Review

Diagnostic Performance of Radiomics-Based Machine Learning in Differentiating Glioblastoma From Brain Metastasis: Systematic Review and Meta-Analysis

Chen Hong, MD; Gemingtian Liu, MD; Zhixu Bie, MD; Yuanhang Guo, MD; Yi Du, MBBS; Yanan Hu, MBBS; Shiwei Li, PhD; Pinan Liu, PhD

Department of Neurosurgery, Beijing Tiantan Hospital, Capital Medical University, Beijing, China

Corresponding Author:
Pinan Liu, PhD
Department of Neurosurgery
Beijing Tiantan Hospital, Capital Medical University
119 South Fourth Ring West Road, Fengtai District
Beijing, 100070
China
Phone: 86 13716372500
Email: pinanliu@ccmu.edu.cn

Abstract

Background: Glioblastoma (GBM) and brain metastasis (BM) are the most common malignant intracranial tumors in adults. Noninvasive differentiation between them remains challenging, as they frequently share similar imaging features. Although radiomics-based machine learning (ML) models show promise, systematic evidence of their performance remains limited.

Objective: We aimed to systematically estimate the diagnostic performance of radiomics-based ML models in differentiating GBM from BM and to explore multiclass classification involving primary central nervous system lymphoma (PCNSL), thereby assessing their potential for clinical use.

Methods: We systematically searched PubMed, Embase, Web of Science, and Cochrane Library databases from inception to January 22, 2026, for studies reporting the diagnostic performance of radiomics-based ML models for differentiating GBM from BM. Multiclass studies involving PCNSL were also retained. Methodological quality was appraised via the Radiomics Quality Score 2.0. Binary diagnostic accuracy was pooled using a bivariate random-effects model, stratified by internal and independent external validation. Multiclass sensitivities and specificities were pooled using separate joint generalized linear mixed models. Physician assessments were evaluated as an exploratory benchmark.

Results: A total of 39 studies (31 binary and 8 multiclass), involving 11,558 patients, were included. For binary classification, pooled sensitivity and specificity were 88.9% (95% CI 84.1%-92.4%) and 83.0% (95% CI 76.8%-87.9%) in internal validation, and 83.8% (95% CI 77.4%-88.7%) and 80.7% (95% CI 71.1%-87.7%) in external validation, respectively. In multiclass analyses, pooled sensitivity was 84.4% (95% CI 72.4%-91.8%) for GBM, 70.6% (95% CI 54.4%-82.8%) for BM, and 75.0% (95% CI 62.7%-84.3%) for PCNSL, with corresponding specificities of 82.5% (95% CI 70.1%-90.4%), 88.8% (95% CI 81.9%-93.3%), and 94.4% (95% CI 88.4%-97.4%), respectively.

Conclusions: Radiomics-based ML models could distinguish between GBM and BM in the included cohorts. However, due to methodological limitations and scarce evidence, their generalizability and clinical utility remain unclear. Further multicenter prospective studies should conduct independent external validation in consecutive patients.

Trial Registration: PROSPERO CRD420261295269; https://www.crd.york.ac.uk/PROSPERO/view/CRD420261295269

(J Med Internet Res 2026;28:e104060) doi: 10.2196/104060

KEYWORDS

glioblastoma; brain metastasis; machine learning; radiomics; differential diagnosis

