BMJ Open

# Accuracy of MRI texture analysis for differentiating high-grade from low-grade gliomas



| | |
|---|---|
| Journal: | *BMJ Open* |
| Manuscript ID | bmjopen-2018-027144 |
| Article Type: | Research |
| Date Submitted by the Author: | 11-Oct-2018 |
| Complete List of Authors: | Qiangping, Wang; Dept. of Neurosurgery, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology<br>Lei, Deqiang; Wuhan Union Hospital, Dept. of Neurosurgery, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology<br>Yuan, Ye; Wuhan Union Hospital, Dept. of Neurosurgery, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology<br>Zhao, Hongyang; Wuhan Union Hospital, Dept. of Neurosurgery, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology; |
| Keywords: | Texture analysis, MRI, glioma, meta-analysis |
| | |

# Accuracy of MRI texture analysis for differentiating high-grade from low-grade gliomas

Qiang-ping Wang, MD[#]; De-qiang Lei, MD[#]; Ye Yuan, MD; Hong-yang Zhao, MD[*].

# Qiang-ping Wang and De-qiang Lei contributed equally to this work.

Dept. of Neurosurgery, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430022, China

[*]Correspondence to Pro. Hong-yang Zhao, Email: zhaohongyangxiehe@163.com; Tel.: +86 134-7685-0638, Fax: +86 028-8712-1111. Address: Dept. of Neurosurgery, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, 1277 JieFang Avenue, Wuhan, China, 430022.

**Key words:** Texture analysis; MRI; glioma; meta-analysis.

**Abstract**

**Objectives:** MRI texture analysis (TA) could be applied to grade gliomas. This meta-analysis was conducted to assess the accuracy of MRI texture analysis in differentiating high-grade from low-grade gliomas.

**Materials and methods:** PubMed, Cochrane library, Science Direct and Embase were searched to identify suitable studies up to Sep 1st, 2018. The quality of studies were evaluated by the quality assessment of diagnostic accuracy studies (QUADAS 2). We estimated the pooled sensitivity, specificity, positive and negative likelihood ratios (LR) and diagnostic accuracy ratio (DOR) using the summary receiver operating characteristic (SROC) to identity the accuracy of MRI texture analysis in grading gliomas.

**Results:** Six studies including 440 patients were included and analyzed. The pooled sensitivity, specificity, PLR, NLR, and DOR with 95% CIs were 0.86 (95% CI 0.81-0.89) and 0.93 (95% CI 0.88-0.96), 12.29 (95% CI 6.95-21.72), 0.16 (95% CI 0.10-0.25), and 88.99 (95% CI 37.92-208.84), respectively. The SROC curve showed an AUC of 0.9718. Deeks testing confirmed no significant publication bias in all studies.

**Conclusions:** This meta-analysis suggested that MRI texture analysis have high accuracy in differentiating high-grade from low-grade gliomas. Standardized methodology is warranted to guide the use of this technique for clinical decision-making.

**Key Words:** Texture analysis; MRI; glioma; meta-analysis.

## Strengths and limitations of this study

1. A meta-analysis assessing the accuracy of MRI texture analysis in differentiating high-grade from low-grade gliomas.

2. The pooled sensitivity and specificity were 0.86 and 0.93 respectively for MRI texture analysis in differentiating high-grade from low-grade gliomas.

3. Standardized methodology is warranted to guide the use of this technique for clinical decision-making.

## 1. Introduction

Gliomas are the most common type of primary malignant brain tumor. According to the World Health Organization (WHO) tumor classification, gliomas are subdivided into grades I-IV, where I-II are low-grade gliomas (LGGs) and III-IV are high-grade gliomas (HGGs). [1] LGG is low malignant tumor associated with a longer life expectancy, while HGG is highly aggressive and have a dismal prognosis despite various therapeutic managements.[2 3, 4] Surgical resection is the preferred treatment for most gliomas. After surgery, HGG normally require 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 tumor is significant for prognosis of patients. Hence, identification of tumor level prior to surgery is of important significance for intraoperative decision-making. The histopathological assessment is the current gold

standard for grading gliomas, which is an invasive procedure and generally performed after surgery. Thus, the potential to accurately ascertain tumor grade by utilizing a non-invasive technique is gaining a lot of attention.[6]

Magnetic resonance imaging (MRI) is the imaging method of first choice for depicting gliomas. With the development of technology, several physiological MRI techniques including MR spectroscopy, diffusion-weighted imaging (DWI), and perfusion-weighted imaging (PWI), have also been applied to grading gliomas.[7 8] Texture analysis (TA) is a method for quantifying the spatial distributions of intensities in images. Some reports have suggested that TA have promise in the field of oncology diagnosis, including quantifying tumor heterogeneity and tumor grading.[9 10] Until now, some reports have been published regarding tumor heterogeneity in glioma using MRI texture analysis.[7 10-14] However, these studies were inconclusive because of insuffcient sample and different diagnostic algorithms. The aim of this meta-analysis was to systematically evaluate the accuracy of TA for discriminating HGGs from LGGs.

## 2. Methods

Patient and public involvement

As this is a meta-analysis, ethical approval was not necessary. Patients' priorities, experience and preferences were not involved in designing the study.

### 2.1. Search strategy

This systematic review and the meta-analysis was performed following the guidelines for the diagnostic studies.[15]

PubMed, Cochrane library, Science Direct and Embase were searched on Aug 25, 2018, and no start date limit was applied. The search key words were "Texture analysis", "glioma", "brain neoplasm" and "brain tumor". 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.

2.2. Study selection criteria

The studies were selected on the basis of the following criteria: (1) Clinical trials assessing the diagnostic accuracy of TA for differentiating HGGs from LGGs; (2) using histopathology as criterion standard; (3) Sufficient information to calculate true positive (TP), false-positive (FP), true-negative (TN), and false negative (FN). Excluded criteria: animal studies， case reports, abstracts, without sufficient calculable data, duplicated reports, or studies based on the same study.

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

2.3. 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 in regards to tumor grading. If the TP, FP, TN and FN were not reported, we calculated backwards using indexes including sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV).

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

2.4. Statistical analysis

Meta-analyses were performed using the software Meta-Disc version 1.4. The pooled sensitivity, specificity, positive and negative likelihood ratios (LRs), and diagnostic odds ratio (DOR) were calculated based on bivariate generalized linear mixed modeling using the extracted data of TP, TN, FP, and FN. The accuracy of the data was determined using a summary receiver operating characteristic plot (SROC) and summarizing that curve by calculating the area under the curve (AUC). In general, a diagnostic tool is regarded failed when AUC values were between 0.5 and 0.6, poor when AUC values were between 0.6 and 0.7, fair when AUC values were between 0.7 and 0.8, good when AUC values were between 0.8 and 0.9, and excellent when AUC values were between 0.9 and 1.[16]

2.5. Subgroup analysis

We calculated the pooled weighted sensitivity and specificity of subgroups to observe the effects caused by substantial heterogeneity of the included studies. Studies were

grouped based on MRI performed at different field strengths (3.0 T vs not 3.0 T), MRI images used (Contrast-enhanced T1 and FLAIR vs DWI) and diagnostic algorithm used ((Gray Level Co-occurrence Matrices (GLCM) vs Laplacian of Gaussian band-pass filtration)).

2.6. Publication bias

The publication bias was assessed using Deeks funnel plot asymmetry test, where P<.05 suggests a potential publication bias. The Deeks funnel plot asymmetry test was performed by Stata 11.0.

## 3. Results

3.1. Literature research

A total of 125 studies were initially identified using the above mentioned search strategy, which were then screened in title and abstract. Of these, 38 articles were further evaluated in full text. According to the inclusion criteria, 6 studies[7 10-14] were retrieved. 29 articles were irrelevant and 3 could not provide enough data to construct the 2x2 table. The study selection process is shown in Figure 1.

