

BMJ Open Accuracy of magnetic resonance imaging texture analysis in differentiating low-grade from high-grade gliomas: systematic review and meta-analysis

Qiangping Wang, Deqiang Lei, Ye Yuan, Hongyang Zhao

To cite: Wang Q, Lei D, Yuan Y, *et al.* Accuracy of magnetic resonance imaging texture analysis in differentiating low-grade from high-grade gliomas: systematic review and meta-analysis. *BMJ Open* 2019;**9**:e027144. doi:10.1136/bmjopen-2018-027144

► Prepublication history and additional material for this paper are available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2018-027144>).

QW and DL contributed equally.

Received 11 October 2018

Revised 11 July 2019

Accepted 19 August 2019



© Author(s) (or their employer(s)) 2019. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

Department of Neurosurgery, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China

Correspondence to

Dr Hongyang Zhao;
zhaohongyangxie@163.com

ABSTRACT

Objectives Texture analysis (TA) is a method used for quantifying the spatial distributions of intensities in images using scanning software. MRI TA could be applied to grade gliomas. This meta-analysis was performed for assessing the accuracy of MRI TA in differentiating low-grade gliomas from high-grade ones.

Methods PubMed, Cochrane Library, Science Direct and Embase were searched for identifying suitable studies from their inception to 1 September 2018. The quality of the studies was evaluated on the basis of the Quality Assessment of Diagnostic Accuracy Studies guidelines. We estimated the pooled sensitivity, specificity, positive likelihood ratio (PLR), negative likelihood ratio (NLR) and diagnostic OR (DOR) using the summary receiver operating characteristic (SROC) for identifying the accuracy of MRI TA in grading gliomas. Fagan nomogram was applied for assessing the clinical utility of TA.

Results Six studies including 440 patients were included and analysed. The pooled sensitivity, specificity, PLR, NLR and DOR with 95% CIs were 0.93 (95% CI 0.88 to 0.96), 0.86 (95% CI 0.81 to 0.89), 6.4 (95% CI 4.8 to 8.6), 0.08 (95% CI 0.05 to 0.15) and 78 (95% CI 39 to 156), respectively. The SROC curve showed an area under the curve of 0.96 (95% CI 0.93 to 0.97). Deeks test confirmed no significant publication bias in all studies. Fagan nomogram revealed that the post-test probability increased by 43% in patients with positive pre-test.

Conclusions The findings of this meta-analysis suggested that MRI TA has high accuracy in differentiating low-grade gliomas from high-grade ones. A standardised methodology is warranted to guide the use of this technique for clinical decision-making.

INTRODUCTION

Gliomas are the most frequently occurring type of primary malignant brain tumour. According to the WHO tumour classification, gliomas are subdivided into grades I to IV, where I to II are low-grade gliomas (LGGs) and III to IV are high-grade gliomas (HGGs).¹ LGG is a low-grade malignant tumour associated with longer life expectancy, while HGG is

Strengths and limitations of this study

- This meta-analysis assesses the accuracy of MRI texture analysis in differentiating low-grade gliomas from high-grade ones.
- The pooled sensitivity and specificity were 0.93 and 0.86, respectively, for MRI texture analysis in differentiating low-grade gliomas from high-grade ones.
- A standardised methodology is warranted to guide the use of this technique for clinical decision-making.

highly aggressive and has a dismal prognosis despite various therapeutic managements.²⁻⁴ Surgical resection is the preferred treatment for most gliomas. Postoperatively, HGG normally requires adjuvant therapy, such as radiotherapy and chemotherapy, to prevent rapid recurrence, while LGG is usually followed by close observation.⁵ Due to the high malignancy of HGG, complete surgical resection of tumour is crucial in the prognosis of patients. Hence, the identification of tumour level before surgery is important for intraoperative decision-making. Histopathological assessment is the current gold standard for grading gliomas, which is an invasive procedure and is generally performed postoperatively. Thus, the potential to accurately ascertain tumour grade by utilising a non-invasive technique is gaining a lot of attention.^{6,7}

MRI is the first-choice of imaging method in detecting gliomas. With the development of technology, several physiological MRI techniques including magnetic resonance (MR) spectroscopy, diffusion-weighted imaging (DWI) and perfusion-weighted imaging (PWI), have also been applied for grading gliomas.^{8,9} Texture analysis (TA) is a method used for quantifying the spatial distributions of intensities in images. Some reports have suggested that TA holds promise in the field

of oncology diagnosis, including quantifying tumour heterogeneity and tumour grading.^{10 11} Until now, some reports have been published regarding tumour heterogeneity in glioma using MRI TA.^{8 11-15} However, these studies were inconclusive because of insufficient samples and different diagnostic algorithms. The present meta-analysis aimed to systematically evaluate the accuracy of TA in discriminating LGGs from HGGs.