XSL•FO
RenderX

Introduction

Glioblastoma (GBM) and brain metastasis (BM) are the most common malignant intracranial neoplasms in adults. It is crucial to distinguish between these 2 entities, as they differ significantly in terms of their natural course, therapeutic strategies, and prognosis [1,2]. GBM has a median overall survival (OS) of 16 to 18 months and a 5-year OS rate of approximately 7%. The standard treatment consists of chemoradiation following maximal safe resection [1,3]. BM affects an estimated 20% of patients with cancer [4]. The management of BM depends heavily on the primary malignancy and typically involves surgery, radiosurgery, or systemic therapy [5]. Although histopathological examination remains the reference standard for definitive diagnosis, tissue sampling is invasive; the reported rate of perioperative complications ranges from 6% to 12%, and the mortality rate ranges from 0% to 1.7% after stereotactic needle biopsy [6]. Consequently, neurosurgeons rely heavily on contrast-enhanced magnetic resonance imaging (MRI) for preoperative diagnosis. GBM and BM frequently share imaging features, including ring enhancement, marked peritumoral edema, and central necrosis. Furthermore, neurosurgeons must distinguish these tumors from other intracranial neoplasms, such as primary central nervous system lymphoma (PCNSL) [7], which has a 10-year survival rate of approximately 30% and requires different therapeutic strategies [1]. These overlapping features complicate preoperative differentiation, even for experienced neuroradiologists.

Advanced MRI techniques, including perfusion-weighted imaging, diffusion-weighted imaging, and magnetic resonance spectroscopy, have been investigated to distinguish these tumors based on vascularity, microstructure, and metabolism [8-10]. Despite considerable research attention, their diagnostic performance remains variable. Their value as stand-alone tools for routinely distinguishing between GBM and BM remains uncertain.

Radiomics-based machine learning (ML) models have shown considerable promise for tumor classification across diverse imaging modalities [11-13]. Deep learning (DL), a branch of ML, can extract high-dimensional quantitative features and capture spatial heterogeneity, reducing reliance on handcrafted feature design. ML approaches have also been applied to the 3-class differentiation of GBM, BM, and PCNSL [14].

However, heterogeneity in imaging inputs, model types, and feature extraction approaches limits comparability and clinical translation. Chen et al [15] evaluated 30 combinations of feature-selection methods and classifiers using conventional MRI, illustrating the potential of MRI-based radiomics for distinguishing between GBM and BM and the variability in performance across modeling approaches. Building on such evidence and previous reviews [16,17], this meta-analysis updated the evidence through January 2026 and estimated the pooled diagnostic performance of radiomics-based ML models for differentiating GBM from BM, separating internal from external validation. We further performed subgroup analyses by model type, MRI input, and feature extraction approach and assessed methodological quality using the Radiomics Quality Score (RQS) 2.0. We also considered multiclass classification (GBM, BM, and PCNSL) and summarized physician assessments as an exploratory benchmark.

Methods

Study Registration

This meta-analysis followed the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 statement and the PRISMA Diagnostic Test Accuracy guidelines [18,19]. The study protocol was registered in PROSPERO (CRD420261295269). Validation stratification was prespecified; its operational definitions were clarified during revision. The completed PRISMA 2020 checklist is provided separately (Multimedia Appendix 1).

Ethical Considerations

Ethics approval was not needed because this review used published data and involved no new participant recruitment or individual patient data.

Eligibility Criteria

This meta-analysis used the PIRD (population, index test, reference standard, and diagnosis of interest) framework to define its research question. The population comprised patients undergoing preoperative evaluation for GBM or BM. The index test was the diagnostic performance of radiomics-based ML and/or DL models. The reference standard was histopathology or, for BM, a clinical and imaging reference standard. The diagnosis of interest was the differentiation of GBM from BM.

Fully published English-language studies that met the following criteria were included: (1) a radiomics-based ML and/or DL model was developed to differentiate GBM from BM or distinguish among GBM, BM, and PCNSL and (2) diagnoses were established by histopathology or, for BM, by a clinical and imaging diagnosis in patients with known systemic malignancy.

Exclusion criteria included the following: (1) conference abstracts, (2) only image segmentation or texture analysis without constructing an ML and/or DL model for differentiating these entities, (3) binary classification studies without sufficient data to report or reconstruct patient-level 2×2 tables, and (4) multiclass studies reporting only overall classification accuracy.

Data Sources and Search Strategy

Web of Science, PubMed, Embase, and Cochrane Library databases were comprehensively searched from inception to January 22, 2026, for eligible studies evaluating the diagnostic performance of radiomics-based models for differentiating relevant brain tumors. The search strategy combined 2 main domains: (1) radiomics, ML, and DL (eg, “radiomics,” “machine learning,” “deep learning,” “support vector machine,” and “convolutional neural network”) and (2) targeted diseases (eg, “glioblastoma,” “brain metastasis,” and “cerebral metastasis”). The search was confined to human studies and full-text articles with peer review. The full search strategy for each database is detailed in Table S1 in Multimedia Appendix 2.