3.2. Study characteristics

Ultimately, 6 studies with 440 participants were enrolled in this meta-analysis. The detailed characteristics of included studies were given in Table 1. All studies were retrospective cohort studies. The MR examinations were performed on a 1.5 T scanner in one study, 3.0 T in four studies and one study not mentioned. Contrast-enhanced T1 images were used for analysis in 2 studies, Contrast-enhanced

T1 combined with fluid-attenuated inversion-recovery sequence (FLAIR) images were used for analysis in 2 studies, and diffusion-weighted imaging (DWI) were used for analysis in 2 studies. As for the TA tools, TexRAD software were used in 2studies, and FSL Library of analysis tools, MISSTA, CAD system, and Fire voxel were used in one research respectively.

3.3. Quality of included studies

The quality assessment of included studies are presented in Table 2 using QUADAS checklist. Overall, the study quality was satisfactory.

3.4. Pooled results

The pooled sensitivity and specificity of TA for discriminating HGGs and LGGs were 0.86 (95% CI 0.81-0.89) and 0.93 (95% CI 0.88-0.96), respectively. The forest plots were shown in Figure 2 and 3. The pooled PLR and NLR were 12.29 (95% CI 6.95-21.72) and 0.16 (95% CI 0.10-0.25), respectively. The DOR was 88.99 (95% CI 37.92-208.84). The SROC curve analysis was used to summarize overall diagnostic accuracy. The AUC was 0.9718. The SROC curve was shown in Figure 4. The results demonstrating high diagnostic performance in discrimination of HGGs from LGGs.

3.5. Subgroup analyses

The results of the subgroup analyses were presented in Table 3. The sensitivity was slightly lower but specificity was higher for studies in which MRI performed by 3.0 T. The sensitivity and specificity were significantly higher for studies using contrast-enhanced T1 and FLAIR images than DWI. The diagnostic performance of GLCM was slightly higher than Laplacian of Gaussian band-pass filtration.

3.6 Publication bias

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 were shown in Figure 5.

## 4. Discussion

We assessed the accuracy of MRI TA in differentiating HGGs from LGGs. The meta-analysis showed the pooled sensitivity and specificity of TA were 0.86 and 0.93, respectively. The PLR and NLR were 12.29 and 0.16, respectively. The AUC was 0.97. The results demonstrated that TA had high diagnostic performance in discrimination of HGGs from LGGs.

The histopathology is the gold standard for diagnosis of gliomas, but it is an invasive procedure. To provide more accurate information and avoid unnecessary operations of gliomas, the role of MR cannot be neglected. With the development of techniques, more and more metabolic and physiologic MR imaging, such as DTI, MRS, DWI, DSC MRI and DCE MRI, have been utilized in the assessment of grading gliomas.[17-19] Textures are complex visual patterns composed of entities that have characteristic size, brightness, intensity, et al. Thus, texture can be regarded as a similarity grouping in an image.[14] The earliest reports indicated that TA based on CT images had the potential of differential diagnosis of tumor heterogeneity.[20 21] To date, there have been some reports on glioma grading using a TA of MRI imaging. However, these studies all had insufficient sample and used different diagnostic

algorithms, thus the results were inconclusive. We conduct this meta-analysis to systematically evaluate the accuracy of TA for discriminating HGGs from LGGs.

This research demonstrated that TA was useful for discrimination between HGGs and LGGs. In a published meta-analysis based on MR PWI for glioma grading, the pooled sensitivity, specificity and diagnostic odds ratio were 93%, 81% and 55%, respectively.[22] However, PWI examination requires injection of contrast medium and the results are influenced by many factors, therefore, it is difficult to widely applicate 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.[23] DWI has specific advantages over PWI that it is easy accessible, nonradiative and less expensive. TA can use any kind of MRI sequences such as PWI, DWI, FLAIR, et al, thus, this technique is easy to widespread applicate.

However, obvious heterogeneity between studies needs further consideration. Different field strength (3.0 T and 1.5 T), different MR imaging used (DWI, contrast-enhanced T1 and FLAIR), different analysis tolls (MISSTA, TexRAD, Fire voxel and the FSL Library of analysis tools) and different diagnostic algorithm (GLCM and Laplacian of Gaussian band-pass filtration) could give unexpected substandard results and affect the accuracy of the conclusion. The technique procedure should be standardized by further researches.

It is worth noting that this study also had several limitations. First, this systematic review just included 6 studies with 440 patients. The number of studies were limited and all studies were based on shortage of participants, which might affected the

accuracy of the results. Second, although no publication bias was detected in this meta-analysis, the test strength was limited by limited l number of studies. Thus, the publication bias was also a concern. Finally, different field strengths, imaging sequence and TA tools were used in the included studies lack of consensus, which influenced the consistency of measurements. Therefore, well-conducted investigations using a standardized methodology are needed to confirm the discrimination value of TA on gliomas.

In conclusion, 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 standardized methodology would improve the accuracy glioma diagnosis and clinical decision making in the future.

## Competing interests

The authors declare that they have no competing interests.

## 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).

## DATA SHARING STATEMENT

The corresponding author could provide original data if requested.

**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.

## References

1. Louis DN, Ohgaki H, Wiestler OD, et al. The 2007 WHO classification of tumours of the central nervous system. *Acta Neuropathol* 2007;114(2):97-109. doi: 10.1007/s00401-007-0243-4 [published Online First: 2007/07/10]
2. Johnson DR, O'Neill BP. Glioblastoma survival in the United States before and during the temozolomide era. *J Neurooncol* 2012;107(2):359-64. doi: 10.1007/s11060-011-0749-4 [published Online First: 2011/11/03]
3. 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(3):159-67. doi: 10.1111/j.1600-0404.2010.01350.x [published Online First: 2010/03/20]
4. 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(5):897-901. doi: 10.3171/jns.2005.102.5.0897 [published Online First: 2005/06/02]
5. 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(2):233-7. doi: 10.3171/jns.2006.104.2.233 [published Online First: 2006/03/03]
6. Server A, Kulle B, Gadmar OB, 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(2):462-70. doi: 10.1016/j.ejrad.2010.07.017 [published Online First: 2010/08/17]
7. 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(9):e108335. doi: 10.1371/journal.pone.0108335 [published Online First: 2014/10/01]
8. Jackson A, O'Connor JP, Parker GJ, et al. Imaging tumor vascular heterogeneity and angiogenesis using dynamic contrast-enhanced magnetic resonance imaging. *Clin Cancer Res* 2007;13(12):3449-59. doi: 10.1158/1078-0432.CCR-07-0238 [published Online First: 2007/06/19]
9. Nie K, Chen JH, Yu HJ, et al. Quantitative analysis of lesion morphology and texture features for diagnostic prediction in breast MRI. *Acad Radiol* 2008;15(12):1513-25. doi:

10.1016/j.acra.2008.06.005 [published Online First: 2008/11/13]

