Neurological Sciences (2019) 40:947–956
https://doi.org/10.1007/s10072-019-03732-7

ORIGINAL ARTICLE

Diagnostic performance of DWI for differentiating primary central nervous system lymphoma from glioblastoma: a systematic review and meta-analysis

Xiaoyang Lu[1] · Weilin Xu[1] · Yuyu Wei[1] · Tao Li[1] · Liansheng Gao[1] · Xiongjie Fu[1] · Yuan Yao[1] · Lin Wang[1]

Received: 17 June 2018 / Accepted: 18 January 2019 / Published online: 31 January 2019

Abstract

Objective The purpose of this meta-analysis was to evaluate the diagnostic performance of diffusion-weighted imaging (DWI) for differentiating primary central nervous system lymphoma (PCNSL) from glioblastoma (GBM).

Materials and methods A thorough search of the databases including PubMed, EMBASE, and Cochrane Library was carried out and the data acquired were up to November 1, 2017. The quality of the studies involved was evaluated using QUADAS-2 (Quality Assessment of Diagnostic Accuracy Studies, revised version). Multiple analytic values including sensitivity (SEN), specificity (SPE), positive likelihood ratio (PLR), negative likelihood ratio (NLR), diagnostic odds ratio (DOR), and the summary receiver operating characteristic (SROC) curve were calculated and pooled for the statistical analysis. The subgroup analysis was also performed to explore the heterogeneity.

Results Eight retrospective studies (461 patients with 461 lesions) were included. The pooled SEN, SPE, PLR, NLR, and DOR with 95% confidence interval (CI) were 0.82 [95% CI 0.70–0.90], 0.84 [95% CI 0.75–0.90], 4.96 [95% CI 3.20–7.69], 0.22 [95% CI 0.13–0.37], and 22.85 [95% CI 10.42–50.11], respectively. The area under the curve (AUC) given by SROC curve was 0.90 [95% CI 0.87–0.92]. The subgroup analysis indicated the slice thickness of the images (> 3 mm versus ≤ 3 mm) was a significant factor affecting the heterogeneity. No existence of significant publication bias was confirmed with Deeks' test.

Conclusions DWI showed moderate diagnostic performance for differentiating primary central nervous system lymphoma (PCNSL) from glioblastoma (GBM). Moreover, it is of clinical significance using DWI combined with conventional MRI to differentiate PCNSL from GBM.

Keywords DWI · Lymphoma · Glioblastoma · Meta-analysis

Introduction

Glioblastoma (GBM) is the most common malignant tumor of the primary malignancies in the central nervous system (CNS) [1]. This malignant neoplasm almost inevitably recurs regardless of a complete surgical resection combined with a score of advanced therapies or aggressive treatment. Primary central nervous system lymphoma (PCNSL) is malignant but rare one of the non-Hodgkin's lymphomas that involve the nervous system; however, the incidence of PCNSL has continued to increase in recent years, especially among elderly patients [2]. This may be explained by the improvement of diagnostic techniques and the increasing population of immune compromised individuals [3].The treatment strategy of PCNSL or GBM is totally different. GBM is treated with extended surgical resection followed by radiation therapy and/or chemotherapy with temozolomide, whereas PCNSL is dealt with chemotherapy using methotrexate [4, 5]. Conventional magnetic resonance imaging (MRI) is capable of differentiating PCNSL from GBM with a satisfactory performance only if the PCNSL and GBM are both typical on imaging since a typical PCNSL demonstrates a nodular homogeneous enhancement without necrosis or hemorrhage while the GBM

Xiaoyang Lu, Weilin Xu and Yuyu Wei contributed equally to this work.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10072-019-03732-7) contains supplementary material, which is available to authorized users.

✉ Lin Wang
dr_wang@zju.edu.cn

[1] Department of Neurosurgery, Second Affiliated Hospital, Zhejiang University School of Medicine, NO.88 Jiefang Rd, Zhejiang 310009, Hangzhou, China

usually appears as an annular enhancement around the central necrosis. However, this mode is lacking in credibility in some cases because an atypical enhancing GBM without visible necrosis may be just alike PCNSL. Similarly, an atypical PCNSL with visible necrosis is often indistinguishable from GBM [6]. Therefore, it is a priority to find a more reliable method to differentiate GBM from PCNSL for the implementation of rational treatment strategies for those patients. Diffusion-weighted imaging (DWI) is a type of MR functional imaging. It is used to measure the Brownian motion of water molecules in tissues, which is affected by cell structures such as membranes and myelin fiber bundles, and the apparent diffusion coefficient (ADC) from DWI can reflect the speed of such motion of water molecules. DWI has been widely used to identify different brain tumors in recent years, for instance, distinguishing between PCNSL and GBM. Some study results [7–9] have reported that the DWI is useful for differentiating PCNSL from GBM. However, the others have presented inconsistent results, which might be due to the divergence of parameters of DWI in individual studies or the limited number of cases involved [10, 11]. Therefore, the purpose of this meta-analysis is to further evaluate the diagnostic performance of the DWI combined with quantitative ADC in differentiating PCNSL from GBMS.

Materials and methods

This systematic review and meta-analysis was performed according to the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-analyses) guidelines (Table S1) [12].

Search strategy

A search for all relevant studies based on online databases including PubMed, EMBASE, and Cochrane Library was initiated, and the results were until November 1, 2017. The MeSH terms and free-text words utilized were as follows: ("diffusion tensor imaging" or "DTI" or "diffusion weighted imaging" or "DWI" or "apparent diffusion coefficient" or "ADC" or "diffusion MR" or "diffusion imaging") and ("glioma" or "glioblastoma" or "brain neoplasm" or "brain tumor" or "central nervous system tumor") and ("lymphoma"). After the initial search, the potential articles were additionally identified by conducting a hand search of the references from the included studies.