XSL•FO
RenderX

Study Selection

ZB and YG independently screened titles, abstracts, and full texts according to the eligibility criteria, using EndNote 21 (Clarivate) for reference management and screening. They also screened the reference lists of relevant articles and reviews to identify additional studies. Interrater reliability for study selection was evaluated using the Cohen κ coefficient (κ=0.68), indicating substantial agreement. Disagreements were resolved through discussion. If consensus could not be reached, PL made the final decision.

Data Extraction

CH, ZB, and YG independently undertook data extraction from eligible studies using a standardized form, including first author, year of publication, country, task category, patient source, metastasis background, segmentation method and software, optimal model types, imaging modalities, validation method, feature extraction approach, comparison with human experts, and diagnostic performance metrics (Table S2 in Multimedia Appendix 2) [15,20-57]. Disagreements were resolved through discussion; unresolved cases were adjudicated by PL.

Assessment of Study Quality

CH, YH, and YD independently assessed methodological quality using the RQS 2.0 [58]. Handcrafted radiomics (HCR) and DL radiomics (DLR) were assessed separately. Radiomics readiness levels (RRLs) were grouped into model development (RRLs 1-5) and clinical translation (RRLs 6-9). Disagreements were resolved through discussion; unresolved cases were adjudicated by PL. Summary scores, item-level ratings, and scoring calculations are provided in Table S3 in Multimedia Appendix 2 [15,20-57].

Synthesis Methods

Binary sensitivity and specificity were pooled using bivariate random-effects models fitted by maximum likelihood, with GBM as the positive class [59]. The same model was used for subgroup analyses and exploratory physician analyses. Summary receiver operating characteristic curves and the area under the curve (AUC) were derived from the fitted models. Approximate AUCs (95% CIs) were calculated using the logit-scale delta method. Heterogeneity was assessed using approximate model-based I^2. The Deeks test and Fagan nomograms were used to assess small-study effects and posttest probabilities [60,61].

Internal and external validation were analyzed separately, retaining 1 optimal model per study within each validation category. Internal validation included resampling and holdout or temporal testing, while external validation used independent institutions or public databases. For subgroup analyses by model type, MRI input, and feature extraction approach, 1 validation cohort was selected per study, separately for HCR and DLR, prioritizing external validation, internal holdout testing, then cross-validation. Within internal validation, bivariate metaregression was used to compare holdout testing with cross-validation.

For multiclass tasks, class-specific sensitivity and specificity were pooled using separate multivariate generalized linear mixed models with a logit link [62]. *P* values were Holm adjusted separately for sensitivity and specificity.

Statistical analyses and postestimation were performed using Stata (version 19.0; StataCorp; *midas*, *gsem*, and *nlcom*), with Python (version 3.14; Python Software Foundation) used for plotting.