10. 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(1):120-30. doi: 10.1016/j.mri.2008.05.017 [published Online First: 2008/07/08]
11. Skogen K, Schulz A, Dormagen JB, et al. Diagnostic performance of texture analysis on MRI in grading cerebral gliomas. *Eur J Radiol* 2016;85(4):824-9. doi: 10.1016/j.ejrad.2016.01.013 [published Online First: 2016/03/15]
12. Li-Chun Hsieh K, Chen CY, Lo CM. Quantitative glioma grading using transformed gray-scale invariant textures of MRI. *Comput Biol Med* 2017;83:102-08. doi: 10.1016/j.compbiomed.2017.02.012 [published Online First: 2017/03/04]
13. 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(5):7297-304. doi: 10.3892/ol.2018.8232 [published Online First: 2018/05/08]
14. Ditmer A, Zhang B, Shujaat T, et al. Diagnostic accuracy of MRI texture analysis for grading gliomas. *J Neurooncol* 2018 doi: 10.1007/s11060-018-2984-4 [published Online First: 2018/08/27]
15. 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. doi: 10.1186/2046-4053-4-1 [published Online First: 2015/01/03]
16. Metz CE. Basic principles of ROC analysis. *Semin Nucl Med* 1978;8(4):283-98. [published Online First: 1978/10/01]
17. Yang D, Korogi Y, Sugahara T, et al. Cerebral gliomas: prospective comparison of multivoxel 2D chemical-shift imaging proton MR spectroscopy, echoplanar perfusion and diffusion-weighted MRI. *Neuroradiology* 2002;44(8):656-66. doi: 10.1007/s00234-002-0816-9 [published Online First: 2002/08/20]
18. Majos C, Julia-Sape 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(10):1696-704. [published Online First: 2004/12/01]
19. 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(3):808-17. doi: 10.1148/radiol.2473070571 [published Online First: 2008/05/20]
20. 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(2):157-64. doi: 10.1016/j.crad.2011.08.012 [published Online First: 2011/09/29]
21. 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(1):177-84. doi: 10.1148/radiol.12120254 [published Online First: 2012/11/16]
22. Min ZG, Liu HJ, Li M, et al. [Accuracy of MR perfusion weighted imaging for cerebral glioma grading: a meta-analysis]. *Zhonghua Yi Xue Za Zhi* 2010;90(41):2927-31. [published Online First: 2011/01/08]
23. 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. doi: 10.1016/j.jns.2016.12.008 [published Online First: 2017/01/31]

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 | Diagnostic algorithm | Reference standard |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Zacharaki EI | 2009 | USA | re | 74 | 52 | 22 | 3.0 T | Contrast-enhanced T1 and FLAIR | FSL Library of analysis tools | NA | Histology |
| Ryu YJ | 2014 | Korea | re | 40 | 32 | 8 | 1.5 T | DWI | MISSTA | GLCM | Histology |
| Skogen K | 2016 | Norway | re | 95 | 68 | 27 | 3.0 T | Contrast-enhanced T1 and FLAIR | TexRAD | Laplacian of Gaussian band-pass filtration | Histology |
| Li-Chun Hsieh K | 2017 | Taiwan | re | 107 | 34 | 73 | NA | Contrast-enhanced T1 | CAD system | GLCM | Histology |
| Ditmer A | 2018 | USA | re | 94 | 80 | 14 | 3.0 T | Contrast-enhanced T1 | TexRAD | Laplacian of Gaussian band-pass filtration | Histology |
| Wang S | 2018 | China | re | 30 | 18 | 12 | 3.0 T | DWI | Fire voxel | GLCM | Histology |

SD: standard deviation; T: Tesla; HGG: high grade gliomas; LGG: low grade gliomas; re: retrospective; pro: prospective; NA: not mentioned. DWI: diffusion-weighted imaging. MISSTA: Medical imaging solution for segmentation and texture analysis; GLCM: Gray Level Co-occurrence Matrices; FLAIR: Fluid-attenuated inversion-recovery sequence.

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

| Study | Risk of bias | | | | Applicability concerns | | |
|---|---|---|---|---|---|---|---|
| | Patient slection | Index test | Reference standard | Flow and timing | Patient selection | Index test | Reference standard |
| Zacharaki EI 2009 | + | - | + | + | + | ? | + |
| Ryu YJ 2014 | + | + | ? | + | + | ? | + |
| Skogen K 2016 | + | + | + | + | + | - | ? |
| Li-Chun Hsieh K 2017 | + | + | - | + | + | + | ? |
| Ditmer A 2018 | + | ? | + | + | + | ? | + |
| Wang S 2018 | + | + | + | + | + | - | + |

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

Table 3 Results of Pooled Estimates of All Studies and of Different Subgroups.

| Studies | N | Sensitivity | Specificity | PLR | NLR | DOR | AUC |
|---|---|---|---|---|---|---|---|
| All studies | 6 | 0.856 | 0.929 | 12.285 | 0.163 | 88.991 | 0.9718 |
| MRI performed at 3.0 T | 4 | 0.849 | 0.933 | 12.458 | 0.169 | 81.446 | 0.9686 |
| MRI performed at not 3.0 T | 2 | 0.879 | 0.925 | 12.144 | 0.095 | 106.63 | - |
| Image used: Contrast-enhanced T1 and FLAIR | 4 | 0.859 | 0.933 | 13.424 | 0.155 | 103.48 | 0.9766 |
| Image used: DWI | 2 | 0.840 | 0.900 | 8.366 | 0.147 | 55.752 | - |
| Diagnostic algorithm: GLCM | 3 | 0.893 | 0.924 | 12.024 | 0.086 | 124.85 | 0.9584 |
| Diagnostic algorithm: Laplacian of Gaussian band-pass filtration | 2 | 0.837 | 0.923 | 11.057 | 0.151 | 70.538 | - |

PLR: positive likelihood ratio; NLR: negative likelihood ratio; DOR: diagnostic odds ratio; AUC: the area under the curve; DWI: diffusion-weighted imaging; FLAIR: Fluid-attenuated inversion-recovery sequence; GLCM: Gray

Level Co-occurrence Matrices.

## Figure legends

Figure1. Results of literature search.

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

Figure 3. Pooled estimates specificity of texture analysis to differentiate high-grade from low-grade gliomas.

Figure 4. Summary receiver operating characteristics (SROC) curve of texture analysis to differentiate high-grade from low-grade gliomas.

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

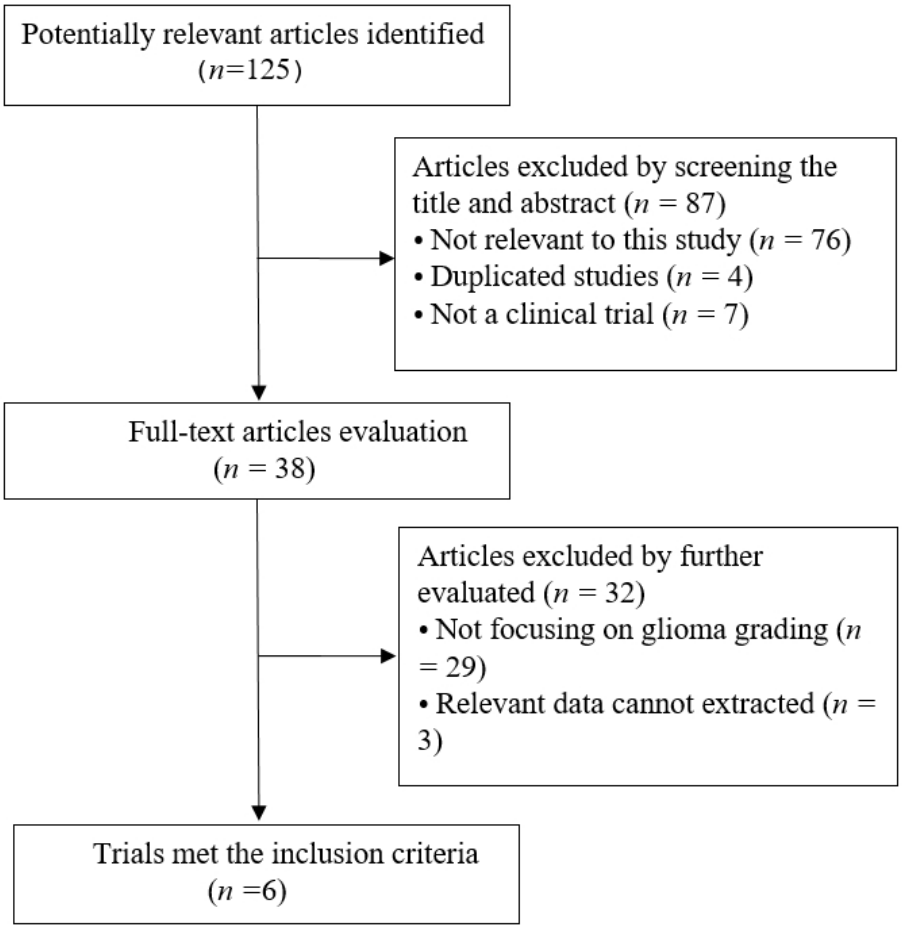


Figure1. Results of literature search.

