliosis in this area, which increases the complexity of the microstructure of local tissues and increases the MK value (15); (II) oxidative stress and chronic inflammation lead to limited water molecule diffusion, resulting in a higher MK value; and (III) it may also be due to the increase of iron content in the substantia nigra, which reduces the signal-to-noise ratio, resulting in an increase in the MK value. However, according to the research of Kamagata *et al.* and Gao *et al.*, the MK value of the basal nucleus in the PD group was lower than that of the healthy control group, which may be related to the loss of cortical-subcortical dopaminergic neurons in PD patients, resulting in complex tissue structure (16,23). Therefore, the mechanism of change of the MK value of the basal nerve nucleus in PD patients requires further investigation.

The MD value reflects the overall diffusion level and diffusion resistance of water molecules in the tissue structure. Eight of the 14 included articles reported that the MD value of PD patients was higher than that of the

healthy controls. The reason may be that although oxidative stress and chronic inflammation can lead to increased local complexity. However, this change may still be lower than the degree of normal brain tissue relationship, so the MD value will also increase.

The FA value mainly reflects the unevenness of the speed and direction of the diffusion of water molecules in tissue. Ten included articles showed that the FA value of the PD group decreased, which may be attributable to the loss of dopaminergic neurons and the destruction of tissue structure. The diffusion of molecules is more inclined to anisotropy, and may also be related to the selection of PD patients. If the selected PD patients are mostly in the early stage, the pathological changes are not significant, which will lead to a lower FA value. At the same time, iron content increases and deposits, and these factors can also cause the FA value to decrease.

Meta-analysis is a systematic evaluation method that integrates the results of similar research topics, expands the sample size, and comprehensive quantitative to improve the efficiency of the experimental findings (36). In recent years, the application field of DKI has become more and more extensive. A previous meta-analysis confirmed the application value of DKI in glioma grading, with a combined AUC value of 0.94, Sen of 0.85, Spe of 0.92, as well as a high diagnostic accuracy (37). To our knowledge, in the context of DKI becoming a hotspot in central nervous system diseases research, there are currently no large sample size studies that evaluate the value of DKI in the diagnosis of PD. In our study, we collected all published Chinese and English articles and satisfied strict inclusion standards and quality assessments, with the aim of increasing the sample size and integrating multiple similar research data to summarize and evaluate the currently available evidence regarding the comprehensive diagnostic efficacy of DKI in PD.

The choice of meta-analytical statistical methods is determined by the heterogeneity between studies. In this study, we used Meta-Disc 1.4 software to test the heterogeneity. By analyzing the ROC plan, we found that

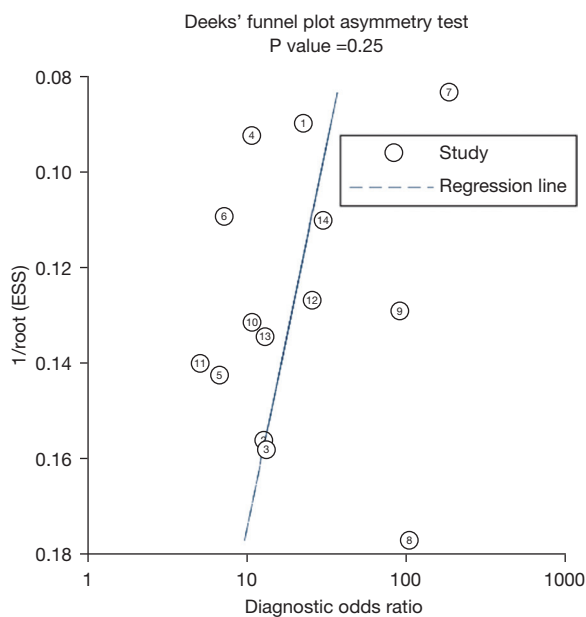


Figure 7 Deeks funnel diagram. ESS, effective sample size.