METHODS

Patient and public involvement

Since this is a meta-analysis, ethical approval was unnecessary. Patients' priorities, experiences and preferences were not involved in the study design.

Search strategy

This systematic review and meta-analysis was performed following the guidelines for the diagnostic studies.¹⁶

PubMed, Cochrane Library, Science Direct and Embase were searched from their inception to 1 September 2018. The search keywords were 'Texture analysis', 'glioma', 'brain neoplasm' and 'brain tumour'. The search strategy used for the retrieval of studies from the Cochrane Library is presented in online supplementary file 1. The search strategy was modified as deemed necessary for other databases. No language restriction was exposed. Reference lists of relevant articles were also manually searched. Two reviewers independently reviewed the articles. Disagreements were resolved by consensus.

Study selection criteria

The studies were selected on the basis of the following criteria: (1) clinical trials assessing the diagnostic accuracy of TA in differentiating LGGs from HGGs, (2) used histopathology as criterion standard and (3) sufficient information for calculating true-positive (TP), false-positive (FP), true-negative (TN) and false negative (FN) results. The exclusion criteria were animal studies, case reports, abstracts, insufficient calculable data, duplicated reports or studies based on the same study.

One author (Wang QP) conducted the initial search according to the inclusion and exclusion criteria. Next, two investigators (Lei DQ and Yuan Y) independently examined all potentially relevant articles. Disagreements were resolved by consensus.

Date extraction and quality assessment

Two investigators (Wang QP and Lei DQ) independently assessed the quality and potential bias and extracted the data of included studies. We extracted the following data: first author, year of publication, country, sample size, study design (retrospective or prospective), patient age, MRI field strengths, TA tools, TP, FP, TN, FN, sensitivity and specificity values according to tumour grading. LGGs (grade I to II gliomas) were considered positive; HGGs (grade III to IV gliomas) were considered negative. If the TP, FP, TN and FN results were not reported, we

calculated backward using indexes including sensitivity, specificity, positive predictive value and negative predictive value.

The quality of each study was assessed on the basis of the Quality Assessment of Diagnostic Accuracy Studies (QUADAS) guidelines,¹⁷ which is an established, evidence-based tool for systematic reviews of diagnostic studies.

Statistical analysis

Meta-analyses were performed using the software Meta-Disc V.1.4 (Metadisc, Unit of Clinical Biostatistics of Ramón y Cajal Hospital, Madrid, Spain) and Stata V.12.0 (StataCorp LP, College Station, Texas, USA). The pooled sensitivity, specificity, positive likelihood ratio (PLR), negative likelihood ratio (NLR) and diagnostic OR (DOR) were calculated on the basis of bivariate generalised linear mixed modelling using the extracted data of TP, TN, FP and FN. The accuracy of the data was determined using a summary receiver operating characteristic (SROC) plot and summarising the curve by calculating the area under the curve (AUC). Cochran-Q method and inconsistency index (I^2) were adopted for investigating heterogeneity among the studies. The significant heterogeneity was indicated by a p value < 0.05 and I^2 < 50%. Generally, a diagnostic tool is regarded as to have failed when AUC values are between 0.5 and 0.6, poor when AUC values are between 0.6 and 0.7, fair when AUC values are between 0.7 and 0.8, good when AUC values are between 0.8 and 0.9 and excellent when AUC values are between 0.9 and 1.¹⁸ Fagan nomogram and likelihood matrix were used for evaluating the clinical utility of TA.

Subgroup analysis

We calculated the pooled weighted sensitivity and specificity of subgroups for observing the effects caused by substantial heterogeneity of the included studies. Studies were grouped on the basis of the MRI performed at different field strengths (3.0 T vs not 3.0 T), MRI images used (contrast-enhanced T1 and fluid-attenuated inversion recovery (FLAIR) vs DWI) and filtration method (grey level co-occurrence matrices (GLCM) vs Laplacian of Gaussian band-pass filtration).

Publication bias

The publication bias was assessed using Deeks funnel plot asymmetry test, where a p value < 0.05 suggests a potential publication bias. Deeks funnel plot asymmetry test was performed using Stata V.12.0.