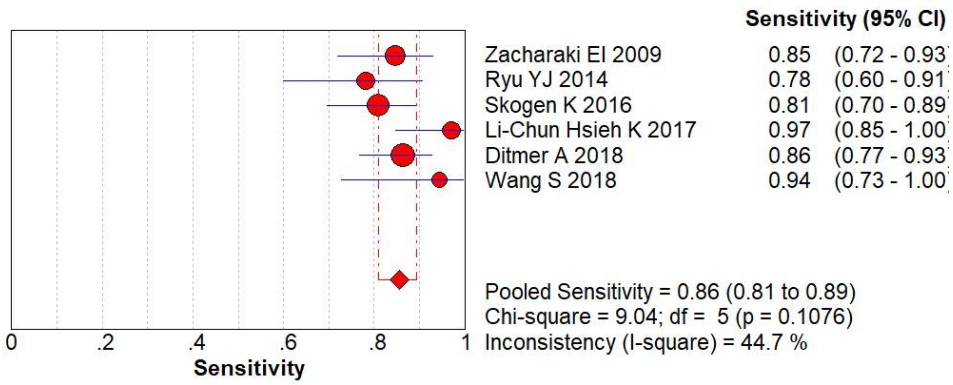


Figure 2. Pooled estimates of sensitivity of texture analysis to differentiate high-grade from low-grade gliomas.

370x265mm (72 x 72 DPI)

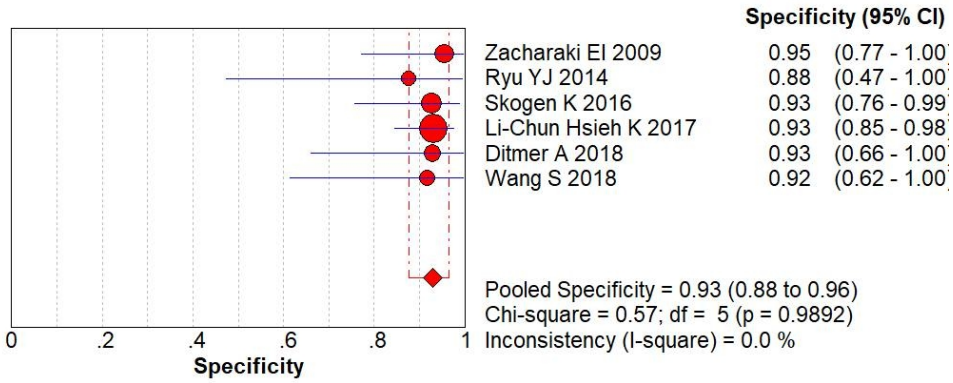


Figure 3. Pooled estimates specificity of texture analysis to differentiate high-grade from low-grade gliomas.

370x265mm (72 x 72 DPI)

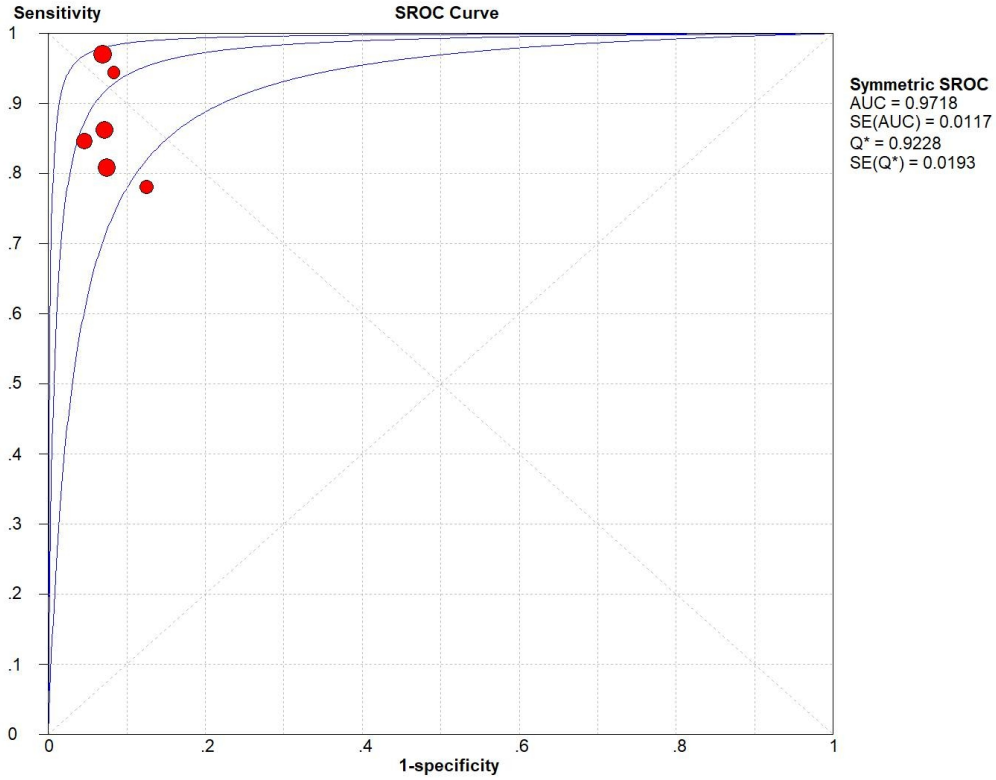


Figure 4. Summary receiver operating characteristics (SROC) curve of texture analysis to differentiate high-grade from low-grade gliomas.

473x376mm (72 x 72 DPI)

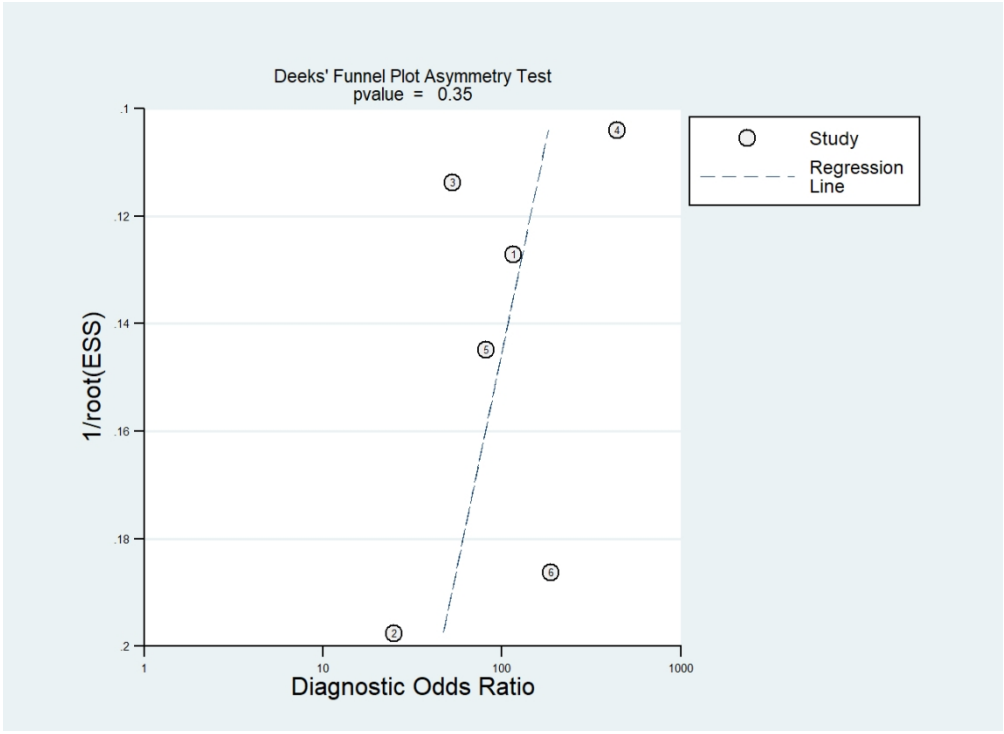


Figure 5. Deeks funnel plots indicating no publication bias (p = 0.35).