Table 3 Deeks analysis results

Var.	Coef.	Std. Err.	t	P> t	95% confidence interval
Bias	-14.32646	11.82401	-1.21	0.249	-40.08876, 11.43585
Intercept	4.800384	1.399584	-3.43	0.005	1.750953, 7.849815

Coef., coefficient; P, P value; Std. Err., standard error; t, Student's *t*-test; Var., variable.

the distribution of each study did not exhibit a “shoulder-arm-like”. The Spearman correlation coefficient was -0.262 ($P=0.365>0.05$), which also showed that there was no obvious threshold effect in this study. The χ^2 value corresponding to the DOR Q statistic test was 30.43, the df value was 13, the P value was 0.0041, and the I^2 statistic test result was 57.3%. Since the I^2 value was greater than 50%, this suggested that there was high heterogeneity, and the random effects model was used to merge the effect size. Regression analysis was carried out using Meta-Disc 1.4 software, and the selection of composite variables included the sample size, reference standard, delineation features of the area of interest, detection location, b-value, etc. Only one compound variable was included in the regression model at a time to ensure the stability of the model estimation results. The results suggested that the reference standard for PD diagnosis may be a factor causing heterogeneity between the included studies ($P<0.05$). After subgroup analysis with reference standard for PD diagnosis in the literature as effect variable, it was found that the heterogeneity of DOR among the subgroups was decreased to different degrees. This also confirms that the heterogeneity may be related to the reference standard of PD diagnosis in each study. Differences in the reference standards between studies will have a certain impact on the selection of patients to be included. In the Sen analysis, we excluded the study of Zhang *et al.* (14) and then conducted a combined effect size analysis. We found that the I^2 of the DOR was significantly reduced ($I^2=25.1\%$), and the heterogeneity was also markedly lower. By excluding the remaining literature one-by-one, we observed that the combined effect size did not change, indicating that the source of heterogeneity may have also been attributable to a single study. However, due to the large sample sizes of the included articles, the quality of the literature evaluation was relatively good. After the overall analysis, this article was considered to have certain research significance, and it was not excluded.

In evaluating the accuracy of diagnosis, the AUC value can intuitively reflect the value of diagnostic tests. The AUC ranges from 0 to 1; when the AUC is infinitely close to 1.00 [i.e., the comprehensive ROC curve (SROC) infinitely approaches the upper left corner of the image], it can be considered that the diagnostic accuracy is higher. Moreover, it is generally believed that an AUC value of between 50% and 100% is meaningful, with 50%, 70%, and 90% considered as cut-off points, indicating low, medium, and high diagnostic accuracy, respectively. Put

simply, the closer the AUC value is to 100%, the higher the diagnostic accuracy. In this study, we found that AUC value of the combined 14 studies was 0.8870, which illustrates that DKI has a high diagnostic accuracy in PD and is an ideal biomarker for the clinical diagnosis of PD. Also, the combined Sen and Spe were 0.78 and 0.83, respectively, which indicates that DKI has a good diagnostic value for the diagnosis of PD. The DOR value is used to evaluate the accuracy of diagnosis, integrating Sen and Spe data into an independent parameter value. A DOR value of 1.0 signifies that the test being evaluated cannot clearly distinguish between diseased individuals. In this meta-analysis, the DOR measurement value of DKI was 17.82 (95% CI: 10.46–30.36), indicating good accuracy. Generally, the DOR value is reasonably constant, regardless of the diagnostic threshold. However, due to the particularity of its value, DOR remains difficult to directly apply in clinical practice.

The LR is not affected by prevalence, and can simultaneously reflect the comprehensive indicators of Sen and Spe, which provides a good evaluation of the value of diagnostic experiments. The +LR of this meta-analysis was 4.26, and the –LR was 0.28, which indicates that the positive test result of PD patients is at least four times that of patients who are misdiagnosed as PD, and thus, negative DKI scanning cannot be used alone. Moreover, further testing and/or continuous monitoring is required to exclude PD inspections and the possibility of disease. The results of each combined index showed that DKI has a good application prospect as a routine MRI examination to evaluate clinical PD.

However, this study had some limitations that should be noted. Firstly, the various diagnostic gold standards used in the included studies may have a certain impact on the selection of cases in the study. Secondly, for studies containing multiple parameter values, there may be a certain deviation in the selection of parameter values included in the study. Considering that this study mainly analyzed the overall diagnostic performance of DKI, we comprehensively considered our study and selected the relevant parameter values with higher AUC values to be included in the study. Among the documents that included the four-grid table data, there were nine articles for the MK value, nine articles for the FA value, and two articles for the MD value, which may be different from the results of the individual parameter evaluation.

Conclusions

The results of this meta-analysis demonstrate that DKI is

a more sensitive examination technique for the diagnosis of PD and has a higher diagnostic accuracy. Owing to its advantages of no radiation and no need to inject exogenous contrast agents, it is more suitable for patients with PD who have renal insufficiency and require long-term follow-up. DKI and other related examinations complement each other and have a high application value for PD diagnosis as well as clinical treatment guidance.