Eligibility criteria and selection of studies

The inclusion criteria were (1) studies assessing the effectiveness of DWI in differentiating between PCNSL and GBM; (2) the quantitative ADC values for PCNSL and GBM already measured; (3) sufficient information available for calculating true-positive (TP), false-positive (FP), false-negative (FN), and true-negative (TN) values for statistical analysis; (4) PCNSL or GBM lesions confirmed by histopathology and/or biopsy; and (5) published studies with original data in peer-reviewed journals.

The exclusion criteria were (1) studies were reviews, abstracts, letters, comments, editorials, animal studies, or case reports and (2) the number of patients in studies was less than 10.

When multiple publications with encompassing study populations were identified, we included only the study with the largest population.

Two researchers (Weilin Xu, Liansheng Gao) examined all relevant studies by reading the titles and abstracts or the full articles independently. Any discrepancies concerning the study selection were resolved by group discussion among three co-first authors.

Data collection and quality assessment

The following information regarding the studies were extracted: (1) general information (first author, date of publication, patient's age, gender, number of patients, number of PCNSL and GBM lesions); (2) types of study design (prospective or retrospective); (3) parameters of imaging used for DWI (magnetic field strength, *b* values, slice thickness, region of interest (ROI) placement); and (4) mean ADC, min ADC, and ADC ratio (ratio of the ADC of the enhancing lesion to that of the contralateral normal-appearing white matter), $ADC_{5\%}$ (fifth percentile value of cumulative ADC histogram), ADCt (ADC of the most strongly enhanced tumor area), cutoff value of ADC, and TP, FP, FN, TN, sensitivity (SEN), and specificity (SPE).

Two authors (Xiaoyang Lu, Yuyu Wei) independently assessed each study for quality and potential bias using the statistical tool, QUADAS-2 [13] (Quality Assessment of Diagnostic Accuracy Studies, revised version), where each item shall be marked as "yes" or "no," or "unclear" if there was no adequate information to make an accurate judgment. Disagreements were resolved by the corresponding author. The process was performed with Revman 5.3.

Statistical analysis

Meta-analyses were performed using the software, STATA 14.0 (StataCorp LP, College Station, TX). First of all, we evaluated the threshold effect by employing the Spearman correlation coefficient between the logit of SEN and the logit of (1 − SPE). The threshold effect was affirmative if the value of $p < 0.05$. Then, a chi-squared test value and the inconsistency index (I^2) were used to assess the heterogeneity across each study, with $I^2 \leq 50\%$ indicating no severe heterogeneity

while $I^2 > 50\%$ indicating significant heterogeneity [14]. The SEN, SPE, positive likelihood ratio (PLR), negative likelihood ratio (NLR), and diagnostic odds ratio (DOR), all of which with 95% confidence interval (CI) of the included studies were pooled using bivariate mixed effects models, setting a $p < 0.05$ to be statistically significant [15]. We also calculated the symmetric receiver operator characteristic curve (SROC) and the area under curve (AUC). AUC values less than 0.50 suggest that the diagnostic test is ineffective; the values ranging from 0.51 to 0.70 indicate that the diagnostic performance is effective but low. The range from 0.71 to 0.90 presents a moderate diagnostic performance, while more than 0.90 means that the performance is high enough for the test to be implemented into real practice [16].

Subgroup analysis

Due to the existence of heterogeneity of the included studies, we made subgroup analysis aimed to find the heterogeneity. We also calculated the pooled weighted sensitivity and specificity of the subgroups: subgroup based on slice thickness (> 3 mm versus ≤ 3 mm), subgroup based on ROI placement (solid portion versus whole tumor), subgroup based on mean age of patients (≥ 60 years old versus < 60 years old), and the one based on ADC measurements (mean ADC value versus the other ADC values). In order to ensure the accuracy of the subgroup analysis, the minimum number of each study required to form a subgroup shall not be less than three.

Publication bias analysis

The probability of publication bias was evaluated using Deeks' funnel plot asymmetry test. Ordinarily, a low probability of publication bias was considered when the p value was less than 0.05 [17].

Results

Literature selection

A thorough search for literature initially yielded 269 articles. After 19 duplicates were removed, we screened the titles and abstracts of the remaining 250 ones. Next, 225 were excluded (167 were not relevant to the field, 33 were review articles, 12 case reports, 5 animal studies and 8 were letters or editorials). Finally, we obtained 25 potentially eligible articles. When a full-text search was done, 17 articles were omitted because 12 of them did not provide sufficient data to calculate the diagnostic test values and 5 of them did not have enough number of patients ($n \leq 10$). Ultimately, 8 studies [18–25] in accordance with the eligibility criteria were adopted in our meta-analysis. The article selection flow was displayed in Fig. 1.