RESULTS

Literature research

A total of 125 studies were initially identified using the abovementioned search strategy, which were then screened by title and abstract. Of these, 38 articles were further evaluated in full text. Twenty-nine articles were irrelevant and three could not provide sufficient data

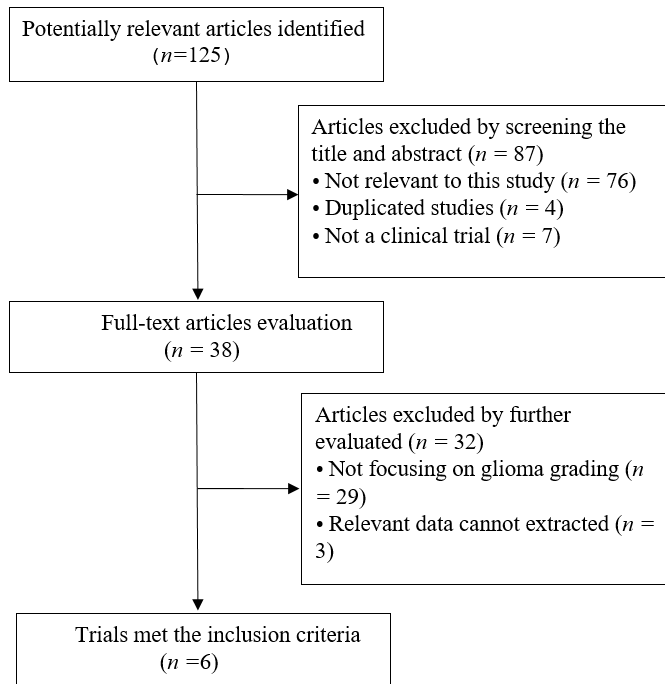


Figure 1 Results of literature search.

to construct the 2x2 table. According to the inclusion criteria, six studies^{8 11–15} were retrieved. The study selection process is shown in [figure 1](#).

Study characteristics

Ultimately, six studies with 440 participants were enrolled in this meta-analysis. The detailed characteristics of included studies are shown in [table 1](#). All studies were retrospective cohort studies. The MR examinations were performed using a 1.5 T scanner in one study, 3.0 T in four studies and one study did not mention the device. Contrast-enhanced T1 images were used for analysis in two studies, contrast-enhanced T1s combined with FLAIR images were used for analysis in two studies and DWI were used for analysis in two studies. Regarding the TA tools, TexRAD software (<http://www.texrad.com>, part of Feedback Plc, Cambridge UK) was used in two studies, and Functional MRI of the Brain’s Software Library (FSL) of analysis tools (Analysis Group, FMRIB, Oxford, UK), Medical Imaging Solution for Segmentation and Texture Analysis (MISSTA, an in-house software of Seoul National University College of Medicine, Seoul, Korea), computer-aided diagnosis (CAD) system and FireVoxel (<https://wp.nyu.edu/firevoxel/>) were used in one research respectively.

Quality of included studies

The quality assessment of included studies using the QUADAS-2 checklist is presented in [table 2](#). For the included studies, ‘index test’ and ‘reference standard’ revealed slight shortcomings (16.7% [1/6] each), which may indicate bias regarding inclusion. Overall, the study quality was satisfactory.

Table 1 Baseline characteristics of included studies

| First author | Year | Country | Study design | No. of patients | No. of HGG | No. of LGG | Field strengths | MRI imaging | Texture analysis | Filteration methods | Reference standard |
|-----------------------------|------|---------|--------------|-----------------|------------|------------|-----------------|--------------------------------|-----------------------|---------------------------------|--------------------|
| Zacharaki ⁸ | 2009 | USA | re | 74 | 52 | 22 | 3.0 T | Contrast-enhanced T1 and FLAIR | FSL of analysis tools | NA | Histology |
| Ryu ⁸ | 2014 | Korea | re | 40 | 32 | 8 | 1.5 T | DWI | MISSTA | GLCM | Histology |
| Skogen ¹² | 2016 | Norway | re | 95 | 68 | 27 | 3.0 T | Contrast-enhanced T1 and FLAIR | TexRAD | Laplacian of Gaussian band-pass | Histology |
| Li-Chun Hsieh ¹³ | 2017 | Taiwan | re | 107 | 34 | 73 | NA | Contrast-enhanced T1 | CAD system | GLCM | Histology |
| Ditmer ¹⁵ | 2018 | USA | re | 94 | 80 | 14 | 3.0 T | Contrast-enhanced T1 | TexRAD | Laplacian of Gaussian band-pass | Histology |
| Wang ¹⁴ | 2018 | China | re | 30 | 18 | 12 | 3.0 T | DWI | FireVoxel | GLCM | Histology |

CAD, computer-aided diagnosis; DWI, diffusion-weighted imaging; FLAIR, Fluid-attenuated inversion-recovery sequence; FSL, Software Library; GLCM, Gray Level Co-occurrence Matrices; HGG, high grade gliomas; LGG, low grade gliomas; MISSTA, Medical imaging solution for segmentation and texture analysis; NA, not mentioned; pro, prospective; re, retrospective; T, Tesla.