452x329mm (72 x 72 DPI)

# PRISMA 2009 Checklist

| Section/topic | # | Checklist item | Reported on page # |
|---|---|---|---|
| **TITLE** | | | 1 |
| Title | 1 | Identify the report as a systematic review, meta-analysis, or both. | 1 |
| **ABSTRACT** | | | 2 |
| Structured summary | 2 | Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number. | 2 |
| **INTRODUCTION** | | | 3-4 |
| Rationale | 3 | Describe the rationale for the review in the context of what is already known. | 3 |
| Objectives | 4 | Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS). | 3 |
| **METHODS** | | | 4-6 |
| Protocol and registration | 5 | Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number. | 4 |
| Eligibility criteria | 6 | Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale. | 4 |
| Information sources | 7 | Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched. | 4 |
| Search | 8 | Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated. | 4 |
| Study selection | 9 | State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis). | 4 |
| Data collection process | 10 | Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators. | 4 |
| Data items | 11 | List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made. | 5 |
| Risk of bias in individual studies | 12 | Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis. | 5 |
| Summary measures | 13 | State the principal summary measures (e.g., risk ratio, difference in means). | 5 |
| Synthesis of results | 14 | Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., $I^2$) for each meta-analysis. | 5 |

# PRISMA 2009 Checklist



| Section/topic | # | Checklist item | Reported on page # |
|---|---|---|---|
| Risk of bias across studies | 15 | Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies). | 6 |
| Additional analyses | 16 | Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified. | 6 |
| **RESULTS** | | | 6-8 |
| Study selection | 17 | Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram. | 6 |
| Study characteristics | 18 | For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations. | 7 |
| Risk of bias within studies | 19 | Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12). | 7 |
| Results of individual studies | 20 | For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot. | 7 |
| Synthesis of results | 21 | Present results of each meta-analysis done, including confidence intervals and measures of consistency. | 7 |
| Risk of bias across studies | 22 | Present results of any assessment of risk of bias across studies (see Item 15). | 8 |
| Additional analysis | 23 | Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]). | 8 |
| **DISCUSSION** | | | 8-108-10 |
| Summary of evidence | 24 | Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers). | 8 |
| Limitations | 25 | Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias). | 10 |
| Conclusions | 26 | Provide a general interpretation of the results in the context of other evidence, and implications for future research. | 10 |
| **FUNDING** | | | 11 |
| Funding | 27 | Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review. | 11 |

*From:* Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(7): e1000097. doi:10.1371/journal.pmed1000097

For more information, visit: **www.prisma-statement.org**.

BMJ Open

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

| | |
|---|---|
| Journal: | *BMJ Open* |
| Manuscript ID | bmjopen-2018-027144.R1 |
| Article Type: | Research |
| Date Submitted by the Author: | 11-Jul-2019 |
| Complete List of Authors: | Qiangping, Wang; Dept. of Neurosurgery, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology<br>Lei, Deqiang; Wuhan Union Hospital, Dept. of Neurosurgery, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology<br>Yuan, Ye; Wuhan Union Hospital, Dept. of Neurosurgery, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology<br>Zhao, Hongyang; Wuhan Union Hospital, Dept. of Neurosurgery, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology; |
| Primary Subject Heading: | Evidence based practice |
| Secondary Subject Heading: | Neurology |
| Keywords: | MRI, glioma, meta-analysis, texture analysis |
| | |

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

Qiang-ping Wang, MD[#]; De-qiang Lei, MD[#]; Ye Yuan, MD; Hong-yang Zhao, MD[*].

# Qiang-ping Wang and De-qiang Lei contributed equally to this work.

Dept. of Neurosurgery, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430022, China

[*]Correspondence to Pro. Hong-yang Zhao, Email: zhaohongyangxiehe@163.com; Tel.: +86 134-7685-0638, Fax: +86 028-8712-1111. Address: Dept. of Neurosurgery, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, 1277 JieFang Avenue, Wuhan, China, 430022.

**Key words:** Texture analysis; MRI; glioma; meta-analysis.

## Abstract

**Objectives:** Texture analysis (TA) is a method used for quantifying the spatial distributions of intensities in images using scanning software. Magnetic resonance imaging (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 accuracy ratio (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% confidence intervals (CIs) were 0.93 (95% CI 0.88-0.96), 0.86 (95% CI 0.81-0.89), 6.4 (95% CI 4.8-8.6), 0.08 (95% CI 0.05-0.15), and 78 (95% CI 39-156), respectively. The SROC curve showed an area under the curve of 0.96 (95% CI 0.93-0.97). 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 standardized methodology is warranted to guide the use of this technique for clinical decision-making.

**Keywords:** texture analysis; MRI; glioma; meta-analysis.

## Strengths and limitations of this study

1. This meta-analysis assesses the accuracy of MRI texture analysis in differentiating low-grade gliomas from high-grade ones.
2. 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.
3. A standardized 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 World Health Organization tumour classification, gliomas are subdivided into grades I–IV, where I-II are low-grade gliomas (LGGs) and III-IV are high-grade gliomas (HGGs). [1] LGG is a low-grade malignant tumour associated with longer life expectancy, while HGG is highly aggressive and has a dismal prognosis despite various therapeutic managements.[2] [3,] [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]

Magnetic resonance imaging (MRI) is the first-choice of imaging method in detecting gliomas. With the development of technology, several physiological MRI techniques including 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 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–II gliomas) were considered positive; HGGs (grade II–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 MetaDisc version 1.4 (Metadisc, Unit of Clinical Biostatistics of Ramón y Cajal Hospital, Madrid, Spain) and Stata version 12.0 (StataCorp LP, College Station, TX, USA). The pooled sensitivity, specificity, positive likelihood ratio (PLR), negative likelihood ratio (NLR) and diagnostic odds ratio (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 < .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 (gray 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 <. 05 suggests a potential publication bias. Deeks funnel plot asymmetry test was performed using Stata version 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 to construct the 2 × 2 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, 6 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 2 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 2 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.

**Pooled results**

The pooled sensitivity and specificity of TA for discriminating LGGs and HGGs were 0.93 (95% confidence interval [CI] 0.88-0.96) and 0.86 (95% CI 0.81-0.89), respectively. The forest plots are shown in Figure 2. The pooled PLR and NLR were 0.86 (95% CI 0.81-0.89) and 6.4 (95% CI 4.8-8.6), respectively. The DOR was 78 (95% CI 39-156). 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.

**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% pretest probability of LGG, the posttest 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 measured using Cochran-Q method and $I^2$. As shown in Figure 2, the P-value of the Cochran-Q method was >.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 computed tomography 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 neglected. With the development of techniques, more and more metabolic and physiologic 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 grayscale 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, CADand 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, standard deviation, 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 researchThe 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 6 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 standardized 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.

## Competing interests

The authors declare that they have no competing interests.

## 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).

## DATA SHARING STATEMENT

The corresponding author could provide original data if requested.