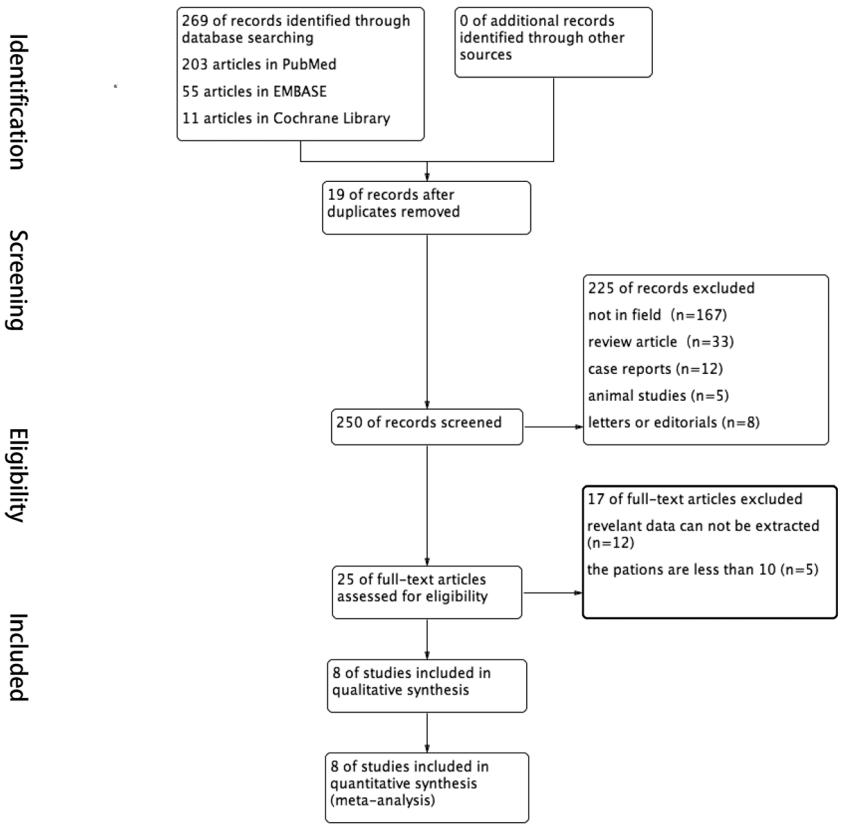

Fig. 1 Flow diagram for study selection process

Characteristics and quality of included studies

In this meta-analysis, a total of 461 patients with 461 lesions were included from 8 studies (133 PCNSL lesions and 328 GBM lesions). Details of the patient information, MRI features, and other study data were documented in Tables 1 and 2. All the eight studies were retrospective cohort studies. The mean age of patients was ≥ 60 years old in four studies while < 60 years old in four studies. The MRI results in four studies were reported to be performed with Siemens scanner, two were with Philips scanner, and two with GE. The field strengths, according to the studies, were described as 3.0 T in five studies, 3.0 T/1.5 T in one study, and 1.5 T in one study. The combination of *b* values was [0, 1000] in 6 studies, [0, 1000, 4000] in one study, and [0, 500, 1000] in one study. Slice thickness of MR scanning was > 3 mm in five studies and ≤ 3 mm in three studies. The ROIs were placed in the solid portion of the tumor in four studies and in whole tumor in the rest four studies. As for the types of diffusion MRI, DWI was used in all studies. The mean ADC value was measured to distinguish PCNSL from GBM in 3 studies; the min ADC value was in 2 studies, the $ADC_{5\%}$ value in one study, the ADC ratio in one study, and the ADCt value in the remaining one. The results of the quality assessment were shown in Figs. 2 and 3. Most of the studies had a low or unclear risk of bias. In general, the quality of all studies was acceptable.

Heterogeneity analysis and quantitative synthesis

There was no threshold effect since the *p* value equaled 0.16. The *Q* test indicated that the heterogeneity was present ($Q = 5.39$, $p = 0.03$). The Higgins I^2 statistics suggested that there was a significant heterogeneity among the studies ($I^2 = 63$). The SEN and SPE were ranging from 0.59 to 1.00 and from 0.64 to 0.92, respectively. The forest plots of SEN, SPE, PLR, NLR, and DOR for 8 studies were depicted in Fig. 4. The pooled statistical values were as follows: SEN = 0.82 [95% CI 0.70–0.90], SPE = 0.84 [95% CI 0.75–0.90], PLR = 4.96 [95% CI 3.20–7.69], NLR = 0.22 [95% CI 0.13–0.37], and DOR = 22.85 [95% CI 10.42–50.11]. The AUC shown by Fig. 5 was 0.90 [95% CI 0.87–0.92], indicating a moderate diagnostic performance in the differentiation of PCNSL from GBM.

Subgroup analysis

The results of subgroup analyses were shown in Table 3. For any potential sources of the heterogeneity, the variations in slice thickness (> 3 mm versus ≤ 3 mm) were confirmed to be a significant one ($p = 0.03$). Other factors, including ROI placement ($p = 0.63$), ADC measurements ($p = 0.18$) and mean age ($p = 0.55$) were insignificantly contributory to the heterogeneity. For the subgroup based on slice thickness,

Table 1 Imaging parameters and technical characteristics of primary diagnostic studies

Author	Year	Country	No. of patients	Gender		Age(years)		No. of PCNSL	No. of GBM	Study design	DWI/DTI	*b* values (s/mm^2)	Magnet strength (T)	Vendor	Slick thickness ((mm)	ROI placement
				F	M	Mean	Range									
Toh et al.	2008	Taiwan	20	11	9	52.6	22–75	10	10	re	DWI	0,1000	3	Siemens,	3	Solid portion
Doskaliyev et al.	2012	Japan	24	13	11	60.0	22–76	10	14	re	DWI	0,1000,4000	3	GE	6	Solid portion
Yamashita et al.	2013	Japan	56	/	/	60.6	42–79	19	37	re	DWI	0,1000	3	Philips	5	Solid portion
Ahn et al.	2014	Korea	87	43	44	57.7	44–77	25	62	re	DWI	0,1000	3	Philips	2	Whole tumor
Ko et al.	2016	Netherlands	126	59	67	59.8	38–77	22	104	re	DWI	0,500,1000	1.5	Siemens	5	Solid portion
Lin et al.	2017	USA	54	21	33	68.6	47–84	18	36	re	DWI	1000	1.5 or 3	GE	5	Whole tumor
Lu et al.	2017	China	60	21	39	56.1	17–75	18	42	re	DWI	0,1000	3	Siemens	5	Whole tumor
Nakajima et al.	2014	Japan	34	17	17	60.9	39–79	11	23	re	DWI	0,1000	3	Siemens	3	Whole tumor

Annotation: *T* tesla, *PCNSL* primary central nervous system lymphoma, *GBM* glioblastoma, *re* retrospective, *F* female, *M* male

Table 2 Characteristic of primary diagnostic studies

Author	Year	Sensitivity	Specificity	TP	FP	FN	TN	Cutoff value
Toh et al.	2008	1.00	0.90	10	1	0	9	meanADC 0.818×10^{-3} mm^2/s
Doskaliyev et al.	2012	0.91	0.92	9	1	1	13	ADCmin 0.50×10^{-3} mm^2/s
Yamashita et al.	2013	0.59	0.87	11	5	8	32	ADCmin 0.62×10^{-3} mm^2/s
Ahn et al.	2014	0.85	0.90	21	6	4	56	ADCmean 0.98×10^{-3} mm^2/s
Ko et al.	2016	0.84	0.64	18	37	4	67	ADCt 0.77×10^{-3} mm^2/s
Lin et al.	2017	0.69	0.89	12	4	6	32	ADCmean 1.3×10^{-3} mm^2/s
Lu et al.	2017	0.76	0.83	14	7	4	35	rADC 1.317
Nakajima et al.	2014	1.00	0.74	11	6	0	17	$ADC_{5\%}$ 0.68×10^{-3} mm^2/s