Table 2 Results of the QUADAS-2 quality assessment of included studies

| Study | Risk of bias | | | Applicability concerns | | | |
|----------------------------------|-------------------|------------|--------------------|------------------------|-------------------|------------|--------------------|
| | Patient selection | Index test | Reference standard | Flow and timing | Patient selection | Index test | Reference standard |
| Zacharaki EI 2009 | + | - | + | + | + | ? | + |
| Ryu 2014 ⁸ | + | + | ? | + | + | ? | + |
| Skogen 2016 ¹² | + | + | + | + | + | - | ? |
| Li-Chun Hsieh 2017 ¹³ | + | + | - | + | + | + | ? |
| Ditmer 2018 ¹⁵ | + | ? | + | + | + | ? | + |
| Wang 2018 ¹⁴ | + | + | + | + | + | - | + |

+: Low risk; -: High risk; ?: Unclear risk.

QUADAS, Quality Assessment of Diagnostic Accuracy Studies.

Pooled results

The pooled sensitivity and specificity of TA for discriminating LGGs and HGGs were 0.93 (95% CI 0.88 to 0.96) and 0.86 (95% CI 0.81 to 0.89), respectively. The forest plots are shown in figure 2. The pooled PLR and NLR were 0.86 (95% CI 0.81 to 0.89) and 6.4 (95% CI 4.8 to 8.6), respectively. The DOR was 78 (95% CI 39 to 156). SROC curve analysis was used to summarise overall diagnostic accuracy. The AUC was 0.96. The SROC curve is shown in figure 3. The results demonstrated high diagnostic performance in discrimination of LGGs from HGGs.

Subgroup analyses

The results of the subgroup analyses are presented in table 3. The specificity was slightly lower, but the AUC was higher in studies wherein MRI was performed using a 3.0 T scanner than in those where MRI was performed using a 1.5 T scanner. The sensitivity and specificity were significantly higher in studies using contrast-enhanced T1 and FLAIR images than in those using DWI. The diagnostic performance of GLCM was slightly higher than that of Laplacian of Gaussian band-pass filtration.

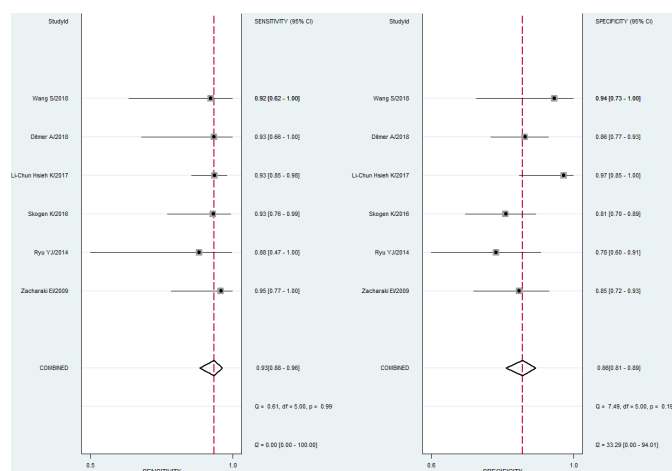


Figure 2 Pooled estimates of sensitivity and specificity of texture analysis to differentiate low-grade gliomas from high-grade ones.

Evaluation of clinical utility

The clinical utility of TA was evaluated by utilising likelihood ratios to simulate a Fagan nomogram. The result is shown in figure 4. With a 25% pre-test probability of LGG, the post-test probabilities of LGG and given positive and negative TA analysis results, are 68% and 3%, respectively. Fagan nomogram revealed that the post-test probability increased by 43% in patients with positive pre-test but decreased by 22% in patients with negative pre-test, which indicated that TA was useful in clinical practice.