**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.

## References


1. Louis DN, Ohgaki H, Wiestler OD, et al. The 2007 WHO classification of tumours of the central nervous system. *Acta Neuropathol* 2007;114(2):97-109. doi: 10.1007/s00401-007-0243-4 [published Online First: 2007/07/10]
2. Johnson DR, O'Neill BP. Glioblastoma survival in the United States before and during the temozolomide era. *J Neurooncol* 2012;107(2):359-64. doi: 10.1007/s11060-011-0749-4 [published Online First: 2011/11/03]
3. 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(3):159-67. doi: 10.1111/j.1600-0404.2010.01350.x [published Online First: 2010/03/20]
4. 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(5):897-901. doi: 10.3171/jns.2005.102.5.0897 [published Online First: 2005/06/02]
5. 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(2):233-7. doi: 10.3171/jns.2006.104.2.233 [published Online First: 2006/03/03]
6. Server A, Kulle B, Gadmar OB, 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(2):462-70. doi: 10.1016/j.ejrad.2010.07.017 [published Online First: 2010/08/17]
7. Chaddad A, Kucharczyk MJ, Daniel P, et al. Radiomics in Glioblastoma: Current Status and Challenges Facing Clinical Implementation. *Front Oncol* 2019;9:374. doi: 10.3389/fonc.2019.00374
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(9):e108335. doi: 10.1371/journal.pone.0108335 [published Online First: 2014/10/01]
9. Jackson A, O'Connor JP, Parker GJ, et al. Imaging tumor vascular heterogeneity and angiogenesis using dynamic contrast-enhanced magnetic resonance imaging. *Clin Cancer Res* 2007;13(12):3449-59. doi: 10.1158/1078-0432.CCR-07-0238 [published Online First: 2007/06/19]
10. Nie K, Chen JH, Yu HJ, et al. Quantitative analysis of lesion morphology and texture features for diagnostic prediction in breast MRI. *Acad Radiol* 2008;15(12):1513-25. doi: 10.1016/j.acra.2008.06.005 [published Online First: 2008/11/13]
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(1):120-30. doi: 10.1016/j.mri.2008.05.017 [published Online First: 2008/07/08]
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(4):824-9. doi: 10.1016/j.ejrad.2016.01.013 [published Online First: 2016/03/15]
13. Li-Chun Hsieh K, Chen CY, Lo CM. Quantitative glioma grading using transformed gray-scale invariant textures of MRI. *Comput Biol Med* 2017;83:102-08. doi: 10.1016/j.compbiomed.2017.02.012 [published Online First: 2017/03/04]
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(5):7297-304. doi: 10.3892/ol.2018.8232 [published Online First: 2018/05/08]
15. Ditmer A, Zhang B, Shujaat T, et al. Diagnostic accuracy of MRI texture analysis for grading gliomas. *J Neurooncol* 2018 doi: 10.1007/s11060-018-2984-4 [published Online First: 2018/08/27]
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. doi: 10.1186/2046-4053-4-1 [published Online First: 2015/01/03]
17. Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011;155(8):529-36. doi: 10.7326/0003-4819-155-8-201110180-00009 [published Online First: 2011/10/19]
18. Metz CE. Basic principles of ROC analysis. *Semin Nucl Med* 1978;8(4):283-98. [published Online First: 1978/10/01]
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(2):157-64. doi: 10.1016/j.crad.2011.08.012 [published Online First: 2011/09/29]
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(1):177-84. doi: 10.1148/radiol.12120254 [published Online First: 2012/11/16]
21. Soni N, Priya S, Bathla G. Texture Analysis in Cerebral Gliomas: A Review of the Literature. *AJNR Am J Neuroradiol* 2019;40(6):928-34. doi: 10.3174/ajnr.A6075
22. Yang D, Korogi Y, Sugahara T, et al. Cerebral gliomas: prospective comparison of multivoxel 2D chemical-shift imaging proton MR spectroscopy, echoplanar perfusion and

diffusion-weighted MRI. *Neuroradiology* 2002;44(8):656-66. doi: 10.1007/s00234-002-0816-9 [published Online First: 2002/08/20]

23. Majos C, Julia-Sape 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(10):1696-704. [published Online First: 2004/12/01]
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(3):808-17. doi: 10.1148/radiol.2473070571 [published Online First: 2008/05/20]
25. Ganeshan B, Goh V, Mandeville HC, et al. Non-small cell lung cancer: histopathologic correlates for texture parameters at CT. *Radiology* 2013;266(1):326-36. doi: 10.1148/radiol.12112428
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. doi: 10.1016/j.ejrad.2019.02.014
27. Min ZG, Liu HJ, Li M, et al. [Accuracy of MR perfusion weighted imaging for cerebral glioma grading: a meta-analysis]. *Zhonghua Yi Xue Za Zhi* 2010;90(41):2927-31. [published Online First: 2011/01/08]
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. doi: 10.1016/j.jns.2016.12.008 [published Online First: 2017/01/31]

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 | Filtration methods | Reference standard |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Zacharaki EI | 2009 | USA | re | 74 | 52 | 22 | 3.0 T | Contrast-enhanced T1 and FLAIR | FSL of analysis tools | NA | Histology |
| Ryu YJ | 2014 | Korea | re | 40 | 32 | 8 | 1.5 T | DWI | MISSTA | GLCM | Histology |
| Skogen K | 2016 | Norway | re | 95 | 68 | 27 | 3.0 T | Contrast-enhanced T1 and FLAIR | TexRAD | Laplacian of Gaussian band-pass | Histology |
| Li-Chun Hsieh K | 2017 | Taiwan | re | 107 | 34 | 73 | NA | Contrast-enhanced T1 | CAD system | GLCM | Histology |
| Ditmer A | 2018 | USA | re | 94 | 80 | 14 | 3.0 T | Contrast-enhanced T1 | TexRAD | Laplacian of Gaussian band-pass | Histology |
| Wang S | 2018 | China | re | 30 | 18 | 12 | 3.0 T | DWI | FireVoxel | GLCM | Histology |

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

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

| Study | Risk of bias | | | | Applicability concerns | | |
|---|---|---|---|---|---|---|---|
| | Patient slection | Index test | Reference standard | Flow and timing | Patient selection | Index test | Reference standard |
| Zacharaki EI 2009 | + | - | + | + | + | ? | + |
| Ryu YJ 2014 | + | + | ? | + | + | ? | + |
| Skogen K 2016 | + | + | + | + | + | - | ? |
| Li-Chun Hsieh K 2017 | + | + | - | + | + | + | ? |
| Ditmer A 2018 | + | ? | + | + | + | ? | + |
| Wang S 2018 | + | + | + | + | + | - | + |

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

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 |

PLR: positive likelihood ratio; NLR: negative likelihood ratio; DOR: diagnostic odds ratio; AUC: the area under the curve; DWI: diffusion-weighted imaging; FLAIR: Fluid-attenuated inversion-recovery sequence; GLCM: Gray

Level Co-occurrence Matrices.

## Figure legends

Figure 1. Results of literature search.

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

Figure 3. Summary receiver operating characteristics (SROC) curve of texture analysis to differentiate low-grade gliomas from high-grade ones.

Figure 4. Fagan nomogram for the elucidation of post-test probabilities with a pretest probability of 25%.