Annotation: *min* minimum, *rADC* ADC ratio, $ADC_{5\%}$ fifth percentile value of cumulative ADC histogram, *ADCt* ADC of most strongly enhanced tumor area, *TP* true-positive, *FP* false-positive, *FN* false-negative, *TN* true-negative

studies using smaller slice thickness (≤ 3 mm) yielded a higher SEN (0.93 [95% CI 0.84–1.00]) than those using larger thickness. (> 3 mm, SEN = 0.74 [95% CI 0.62–0.85]; $p = 0.01$). The pooled SPE tended to be also higher in studies using smaller slice thickness; however, it did not exhibit a statistical significance ($p = 0.08$). Other subgroup studies, however, did not show convincing statistically significant sensitivity or specificity ($p > 0.05$).

Publication bias

Deek's funnel plot asymmetry test (Fig. 6) for the overall analysis showed that there was no significant publication bias ($p = 0.08$). Circles in Fig. 6 represented individual studies; the dashed line denoted the regression line.

Discussion

PCNSL and GBM have some differences on conventional imaging due to the different infiltrative characters. GBM is a kind of highly aggressive tumor, and it spreads most commonly through white matter. The origin of GBM and its spread are both mainly within the subcortical white matter. GBM can spread through all the fibrous bundles, such as callosum, uncinate fasciculus, longitudinal bands, audiovisual fiber bundle, and corona radiata [26]. Among which, the corpus callosum is the main fibrous bundle that facilitates the tumors to spread to the opposite brain hemisphere [27]. In addition, GBM is predominant in the frontal and temporal lobes, which may be associated with a higher density of fiber content and fibrous connections in those lobes. On the other hand, PCNSL often involves peri-vascular tissue in the brain, suggesting that the tumor cells may originate from blood lymphocytes [28]. Moreover, its selectivity for such tissue might be related to the specific adhesion molecules in vascular tissue and brain parenchyma. This could explain why PCNSL usually occurs in the pia mater, brain parenchyma, or subependymal space. Although it is uncommon for PCNSL to have necrosis or cystic degeneration, both can happen in some cases and make PCNSL and GBM share same characters on conventional imaging. Therefore, we hope to explore whether DWI is better to differ them when such imaging overlaps exist. In this meta-analysis, we appraised the diagnostic performance of DWI for differentiating PCNSL from GBM. As a result, the SEN, SPE, PLR, NLR, and DOR were 0.82, 0.84, 4.96, 0.22, and 22.85 respectively. The AUC was 0.90, which suggested that DWI could distinguish PCNSL and GBM efficaciously. The DOR is a single performance indicator that combines SEN and SPE of the diagnostic test. Since the pooled DOR of the 8 studies was 22.85 in our research, we had good reasons to deduce that it was of advantage to use the diffusion MR to differentiate PCNSL from GBM. It seems that the PLR and NLR are more predominant in clinical practice for evaluating the diagnostic performance than the SROC curve and the DOR. If the PLR > 10 or the NLR < 0.1 for a test, we can always conclude that the test has a great diagnostic accuracy [29]. Since the value of the PLR was 4.96 in our study, it could be interpreted that the

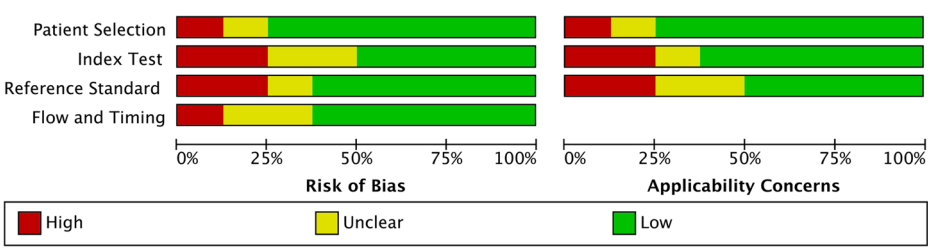

Fig. 2 Methodological quality analysis of the 12 eligible studies using QUADAS-2 tool

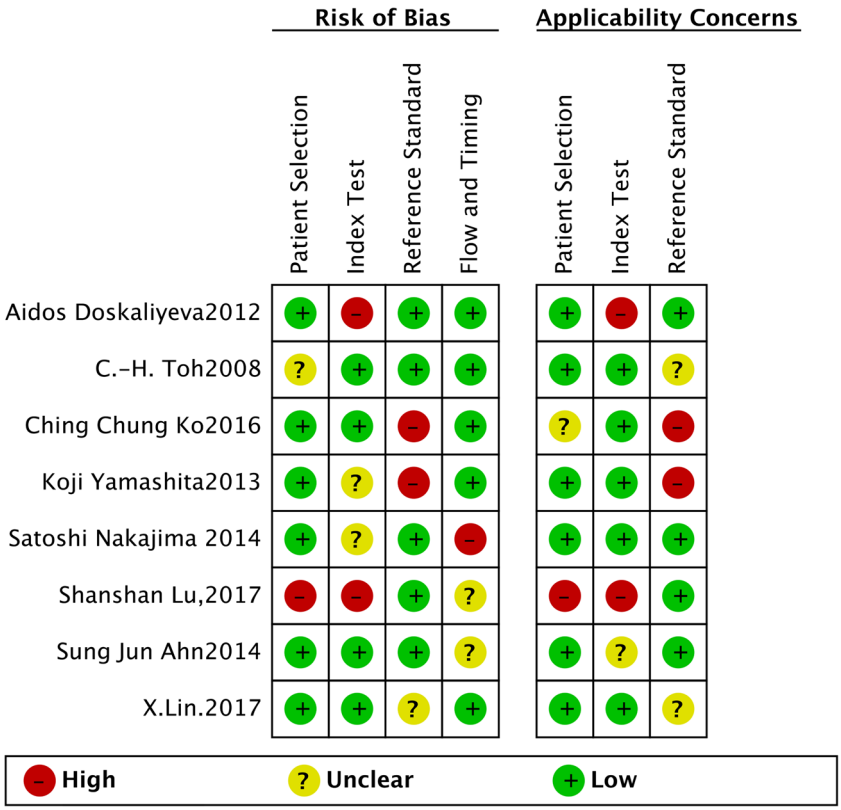

Fig. 3 Methodological quality analysis of the 12 eligible studies using QUADAS-2 tool