Publication bias and heterogeneity

Publication bias was examined using Deeks plot asymmetry test, and the funnel plot did not reveal significant publication bias ($p=0.35$). The funnel plots are shown in figure 5. Heterogeneity among the included studies was

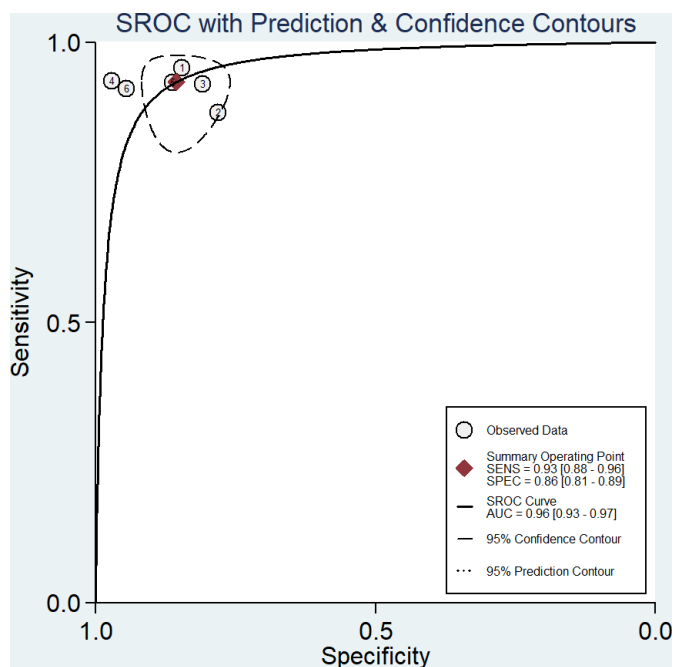


Figure 3 Summary receiver operating characteristics (SROC) curve of texture analysis to differentiate low-grade gliomas from high-grade ones. AUC, area under the curve; SENS, sensitivity; SPEC, specificity.

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

| Studies | N | Sensitivity | Specificity | PLR | NLR | DOR | AUC |
|--|---|-------------|-------------|------|------|-----|------|
| All studies | 6 | 0.93 | 0.86 | 6.4 | 0.08 | 78 | 0.96 |
| MRI performed at 3.0 T | 4 | 0.93 | 0.85 | 6.2 | 0.08 | 78 | 0.96 |
| MRI performed at not 3.0 T | 2 | 0.93 | 0.88 | 10.5 | 0.08 | 107 | 0.50 |
| Image used: contrast-enhanced T1 and FLAIR | 4 | 0.93 | 0.86 | 6.6 | 0.08 | 85 | 0.96 |
| Image used: DWI | 2 | 0.90 | 0.84 | 6.8 | 0.12 | 56 | 0.50 |
| Diagnostic algorithm: GLCM | 3 | 0.92 | 0.89 | 11.6 | 0.08 | 125 | 0.96 |
| Diagnostic algorithm: Laplacian of Gaussian band-pass filtration | 2 | 0.93 | 0.84 | 5.6 | 0.09 | 62 | 0.50 |

AUC, the area under the curve; DOR, diagnostic OR; DWI, diffusion-weighted imaging; FLAIR, Fluid-attenuated inversion-recovery sequence; GLCM, Grey Level Co-occurrence Matrices; NLR, negative likelihood ratio; PLR, positive likelihood ratio.

measured using Cochran-Q method and I^2 . As shown in figure 2, the p value of the Cochran-Q method was >0.05 . The I^2 value of the pooled specificity analysis was 33.29%, which showed slight heterogeneity. The potential

source of the observed heterogeneity was assessed using subgroup analyses.

DISCUSSION

The earliest reports have indicated that TA based on CT images has the potential of differential diagnosis of tumour heterogeneity.^{19,20} To date, there have been some reports on glioma grading using MRI TA.²¹ However, the results have been inconclusive. We conducted this meta-analysis for systematically evaluating the accuracy of TA in discriminating LGGs from HGGs. The findings of the meta-analysis showed that the pooled sensitivity and specificity of TA were 0.93 and 0.86, respectively. The PLR and NLR were 6.4 and 0.08, respectively. The AUC was 0.96. The results demonstrated that TA had high diagnostic performance in ruling out HGGs in discriminating gliomas.