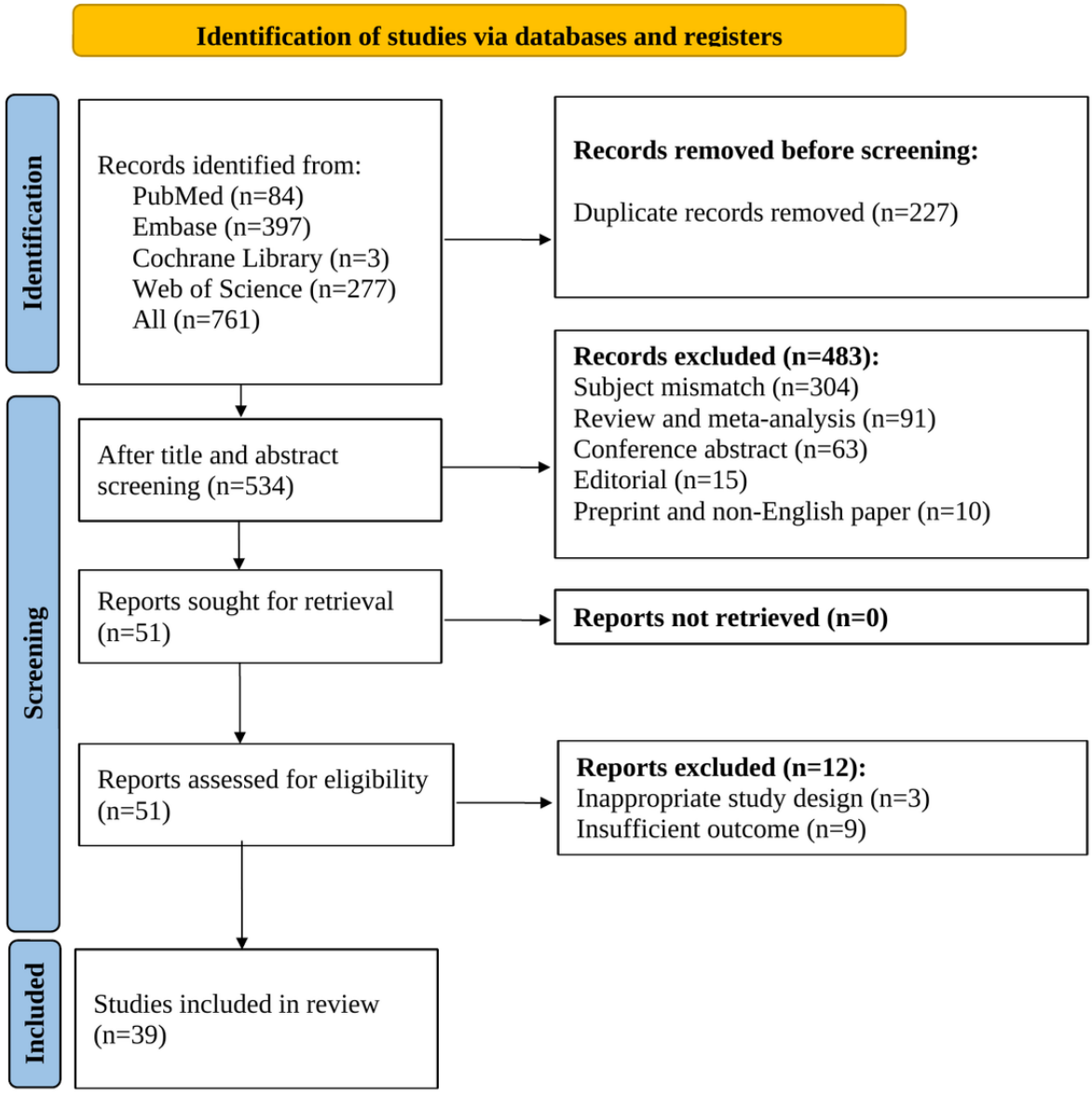
Results

Study Selection

Our search identified 761 records from 4 databases: Embase (n=397, 52.2%), Web of Science (n=277, 36.4%), PubMed (n=84, 11%), and Cochrane Library (n=3, 0.4%). After 227 duplicate records were eliminated, 534 records underwent title and abstract screening, of which 483 articles were excluded due to subject mismatch (n=304), reviews or meta-analyses (n=91), conference abstracts (n=63), editorials (n=15), and preprints or non-English papers (n=10). Full texts were retrieved for the remaining 51 documents. Following rigorous full-text assessment, we excluded 12 articles for inappropriate study designs (n=3) and insufficient outcome data (n=9). A total of 39 studies were included [15,20-57] (Figure 1; Table S2 in Multimedia Appendix 2). Full exclusion reasons are provided in Table S4 in Multimedia Appendix 2 [63-74].

XSL•FO
RenderX

Figure 1. Study selection flow diagram.