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

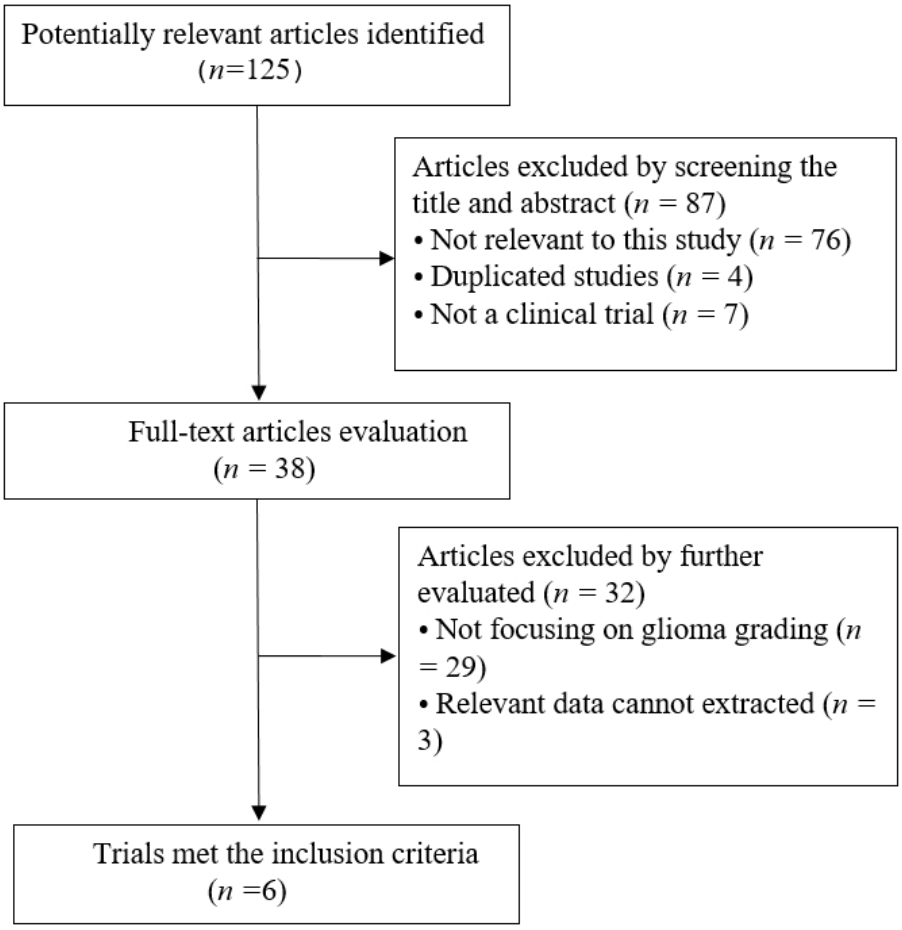


Figure 1. Results of literature search.

195x188mm (96 x 96 DPI)

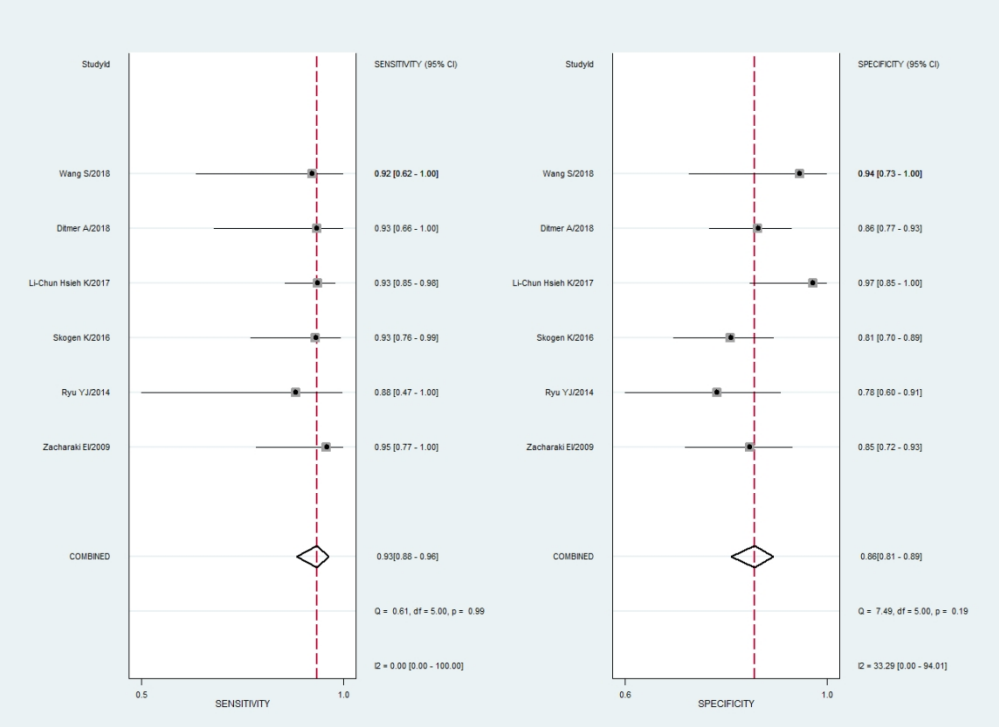


Figure 2. Pooled estimates of sensitivity and specificity of texture analysis to differentiate low-grade gliomas from high-grade ones.

452x329mm (72 x 72 DPI)

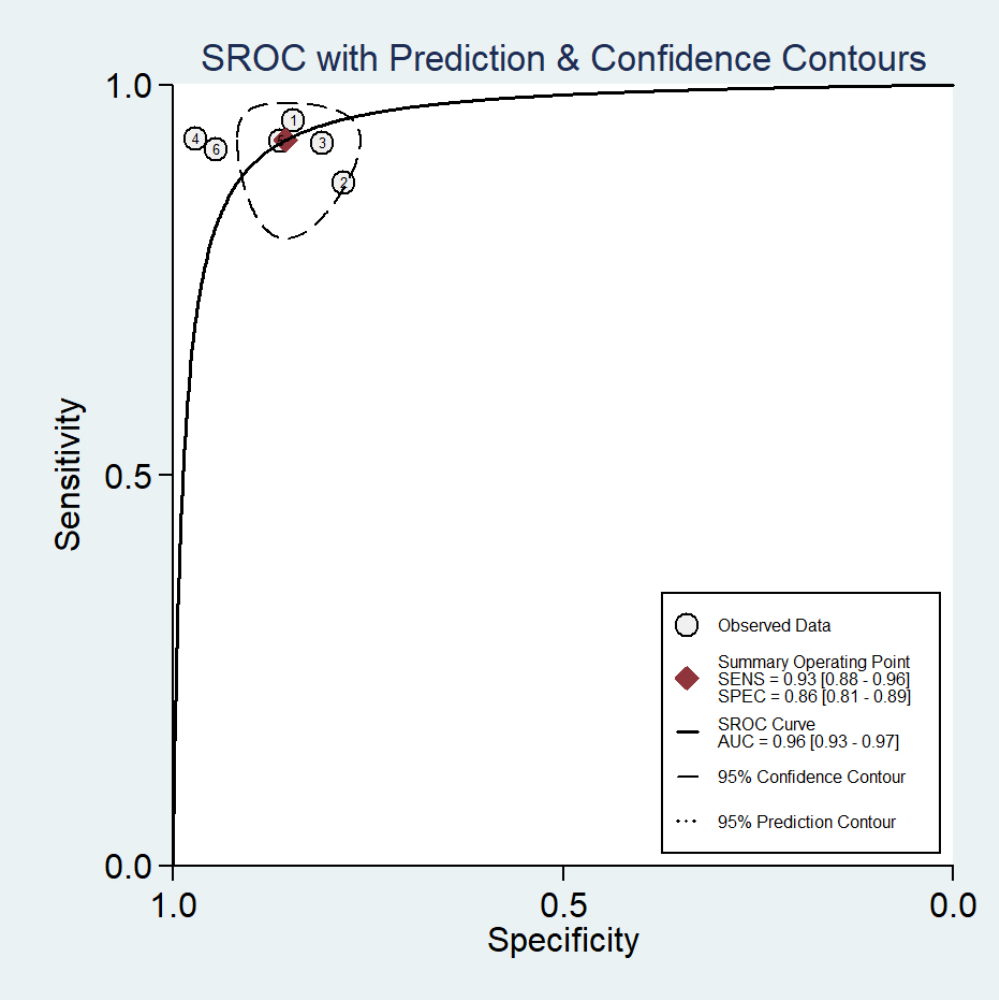


Figure 3. Summary receiver operating characteristics (SROC) curve of texture analysis to differentiate low-grade gliomas from high-grade ones.