Histopathology assessment is the gold standard for the diagnosis of gliomas, but it is an invasive procedure. To provide accurate information and to avoid unnecessary operations for gliomas, the role of MRI cannot be

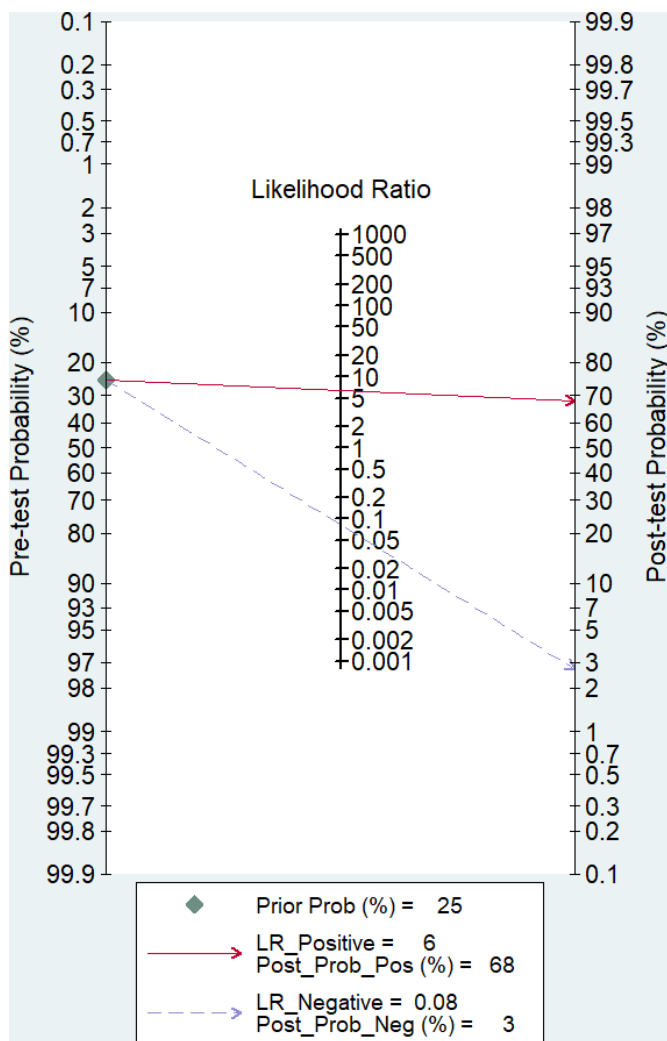


Figure 4 Fagan nomogram for the elucidation of post-test probabilities with a pre-test probability of 25%. LR, likelihood ratio; Prob, probability.

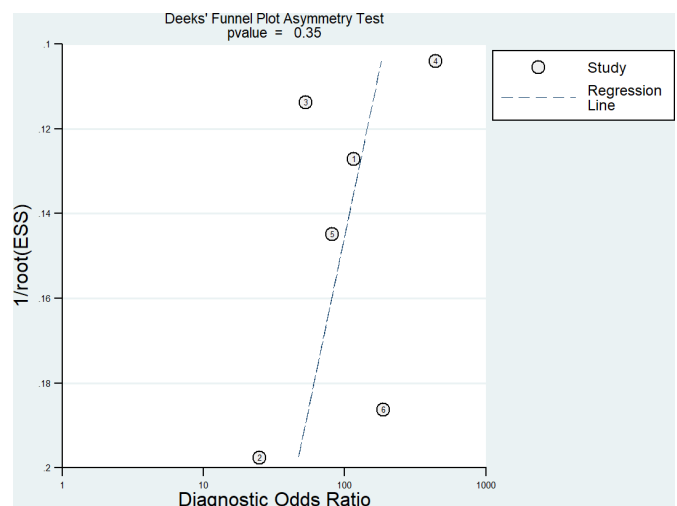


Figure 5 Deeks funnel plots indicating no publication bias ($p=0.35$).

neglected. With the development of techniques, more and more metabolic and physiological MRI, such as diffusion tensor imaging, magnetic resonance spectroscopy, DWI, dynamic susceptibility contrast MRI and dynamic contrast-enhanced MRI, have been utilised in grading gliomas.^{22–24} All these examinations assessed the malignancy of tumours by identifying the difference of in characteristics in the images, such as greyscale brightness and contrast of image pixels. Textures are complex visual patterns composed of entities that have characteristic size, brightness, intensity and so on. Thus, texture can be regarded as a similarity grouping in an image.¹⁵ TA is an integrated analysis of texture using special tools, such as TexRAD, MISSTA, CAD and FireVoxel. Therefore, TA has more powerful diagnostic capability than the ordinary examination method.