Study Characteristics

General Features and Patient Source

The 39 included studies (31 binary [GBM vs BM] and 8 multiclass [GBM, BM, and PCNSL]) were published between 2019 and 2025 and conducted in China (n=19), South Korea (n=5), the United States (n=5), Spain (n=2), Italy (n=2), and 6 other countries (n=1 each). Studies used single-center (n=22) or multicenter (n=17) datasets; 5 also incorporated public databases, such as the Cancer Imaging Archive (Table S2 in Multimedia Appendix 2) [15,20-57].

Imaging Inputs, Model Types, and Feature Extraction Approaches

At the study level, imaging inputs were classified into conventional MRI (n=26; T1-weighted imaging, T2-weighted imaging, fluid-attenuated inversion recovery, and contrast-enhanced T1-weighted imaging) and advanced MRI (n=11; including diffusion, perfusion, susceptibility-sensitive imaging, and quantitative physiological maps) [15,20-22,24-40,42-57]. Dual-energy computed tomography and MRI combined with positron emission tomography were each used in 1 study [23,41]. MRI-only models using any advanced input were classified as advanced MRI (Table S2 in Multimedia Appendix 2) [15,20-57].

In total, 25 studies evaluated HCR alone, 7 evaluated DLR alone, and 7 evaluated both approaches. Model types were grouped into neural networks, tree-based and ensemble models, and other classical classifiers (Table S2 in Multimedia Appendix 2) [15,20-57].

Comparison With Human Experts

A total of 8 binary classification studies reported unassisted physician diagnostic performance [21,26,34,42,43,49,54,56]. Assessments were grouped by experience (>10 or ≤10 years). One study reported assessments by a multidisciplinary team of

XSL•FO
RenderX

3 physicians, all with >10 years of experience; these assessments were included in the >10-year group [42].

Assessment of Study Quality

RQS 2.0 assessment comprised 46 evaluations from 39 studies, including 7 studies that developed separate HCR and DLR models: 32 HCR and 14 DLR evaluations. Mean total scores were 16.8 (SD 4) of a maximum of 56 points (30.1% of the maximum RQS 2.0 score) for HCR and 16.4 (SD 4) of a maximum of 52 points (31.6% of the maximum RQS 2.0 score) for DLR. Mean scores were 14.7 points (45.0% of maximum) for model development (RRLs 1-5) and 2 points (9.0% of maximum) for clinical translation (RRLs 6-9; Table S3 in Multimedia Appendix 2) [15,20-57].

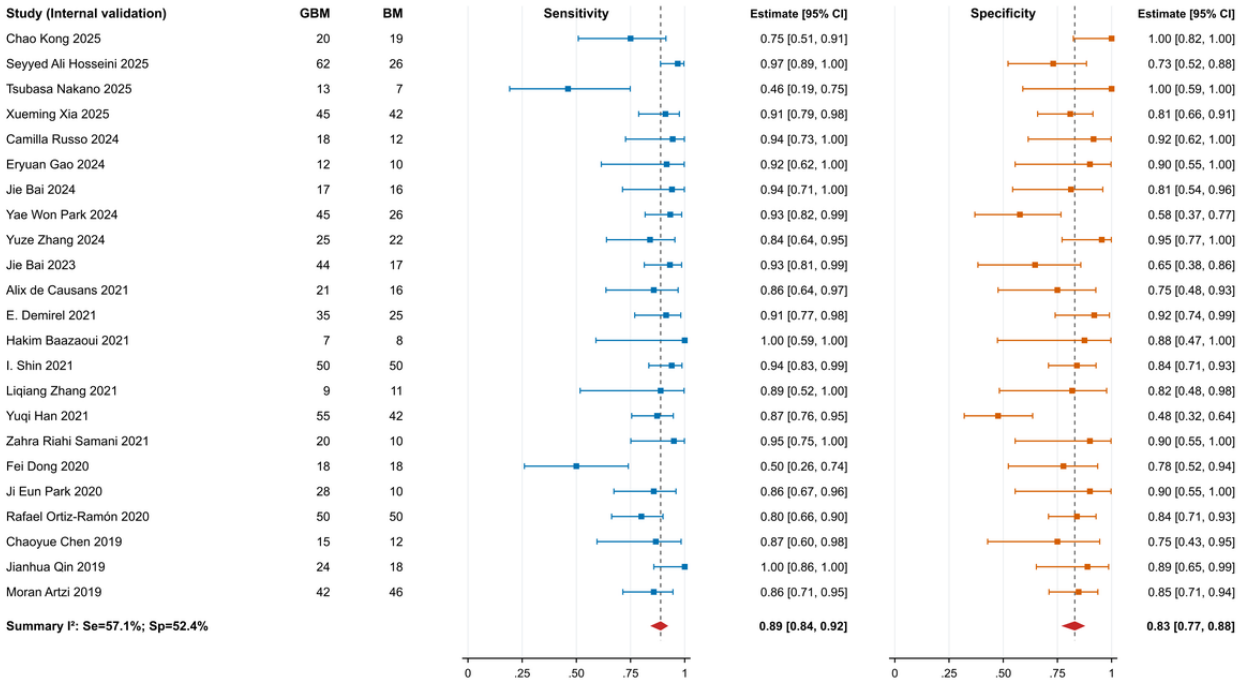
Meta-Analysis of ML for Binary Classification