Acknowledgments

Funding: The study was supported by Zibo Key Research and Development Plan (Nos. 2021ZC010169, 2021ZC010263).

Footnote

Reporting Checklist: The authors have completed the PRISMA-DTA reporting checklist. Available at <https://atm.amegroups.com/article/view/10.21037/atm-22-1461/rc>

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at <https://atm.amegroups.com/article/view/10.21037/atm-22-1461/coif>). All authors report that this study was supported by Zibo Key Research and Development Plan (Nos. 2021ZC010169, 2021ZC010263). The authors have no other conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

References

1. Tolosa E, Wenning G, Poewe W. The diagnosis of Parkinson's disease. *Lancet Neurol* 2006;5:75-86.
2. de Lau LM, Breteler MM. Epidemiology of Parkinson's disease. *Lancet Neurol* 2006;5:525-35.
3. Adler CH, Beach TG, Hentz JG, et al. Low clinical diagnostic accuracy of early vs advanced Parkinson disease: clinicopathologic study. *Neurology* 2014;83:406-12.
4. Rizzo G, Copetti M, Arcuti S, et al. Accuracy of clinical diagnosis of Parkinson disease: A systematic review and meta-analysis. *Neurology* 2016;86:566-76.
5. Bingbing G, Yujing Z, Yanwei M, et al. Diffusion Kurtosis Imaging of Microstructural Changes in Gray Matter Nucleus in Parkinson Disease. *Front Neurol* 2020;11:252.
6. Berg D, Godau J, Walter U. Transcranial sonography in movement disorders. *Lancet Neurol* 2008;7:1044-55.
7. Bouwmans AE, Vlaar AM, Mess WH, et al. Specificity and sensitivity of transcranial sonography of the substantia nigra in the diagnosis of Parkinson's disease: prospective cohort study in 196 patients. *BMJ Open* 2013;3:e002613.
8. Huang KH, Wu L, Tian YY. Application Progress of Positron Emission Computed Tomography in the Diagnosis of Parkinson's Disease. *Medical Recapitulate* 2020;26:559-64.
9. Viallon M, Cuvinciu V, Delattre B, et al. State-of-the-art MRI techniques in neuroradiology: principles, pitfalls, and clinical applications. *Neuroradiology* 2015;57:441-67.
10. Jensen JH, Helpern JA. MRI quantification of non-Gaussian water diffusion by kurtosis analysis. *NMR Biomed* 2010;23:698-710.
11. Arab A, Wojna-Pelczar A, Khairnar A, et al. Principles of diffusion kurtosis imaging and its role in early diagnosis of neurodegenerative disorders. *Brain Res Bull* 2018;139:91-8.
12. Guglielmetti C, Veraart J, Roelant E, et al. Diffusion kurtosis imaging probes cortical alterations and white matter pathology following cuprizone induced demyelination and spontaneous remyelination. *Neuroimage* 2016;125:363-77.
13. Zhuo J, Xu S, Proctor JL, et al. Diffusion kurtosis as an in vivo imaging marker for reactive astrogliosis in traumatic brain injury. *Neuroimage* 2012;59:467-77.
14. Zhang G, Zhang Y, Zhang C, et al. Diffusion Kurtosis Imaging of Substantia Nigra Is a Sensitive Method for Early Diagnosis and Disease Evaluation in Parkinson's Disease. *Parkinsons Dis* 2015;2015:207624.
15. Wang JJ, Lin WY, Lu CS, et al. Parkinson disease: diagnostic utility of diffusion kurtosis imaging. *Radiology* 2011;261:210-7.
16. Kamagata K, Zalesky A, Hatano T, et al. Gray Matter Abnormalities in Idiopathic Parkinson's Disease:

- Evaluation by Diffusional Kurtosis Imaging and Neurite Orientation Dispersion and Density Imaging. *Hum Brain Mapp* 2017;38:3704-22.
17. Sun QY. The application value of various magnetic resonance imaging techniques on the gray matter nuclei of patients with early Parkinson's disease. Dissertation. Jilin: Beihua University, 2019.
 18. Sejnoha Minsterova A, Klobusiakova P, Pies A, et al. Patterns of diffusion kurtosis changes in Parkinson's disease subtypes. *Parkinsonism Relat Disord* 2020;81:96-102.
 19. Liu Y, Wang H, Ma J. The value of magnetic resonance diffusion tensor imaging and diffusion kurtosis imaging in the diagnosis of Parkinson's disease. *Journal of Medical Imaging* 2020;30:358-62.
 20. Cai CC. Study on the structural changes of brain tissue in patients with different motor subtypes of Parkinson's disease based on magnetic resonance diffusion kurtosis imaging. Dissertation. Fuzhou: Fujian Medical University, 2019.
 21. Li Y, Si H, Tian Y, et al. Application value of diffusion kurtosis imaging in the early diagnosis of Parkinson's disease. *Chinese Journal of Magnetic Resonance Imaging* 2019;10:486-90.
 22. Kamagata K, Tomiyama H, Motoi Y, et al. Diffusional kurtosis imaging of cingulate fibers in Parkinson disease: comparison with conventional diffusion tensor imaging. *Magn Reson Imaging* 2013;31:1501-6.
 23. Gao H. Voxel-based analysis of brain magnetic resonance diffusion kurtosis imaging in patients with Parkinson's disease. Dissertation. Dalian: Dalian Medical University, 2020.
 24. Sun YQ, Zhang Q, Jiang Z, et al. Magnetic resonance diffusion kurtosis imaging analysis and clinical significance of deep brain nuclei in patients with Parkinson's disease. *National Medical Journal of China* 2017;97:3534-7.
 25. Yao JQ. Changes in the microstructure of important nuclei and cingulate fibers in the brain of Parkinson's patients in DKI. Dissertation. Urumchi: Xinjiang Medical University, 2020.
 26. Zhang ZW, Li YF, Pan JL, et al. The value of diffusion kurtosis imaging parameters in the diagnosis of Parkinson's disease. *Chinese Journal of General Practice* 2020;18:273-6.
 27. Wu L, Zhang Y, Zeng X. Application of QUADAS-2 as a quality assessment tool for diagnostic accuracy research. *Journal of Hubei University of Medicine*. 2013;32:201-8.
 28. Zhang TS, Zhong WZ. Meta-DiSc Software in Meta-Analysis of Diagnostic Test. *The Journal of Evidence-Based Medicine* 2008;(2):97-100.
 29. Kang D, Hong Q, Liu G, et al. Identification and treatment of publication bias in meta-analysis. *Chinese Journal of Evidence-Based Medicine* 2003;3:45-9.
 30. Zhang J, Xu Z, Li K. Evaluation of effect indicators in meta-analysis of diagnostic tests. *Chinese Journal of Evidence-Based Medicine* 2013;13:890-5.
 31. Veraart J, Van Hecke W, Sijbers J. Constrained maximum likelihood estimation of the diffusion kurtosis tensor using a Rician noise model. *Magn Reson Med* 2011;66:678-86.
 32. Rosenkrantz AB, Padhani AR, Chenevert TL, et al. Body diffusion kurtosis imaging: Basic principles, applications, and considerations for clinical practice. *J Magn Reson Imaging* 2015;42:1190-202.
 33. Jensen JH, Helpert JA, Ramani A, et al. Diffusional kurtosis imaging: the quantification of non-gaussian water diffusion by means of magnetic resonance imaging. *Magn Reson Med* 2005;53:1432-40.
 34. Zheng H, Zhang H, Wang X, et al. Diffusion kurtosis imaging study of age-related changes in normal adult brain structure. *Chinese Journal of Integrative Medicine on Cardio-Cerebrovascular Disease* 2013;11:564-7.
 35. Lätt J, Nilsson M, Wirestam R, et al. Regional values of diffusional kurtosis estimates in the healthy brain. *J Magn Reson Imaging* 2013;37:610-8.
 36. Zou X, Chen W, Ling L. Statistical problems and doubts in common medical papers (5): key issues of meta-analysis. *Organ Transplantation* 2014;5:380-2.
 37. Falk Delgado A, Nilsson M, van Westen D, et al. Glioma Grade Discrimination with MR Diffusion Kurtosis Imaging: A Meta-Analysis of Diagnostic Accuracy. *Radiology* 2018;287:119-27.
- (English Language Editor: A. Kassem)

Cite this article as: Li Q, Cao J, Liu X, Luo X, Su G, Wang D, Lin B. The diagnostic value of diffusion kurtosis imaging in Parkinson's disease: a systematic review and meta-analysis. *Ann Transl Med* 2022;10(8):474. doi: 10.21037/atm-22-1461