In performing TA, the first step is image filtration. The two methods used in the included studies were GLCM and Laplacian of Gaussian band-pass filtration. Although the superiority of the two remains undetermined, the meta-analysis found that the diagnostic performance of GLCM was slightly higher than that of Laplacian of Gaussian band-pass filtration. Quantitative analysis of the filtered pixel values is conducted after the image-filtration step. The parameters include mean of positive pixel values, mean intensity, SD, entropy, skewness and kurtosis.^{25 26} Next, the AUC of the parameters to distinguish tumour grades were calculated by receiver operating characteristic curve analysis.

This review demonstrated that TA was useful in discriminating LGGs and HGGs. In a published meta-analysis based on MR PWI for glioma grading, the pooled sensitivity, specificity and DOR were 93%, 81% and 55%, respectively.²⁷ However, PWI requires the injection of contrast medium and the results are influenced by many factors; therefore, it is difficult to widely use of PWI. In another meta-analysis on the accuracy of MR DWI for glioma grading, the pooled sensitivity, specificity and AUC were 0.85, 0.80 and 0.90, respectively.²⁸ DWI has specific advantages over PWI; it is easily accessible, less expensive and does not require a contrast agent. TA can use any kind of MRI sequences such as PWI, DWI and FLAIR; thus, this technique is easy to use.

However, obvious heterogeneity between studies was noted. Different field strengths (3.0 T and 1.5 T), MRI used (DWI, contrast-enhanced T1 and FLAIR), analysis tools (MISSTA, TexRAD, FireVoxel and FSL of analysis tools) and filtration methods (GLCM and Laplacian of Gaussian band-pass filtration) could affect the accuracy of the conclusion. The procedure should be standardised by conducting further research. The meta-analysis showed that studies employed higher strength (3.0 T), contrast-enhanced T1 and FLAIR imaging and GLCM to perform TA yielding higher diagnostic performance in the discrimination of LGGs from HGGs. Therefore, it is recommended to adopt these techniques for TA in future studies.

It is worth noting that this study had several limitations. First, this systematic review included six studies with 440 patients. Limited studies and participants might have affected the accuracy of the results. Second, although no publication bias was detected in this meta-analysis, the test strength may have been affected by the limited number of studies. Thus, publication bias was also a concern. Lastly, different field strengths, imaging sequences and TA tools were used in the included studies that lack consensus, which influenced the consistency of measurements. Therefore, well-conducted investigations using a standardised methodology are required to confirm the discrimination value of TA on gliomas.

Therefore, our study suggested that TA could be an accurate tool for discriminating gliomas. However, more studies are warranted to verify the most suitable technique. The application of TA with a standardised methodology would improve the accuracy of glioma diagnosis and clinical decision-making in the future.

Contributors Wang QP and Zhao HY conceived and designed the work. Wang QP, Lei DQ and Yuan Y were involved in data collection, data analysis and interpretation. Wang QP drafted the manuscript. Zhao HY involved in critical revision of the article and final approval of the version to be published. All authors have agreed to be accountable for all aspects of the work.

Funding The study was funded by The Funds for Creative Research of Union Hospital, Tongji Medical College, Huazhong University of Science and Technology (02.03.2017-65).

Competing interests None declared.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement All data relevant to the study are included in the article or uploaded as supplementary information.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

REFERENCES

- Louis DN, Ohgaki H, Wiestler OD, *et al*. The 2007 WHO classification of tumours of the central nervous system. *Acta Neuropathol* 2007;114:97–109.
- Johnson DR, O'Neill BP. Glioblastoma survival in the United States before and during the temozolomide era. *J Neurooncol* 2012;107:359–64.
- Helseth R, Helseth E, Johannesen TB, *et al*. Overall survival, prognostic factors, and repeated surgery in a consecutive series of 516 patients with glioblastoma multiforme. *Acta Neurol Scand* 2010;122:159–67.
- McGirt MJ, Woodworth GF, Coon AL, *et al*. Independent predictors of morbidity after image-guided stereotactic brain biopsy: a risk assessment of 270 cases. *J Neurosurg* 2005;102:897–901.
- Woodworth GF, McGirt MJ, Samdani A, *et al*. Frameless image-guided stereotactic brain biopsy procedure: diagnostic yield, surgical morbidity, and comparison with the frame-based technique. *J Neurosurg* 2006;104:233–7.
- Server A, Kulle B, Gadmar Øystein B, *et al*. Measurements of diagnostic examination performance using quantitative apparent diffusion coefficient and proton MR spectroscopic imaging in the preoperative evaluation of tumor grade in cerebral gliomas. *Eur J Radiol* 2011;80:462–70.
- Chaddad A, Kucharczyk MJ, Daniel P, *et al*. Radiomics in glioblastoma: current status and challenges facing clinical implementation. *Front Oncol* 2019;9:374.