patients with PCNSL were approximately five times more likely to have a positive test result than those with GBM. Taking the NLR = 0.22 as a contrast, the probability of a patient with PCNSL being marked as negative (instead of

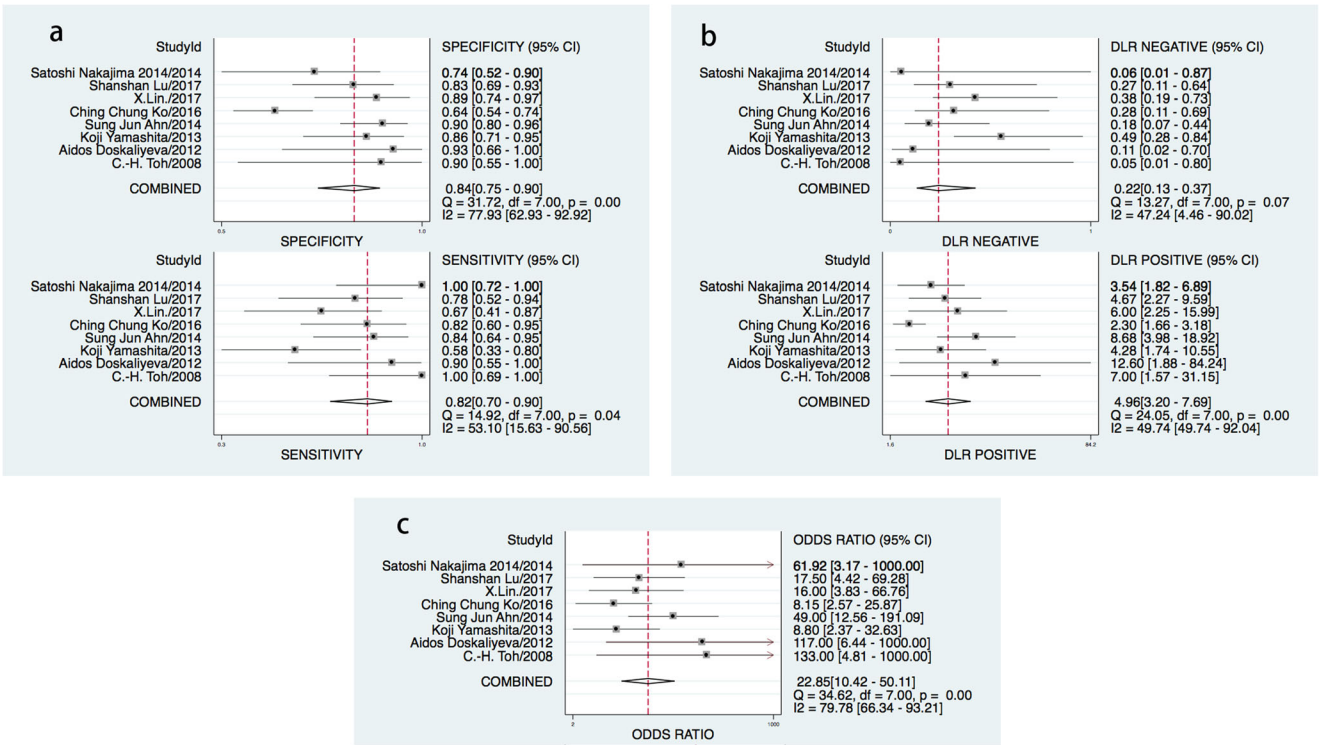

Fig. 4 Coupled forest plots of pooled sensitivity and specificity (**a**), coupled forest plots of pooled positive likelihood ratio, negative likelihood ratio (**b**), coupled forest plots of diagnostic odds ratio (**c**)

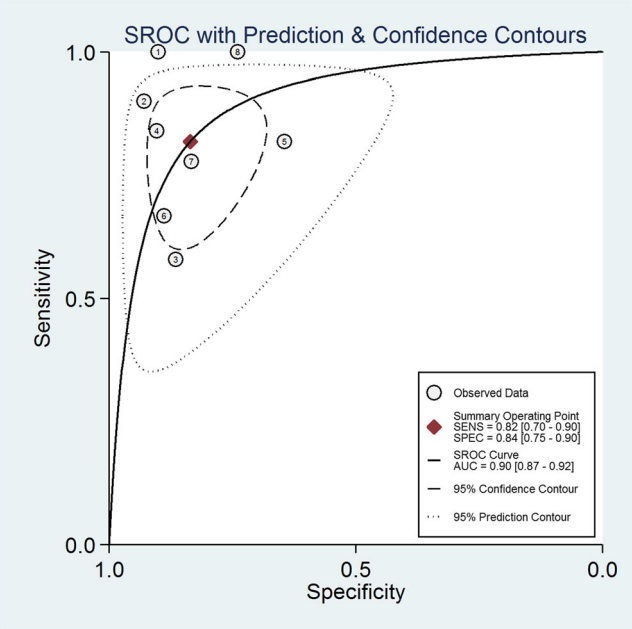

Fig. 5 Summary receiver operating characteristic curve (SROC) of diagnostic performance of DWI for differentiating PCNSL from GBM

positive) was as high as 22%, which was definitely too high to exclude PCNSL.