Overall Diagnostic Performance

To examine generalizability, we stratified the analysis by validation setting (internal vs independent external validation).

In internal validation cohorts (23 studies; 1188 patients), pooled sensitivity and specificity were 88.9% (95% CI 84.1%-92.4%) and 83.0% (95% CI 76.8%-87.9%), respectively (Figure 2) [15,20,23-25,27-29,32,35,39,41,43,44,47-53,55,57]. Pooled positive and negative likelihood ratios were 5.24 (95% CI 3.80-7.22) and 0.13 (95% CI 0.09-0.19), respectively. The diagnostic odds ratio was 39.17 (95% CI 23.35-65.71). The Deeks funnel plot asymmetry test did not detect small-study effects (*P*=.36; Figure S1 in Multimedia Appendix 2). Assuming a pretest probability of 33%, the Fagan nomogram yielded posttest GBM probabilities of 72.1% after a positive result and 6.2% after a negative result (Figure S2 in Multimedia Appendix 2).

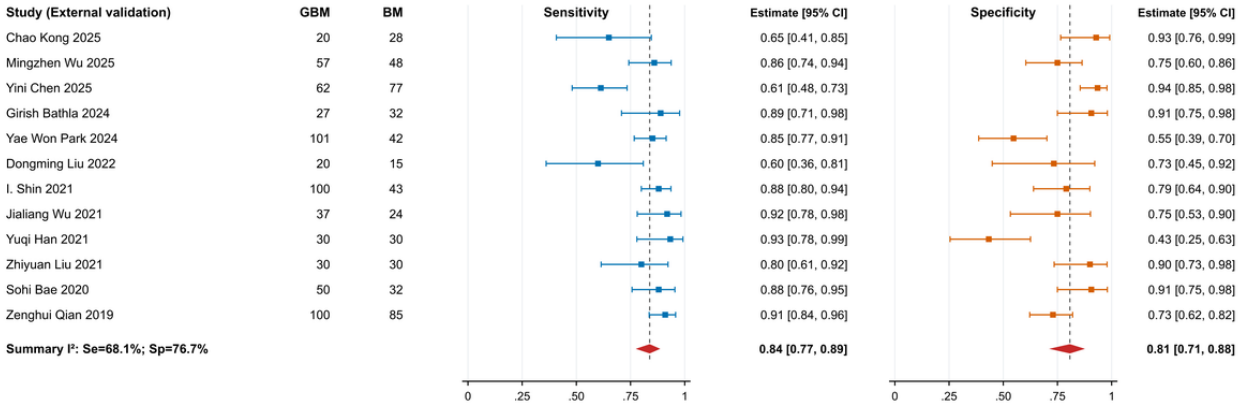
Figure 2. Forest plots of sensitivity and specificity in internal validation cohorts [15,20,23-25,27-29,32,35,39,41,43,44,47-53,55,57]. BM: brain metastasis; GBM: glioblastoma.