328x328mm (72 x 72 DPI)

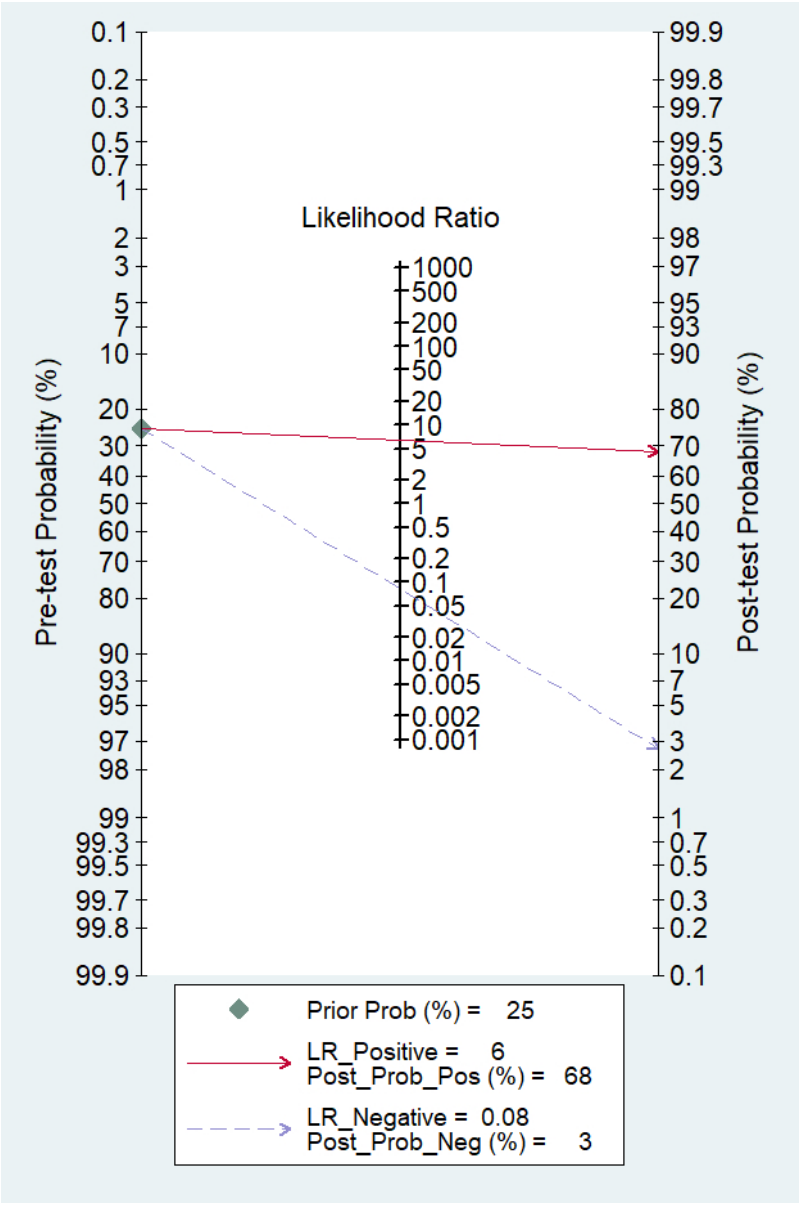


Figure 4. Fagan nomogram for the elucidation of post-test probabilities with a pretest probability of 25%.

219x329mm (72 x 72 DPI)

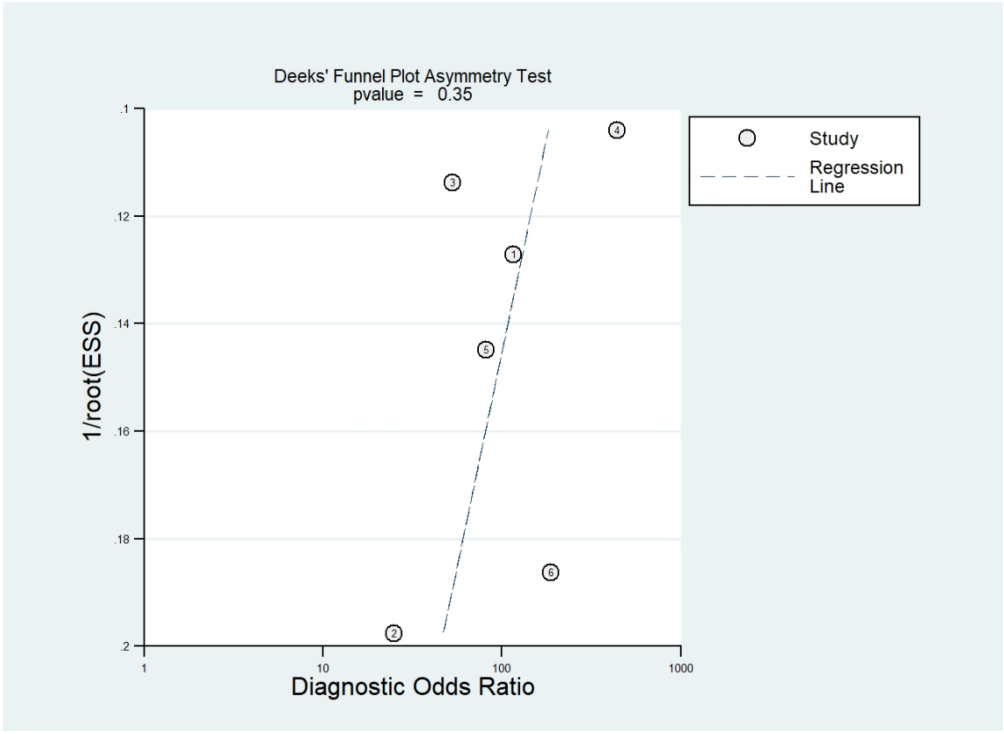


Figure 5 Deeks funnel plots indicating no publication bias (P = 0.35).

452x329mm (72 x 72 DPI)

Search Strategy：

Cochrane Library search strategy:
#1 glioma* or brain neoplasm* or brain tumor*
#2 texture analysis or TA
#3 diagnostic accuracy OR sensitivity OR specificity OR AUC
#1 and #2 and #3

# PRISMA 2009 Checklist

| Section/topic | # | Checklist item | Reported on page # |
|---|---|---|---|
| **TITLE** | | | 1 |
| Title | 1 | Identify the report as a systematic review, meta-analysis, or both. | 1 |
| **ABSTRACT** | | | 2 |
| Structured summary | 2 | Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number. | 2 |
| **INTRODUCTION** | | | 3-4 |
| Rationale | 3 | Describe the rationale for the review in the context of what is already known. | 3 |
| Objectives | 4 | Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS). | 3 |
| **METHODS** | | | 4-6 |
| Protocol and registration | 5 | Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number. | 4 |
| Eligibility criteria | 6 | Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale. | 4 |
| Information sources | 7 | Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched. | 4 |
| Search | 8 | Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated. | 4 |
| Study selection | 9 | State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis). | 4 |
| Data collection process | 10 | Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators. | 4 |
| Data items | 11 | List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made. | 5 |
| Risk of bias in individual studies | 12 | Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis. | 5 |
| Summary measures | 13 | State the principal summary measures (e.g., risk ratio, difference in means). | 5 |
| Synthesis of results | 14 | Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., $I^2$) for each meta-analysis. | 5 |

# PRISMA 2009 Checklist



| Section/topic | # | Checklist item | Reported on page # |
|---|---|---|---|
| Risk of bias across studies | 15 | Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies). | 6 |
| Additional analyses | 16 | Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified. | 6 |
| **RESULTS** | | | 6-8 |
| Study selection | 17 | Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram. | 6 |
| Study characteristics | 18 | For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations. | 7 |
| Risk of bias within studies | 19 | Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12). | 7 |
| Results of individual studies | 20 | For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot. | 7 |
| Synthesis of results | 21 | Present results of each meta-analysis done, including confidence intervals and measures of consistency. | 7 |
| Risk of bias across studies | 22 | Present results of any assessment of risk of bias across studies (see Item 15). | 8 |
| Additional analysis | 23 | Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]). | 8 |
| **DISCUSSION** | | | 8-108-10 |
| Summary of evidence | 24 | Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers). | 8 |
| Limitations | 25 | Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias). | 10 |
| Conclusions | 26 | Provide a general interpretation of the results in the context of other evidence, and implications for future research. | 10 |
| **FUNDING** | | | 11 |
| Funding | 27 | Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review. | 11 |

*From:* Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(7): e1000097. doi:10.1371/journal.pmed1000097

For more information, visit: www.prisma-statement.org.