8. Ryu YJ, Choi SH, Park SJ, *et al.* Glioma: application of whole-tumor texture analysis of diffusion-weighted imaging for the evaluation of tumor heterogeneity. *PLoS One* 2014;9:e108335.
9. Jackson A, O'Connor JPB, Parker GJM, *et al.* Imaging tumor vascular heterogeneity and angiogenesis using dynamic contrast-enhanced magnetic resonance imaging. *Clin Cancer Res* 2007;13:3449–59.
10. Nie K, Chen J-H, Yu HJ, *et al.* Quantitative analysis of lesion morphology and texture features for diagnostic prediction in breast MRI. *Acad Radiol* 2008;15:1513–25.
11. Georgiadis P, Cavouras D, Kalatzis I, *et al.* Enhancing the discrimination accuracy between metastases, gliomas and meningiomas on brain MRI by volumetric textural features and ensemble pattern recognition methods. *Magn Reson Imaging* 2009;27:120–30.
12. Skogen K, Schulz A, Dormagen JB, *et al.* Diagnostic performance of texture analysis on MRI in grading cerebral gliomas. *Eur J Radiol* 2016;85:824–9.
13. Li-Chun Hsieh K, Chen C-Y, Lo C-M. Quantitative glioma grading using transformed gray-scale invariant textures of MRI. *Comput Biol Med* 2017;83:102–8.
14. Wang S, Meng M, Zhang X, *et al.* Texture analysis of diffusion weighted imaging for the evaluation of glioma heterogeneity based on different regions of interest. *Oncol Lett* 2018;15:7297–304.
15. Ditmer A, Zhang B, Shujaat T, *et al.* Diagnostic accuracy of MRI texture analysis for grading gliomas. *J Neurooncol* 2018;140:583–9.
16. Moher D, Shamseer L, Clarke M, *et al.* Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Syst Rev* 2015;4:1.
17. Whiting PF, Rutjes AWS, Westwood ME, *et al.* QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011;155:529–36.
18. Metz CE. Basic principles of ROC analysis. *Semin Nucl Med* 1978;8:283–98.
19. Ganeshan B, Skogen K, Pressney I, *et al.* Tumour heterogeneity in oesophageal cancer assessed by CT texture analysis: preliminary evidence of an association with tumour metabolism, stage, and survival. *Clin Radiol* 2012;67:157–64.
20. Ng F, Ganeshan B, Kozarski R, *et al.* Assessment of primary colorectal cancer heterogeneity by using whole-tumor texture analysis: contrast-enhanced CT texture as a biomarker of 5-year survival. *Radiology* 2013;266:177–84.
21. Soni N, Priya S, Bathla G. Texture analysis in cerebral gliomas: a review of the literature. *AJNR Am J Neuroradiol* 2019;40:928–34.
22. Yang D, Korogi Y, Sugahara T, *et al.* Cerebral gliomas: prospective comparison of multivoxel 2D chemical-shift imaging proton MR spectroscopy, echo-planar perfusion and diffusion-weighted MRI. *Neuroradiology* 2002;44:656–66.
23. Majós C, Julià-Sapé M, Alonso J, *et al.* Brain tumor classification by proton MR spectroscopy: comparison of diagnostic accuracy at short and long TE. *AJNR Am J Neuroradiol* 2004;25:1696–704.
24. Emblem KE, Nedregaard B, Nome T, *et al.* Glioma grading by using histogram analysis of blood volume heterogeneity from MR-derived cerebral blood volume maps. *Radiology* 2008;247:808–17.
25. Ganeshan B, Goh V, Mandeville HC, *et al.* Non-small cell lung cancer: histopathologic correlates for texture parameters at CT. *Radiology* 2013;266:326–36.
26. Lewis MA, Ganeshan B, Barnes A, *et al.* Filtration-histogram based magnetic resonance texture analysis (MRTA) for glioma IDH and 1p19q genotyping. *Eur J Radiol* 2019;113:116–23.
27. Min Z-gang, Liu H-juan, Li M, *et al.* [Accuracy of MR perfusion weighted imaging for cerebral glioma grading: a meta-analysis]. *Zhonghua Yi Xue Za Zhi* 2010;90:2927–31.
28. Zhang L, Min Z, Tang M, *et al.* The utility of diffusion MRI with quantitative ADC measurements for differentiating high-grade from low-grade cerebral gliomas: evidence from a meta-analysis. *J Neurol Sci* 2017;373:9–15.