We also noted that the ADC value of PCNSL was lower than that of GBM. Sugahara et al. [30] reported that the diffusion ability was related in inverse to the cellularity and restricted diffusion was observed in tissue with high cellularity. The PCNSL with a higher cellularity than GBM restricts the motion of water molecules and consequently has a lower ADC value [18, 31]. The difference of the locations of ROIs may lead to the heterogeneity. Setting the ROIs to embrace a whole tumor makes the ADCs available to systematically clarify the issue characteristic and the heterogeneity [32]. It was reported by Ahn et al. [18] that the ADC values extracted from the whole tumor had a higher diagnostic performance. However, results of subgroup analysis in our study did not side with this conclusion; in fact, there was no statistical significance in terms of SEN ($p = 0.5$) or SPE ($p = 0.4$) between ROI placement (in whole tumor versus in solid portion).

Actually, an unnegligible heterogeneity was observed in the studies. The subgroup analyses suggested that the different slice thickness might be the sources of heterogeneity; in the meantime, studies having the slice thickness ≤ 3 mm had significantly higher SEN. As it is aforementioned, the incidence of PCNSL in the elderly is on the rise, partly due to a possible decline in immunological surveillance or an accumulation of somatic mutations over a lifetime [33]. However, in the subgroup based on mean age, we did not find statistical significance of sensitivity or specificity between patients who are 60 years and over and those are below 60. Furthermore, this meta-analysis was the first one to have the diagnostic performance of DWI in differentiating PCNSL from GBM analyzed and have the heterogeneity explored.

In our study, the evaluation of the diagnostic performance of DWI was conducted using a single ADC value (meanADC, minADC, rADC, $ADC_{5\%}$, or ADCt). Lu et al. [22] reported that the ADC ratio in combination with the dynamic contrast-enhanced MRI (DCE-MRI) could boost the diagnostic performance. Nakajima et al. [28] claimed that the $ADC_{5\%}$ was excellent to differentiate PCNSL from GBM. The diagnostic performance using a single ADC value in our meta-analysis was moderate in distinguishing the two types of tumors; hence, a more comprehensive and efficient property was in demand for its use in clinical practice.

In a recent published meta-analysis [34] on the performance of MR perfusion-weighted imaging (PWI) for differentiating of high-grade glioma from PCNSL, the pooled SEN, SPE, and DOR were 0.883 [95% CI 0.848–0.912], 0.837 [95% CI 0.777–0.886], and 53.83 [95% CI 20.048–131.43],

Table 3 Results of pooled estimates of all studies and of different subgroups

Covariate	No. of studies	Sensitivity (95%CI)	P1	Specificity (95% CI)	P2	χ^2	*P*
Slice thickness (mm)							
≥ 5	5	0.74 [0.62–0.85]	0.01	0.82 [0.73–0.91]	0.08	7.14	0.03
< 5	3	0.93 [0.84–1.00]		0.87 [0.80–0.94]			
ROI placement							
Whole tumor	4	0.82 [0.74–0.94]	0.50	0.86 [0.78–0.93]	0.43	0.91	0.63
Solid portion	4	0.80 [0.66–0.95]		0.79 [0.66–0.92]			
ADC measurements							
Mean ADC	3	0.83 [0.69–0.97]	0.51	0.90 [0.82–0.97]	0.67	3.43	0.18
The other ADC values	5	0.80 [0.68–0.93]		0.78 [0.69–0.87]			
Mean age (years)							
≥ 60	4	0.76 [0.61–0.91]	0.07	0.86 [0.77–0.95]	0.30	1.19	0.55
< 60	4	0.85 [0.75–0.95]		0.81 [0.71–0.91]			

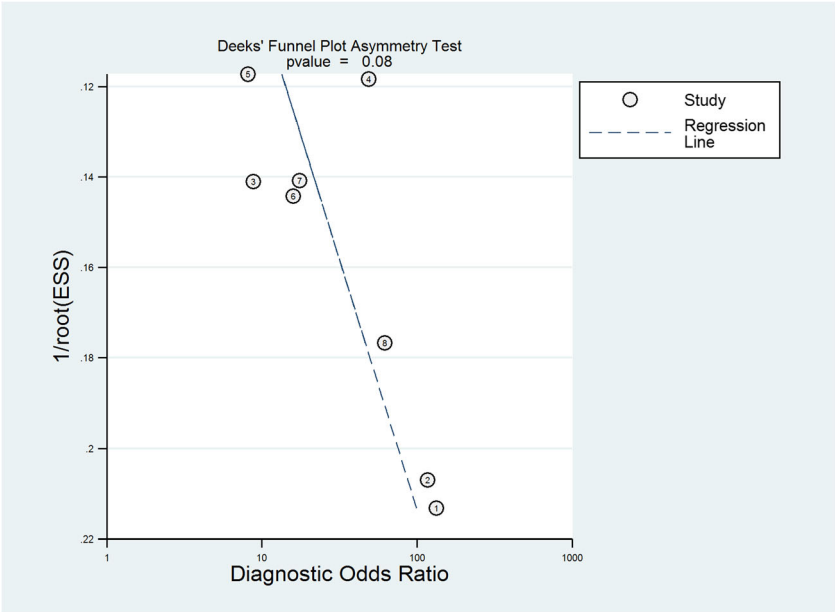

Fig. 6 Funnel plot of Deeks' test results for assessment of publication bias. ESS = effective sample size

respectively. However, the PWI technique requires a contrast medium, and its diagnostic performance can be affected by a variety of factors, therefore having a limited application. In contrast, DWI is irradiative, easily accessible, and less expensive.

One study [35] discussed about the differences between GBM and PCNSL showed in magnetic resonance spectroscopy (MRS). The authors concluded that PCNSL presented significantly higher Glu/Cr ratios and Glu/Glu + Gln ratios when compared to GBM. PCNSL also had significantly higher Cho/Cr ratios as compared to GBM, although the difference for Cho/Cr ratios seemed less significant than that for Glu/Cr or Glu/Glu + Gln ratios. The increased Cho peak in PCNSL was reported to be related to increased cellularity [36]. As we mentioned before, PCNSL has higher cellularity than GBM. However, another study [37] found that no significant difference could be confirmed between PCNSL and GBM for Cho/Cr ratios. Therefore, more studies are needed to evaluate the diagnostic performance of MRS on differentiating PCNSL from GBM.

Clinical practice

This meta-analysis provided reliable evidence that DWI showed moderate diagnostic performance in differentiating PCNSL from GBM. Although the gold standard for the diagnosis of PCNSL still is biopsy, it is an invasive test that can cause severe damage to the patient. Let alone DWI's being noninvasive, it should be pointed out that it is also an convenient, budget, and time-saving method, which makes it more suitable for clinical application.

Limitations

There are some limitations in our study that could not be ignored. First, the number of involved studies was small. Some studies regarding the use of DWI for the differentiation had to be excluded due to a lack of statistics to calculate SEN and SPE [7, 8, 38, 39]. Despite that, those studies showed a significant difference of DWI characteristics (mean ADC value, ADC ratio, and mean diffusivity) between PCNSL and GBM, which is consistent with our results. Second, all the included studies were retrospective. Pooling analysis from retrospective studies could have the diagnostic performance raised [40]. Thus, further prospective studies concerning the diagnostic performance of DWI are needed. Third, unified methodology using DWI should be established, for the divergence of parameters of imaging, variations in field strengths and differences of post-processing software could all lead to a discrepancy of ADC measurement. Fourth, we do not do subgroup analysis based on gender because one of the studies did not discuss the number of male patients and female patients.

Conclusions

This meta-analysis revealed a reliable level of diagnostic performance of the DWI to distinguish PCNSL from GBM. DWI with the slice thickness ≤ 3 mm has significantly higher sensitivity. It is of clinical significance using DWI and quantitative ADC combined with conventional MRI to differentiate PCNSL from GBM. Further prospective studies concerning the diagnostic performance of DWI with more studies and larger population are needed. Along with this, the unified

methodology regarding the use of DWI shall be employed in clinical practice and be included in future analysis.

Acknowledgments This work was supported by the Zhejiang Provincial Natural Science Foundation of China (grant LY18H090007). All the authors confirmed that there is no conflict of interest.

References

1. Deangelis LM (2001) Brain tumors. N Engl J Med 344(344):114–123
2. Norden AD, Jan D, Wen PY, Claus EB (2011) Survival among patients with primary central nervous system lymphoma, 1973-2004. J Neuro-Oncol 101(3):487–493
3. Schultz CJ, Bovi J (2010) Current management of primary central nervous system lymphoma. Int J Radiat Oncol Biol Phys 76(3): 666–678
4. Schlegel U (2009) Primary CNS lymphoma. Ther Adv Neurol Disord 2(2):93–104
5. Stupp R, Mason WP, Van DB, Martin J et al (2008) Radiotherapy plus concomitant and adjuvant temozolomide for glioblastoma. Clin Med Oncol 2(10):421–422
6. Philipp K, Benedikt W, Felix S et al (2014) Primary central nervous system lymphoma and atypical glioblastoma: multiparametric differentiation by using diffusion-, perfusion-, and susceptibility-weighted MR imaging. Radiology 272(3):843–850
7. Kitis O, Altay H, Calli C, Yunten N, Akalin T, Yurtseven T (2005) Minimum apparent diffusion coefficients in the evaluation of brain tumors. Eur J Radiol 55(3):393–400
8. Yamasaki F, Kurisu K, Satoh K, Arita K, Sugiyama K, Ohtaki M, Takaba J, Tominaga A, Hanaya R, Yoshioka H, Hama S, Ito Y, Kajiwara Y, Yahara K, Saito T, Thohar MA (2005) Apparent diffusion coefficient of human brain tumors at MR imaging. Radiology 235(3):985–991
9. Server A, Kulle B, Maehlen J et al (2009) Quantitative apparent diffusion coefficients in the characterization of brain tumors and associated peritumoral edema. Acta Radiol 50(6):682–689
10. Rizzo L, Crasto SG, Moruno PG, Cassoni P, Rudà R, Boccaletti R, Brosio M, de Lucchi R, Fava C (2009) Role of diffusion- and perfusion-weighted MR imaging for brain tumour characterisation. Radiol Med 114(4):645–659
11. Maeda M, Matsushima N, Takeda K (2010) Relationship between FDG uptake and apparent diffusion coefficients in solitary brain tumor. Neuroradiol J 23:390–391
12. Moher D, Liberati A, Tetzlaff J, Altman DG (2009) Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Open Med 3(3):e123
13. Whiting PF, Rutjes AW, Westwood ME et al (2011) QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med 155(8):529–536
14. Higgins JP, Thompson SG (2002) Quantifying heterogeneity in a meta-analysis. Stat Med 21(11):1539–1558
15. Rutjes AWS, Zwinderman AH, Reitsma JB, Glas AS, Bossuyt PM, Scholten RJPM (2005) Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. J Clin Epidemiol 58(10):982–990
16. Deeks JJ, Petra M, Les I (2005) The performance of tests of publication bias and other sample size effects in systematic reviews of diagnostic test accuracy was assessed. J Clin Epidemiol 58(9):882–893
17. Deeks JJ, Macaskill P, Irwig L (2005) The performance of tests of publication bias and other sample size effects in systematic reviews of diagnostic test accuracy was assessed. J Clin Epidemiol 58(9): 882–893
18. Ahn SJ, Shin HJ, Chang JH, Lee SK (2014) Differentiation between primary cerebral lymphoma and glioblastoma using the apparent diffusion coefficient: comparison of three different ROI methods. PLoS One 9(11):e112948
19. Doskaliyev A, Yamasaki F, Ohtaki M, Kajiwara Y, Takeshima Y, Watanabe Y, Takayasu T, Amatya VJ, Akiyama Y, Sugiyama K, Kurisu K (2012) Lymphomas and glioblastomas: differences in the apparent diffusion coefficient evaluated with high b-value diffusion-weighted magnetic resonance imaging at 3T. Eur J Radiol 81(2):339–344
20. Ko CC, Tai MH, Li CF, Chen TY, Chen JH, Shu G, Kuo YT, Lee YC (2016) Differentiation between glioblastoma multiforme and primary cerebral lymphoma: additional benefits of quantitative diffusion-weighted MR imaging. PLoS One 11(9):e0162565
21. Lin X, Lee M, Buck O, Woo KM, Zhang Z, Hatzoglou V, Omuro A, Arevalo-Perez J, Thomas AA, Huse J, Peck K, Holodny AI, Young RJ (2017) Diagnostic accuracy of T1-weighted dynamic contrast-enhanced-MRI and DWI-ADC for differentiation of glioblastoma and primary CNS lymphoma. AJNR Am J Neuroradiol 38(3):485–491
22. Lu S, Wang S, Gao Q, Zhou M, Li Y, Cao P, Hong X, Shi H (2017) Quantitative evaluation of diffusion and dynamic contrast-enhanced magnetic resonance imaging for differentiation between primary central nervous system lymphoma and glioblastoma. J Comput Assist Tomogr 41(6):898–903
23. Nakajima S, Okada T, Yamamoto A, Kanagaki M, Fushimi Y, Okada T, Arakawa Y, Takagi Y, Miyamoto S, Togashi K (2015) Primary central nervous system lymphoma and glioblastoma: differentiation using dynamic susceptibility-contrast perfusion-weighted imaging, diffusion-weighted imaging, and (18)F-fluorodeoxyglucose positron emission tomography. Clin Imaging 39(3):390–395
24. Toh CH, Castillo M, Wong AM et al (2008) Primary cerebral lymphoma and glioblastoma multiforme: differences in diffusion characteristics evaluated with diffusion tensor imaging. AJNR Am J Neuroradiol 29(3):471–475
25. Yamashita K, Yoshiura T, Hiwatashi A, Togao O, Yoshimoto K, Suzuki SO, Abe K, Kikuchi K, Maruoka Y, Mizoguchi M, Iwaki T, Honda H (2013) Differentiating primary CNS lymphoma from glioblastoma multiforme: assessment using arterial spin labeling, diffusion-weighted imaging, and (1)(8)F-fluorodeoxyglucose positron emission tomography. Neuroradiology 55(2):135–143
26. Matsukado Y, Maccarty CS, Kernohan JW (1961) The growth of glioblastoma multiforme (astrocytomas, grades 3 and 4) in neurosurgical practice. J Neurosurg 18(5):636–644
27. Scherer HJ (1940) The forms of growth in gliomas and their practical significance. Brain Part 63(1):1–35
28. Aho R, Ekfors T, Haltia M, Kalimo H (1993) Pathogenesis of primary central nervous system lymphoma: invasion of malignant lymphoid cells into and within the brain parenchyme. Acta Neuropathol 86(1):71–76
29. Zhang H, Ma L, Wang Q, Zheng X, Wu C, Xu BN (2014) Role of magnetic resonance spectroscopy for the differentiation of recurrent glioma from radiation necrosis: a systematic review and meta-analysis. Eur J Radiol 83(12):2181–2189
30. Sugahara T, Korogi Y, Kochi M, Ikushima I, Shigematu Y, Hirai T, Okuda T, Liang L, Ge Y, Komohara Y, Ushio Y, Takahashi M (1999) Usefulness of diffusion-weighted MRI with echo-planar technique in the evaluation of cellularity in gliomas. J Magn Reson Imaging 9(1):53–60
31. Horger M, Fenchel M, Nagele T et al (2009) Water diffusivity: comparison of primary CNS lymphoma and astrocytic tumor

infiltrating the corpus callosum. AJR Am J Roentgenol 193(5): 1384–1387
32. Kang Y, Choi SH, Kim YJ, Kim KG, Sohn CH, Kim JH, Yun TJ, Chang KH (2011) Gliomas: histogram analysis of apparent diffusion coefficient maps with standard- or high-b-value diffusion-weighted MR imaging–correlation with tumor grade. Radiology 261(3):882–890
33. Villano JL, Koshy M, Shaikh H, Dolecek TA, McCarthy BJ (2011) Age, gender, and racial differences in incidence and survival in primary CNS lymphoma. Br J Cancer 105(9):1414–1418
34. Xu W, Wang Q, Shao A, Xu B, Zhang J (2017) The performance of MR perfusion-weighted imaging for the differentiation of high-grade glioma from primary central nervous system lymphoma: a systematic review and meta-analysis. PLoS One 12(3):e0173430
35. Aburano H, Ueda F, Yoshie Y et al (2015) Differences between glioblastomas and primary central nervous system lymphomas in 1 H-magnetic resonance spectroscopy. Jpn J Radiol 33(7):1–12
36. Poptani H, Gupta RK, Roy R, Pandey R, Jain VK, Chhabra DK (1995) Characterization of intracranial mass lesions with in vivo proton MR spectroscopy. AJNR Am J Neuroradiol 16(8):1593–1603
37. Chawla S, Zhang Y, Wang S, Chaudhary S, Chou C, O'Rourke DM, Vossough A, Melhem ER, Poptani H (2010) Proton magnetic resonance spectroscopy in differentiating glioblastomas from primary cerebral lymphomas and brain metastases. J Comput Assist Tomogr 34(6):836–841
38. Guo AC, Cummings TJ, Dash RC, Provenzale JM (2002) Lymphomas and high-grade astrocytomas: comparison of water diffusibility and histologic characteristics. Radiology 224(1):177–183
39. Calli C, Kitis O, Yunten N, Yurtseven T, Islekel S, Akalin T (2006) Perfusion and diffusion MR imaging in enhancing malignant cerebral tumors. Eur J Radiol 58(3):394–403
40. Deeks JJ (2001) Systematic reviews in health care: systematic reviews of evaluations of diagnostic and screening tests. BMJ 323(7322):157–162