In external validation cohorts (12 studies; 1120 patients), pooled sensitivity and specificity were 83.8% (95% CI 77.4%-88.7%) and 80.7% (95% CI 71.1%-87.7%), respectively (Figure 3) [21,24,26,29,34,40,42,43,46,47,54,56]. Positive likelihood ratio, negative likelihood ratio, and diagnostic odds ratio were 4.35 (95% CI 2.93-6.45), 0.20 (95% CI 0.15-0.27), and 21.71 (95% CI 13.93-33.83), respectively. The Deeks funnel plot asymmetry test did not detect small-study effects (*P*=.96; Figure S3 in Multimedia Appendix 2). Assuming a pretest probability of 33%, the Fagan nomogram yielded posttest GBM probabilities of 68.2% after a positive result and 9.0% after a negative result (Figure S4 in Multimedia Appendix 2). Pooled sensitivity and specificity were lower than those in internal validation.

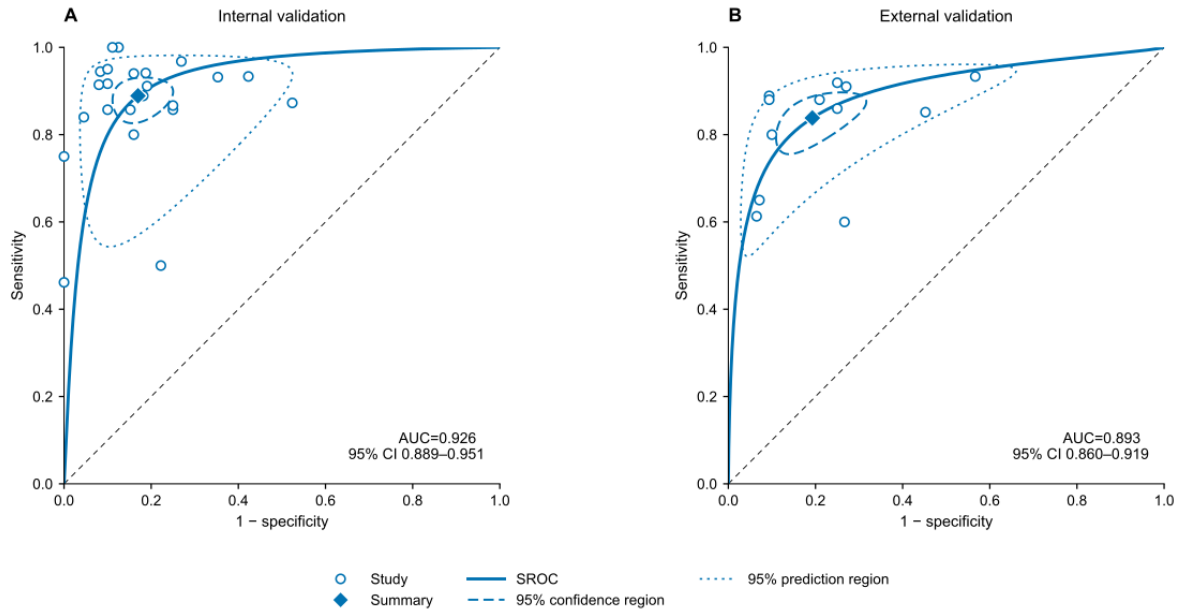
XSL•FO
RenderX

Figure 3. Forest plots of sensitivity and specificity in external validation cohorts [21,24,26,29,34,40,42,43,46,47,54,56]. BM: brain metastasis; GBM: glioblastoma.



The AUCs were 0.926 (95% CI 0.889-0.951) for internal validation (Figure 4A) and 0.893 (95% CI 0.860-0.919) for external validation (Figure 4B).

Figure 4. Summary receiver operating characteristic (SROC) curves by validation setting: (A) internal validation and (B) external validation. AUC: area under the curve.



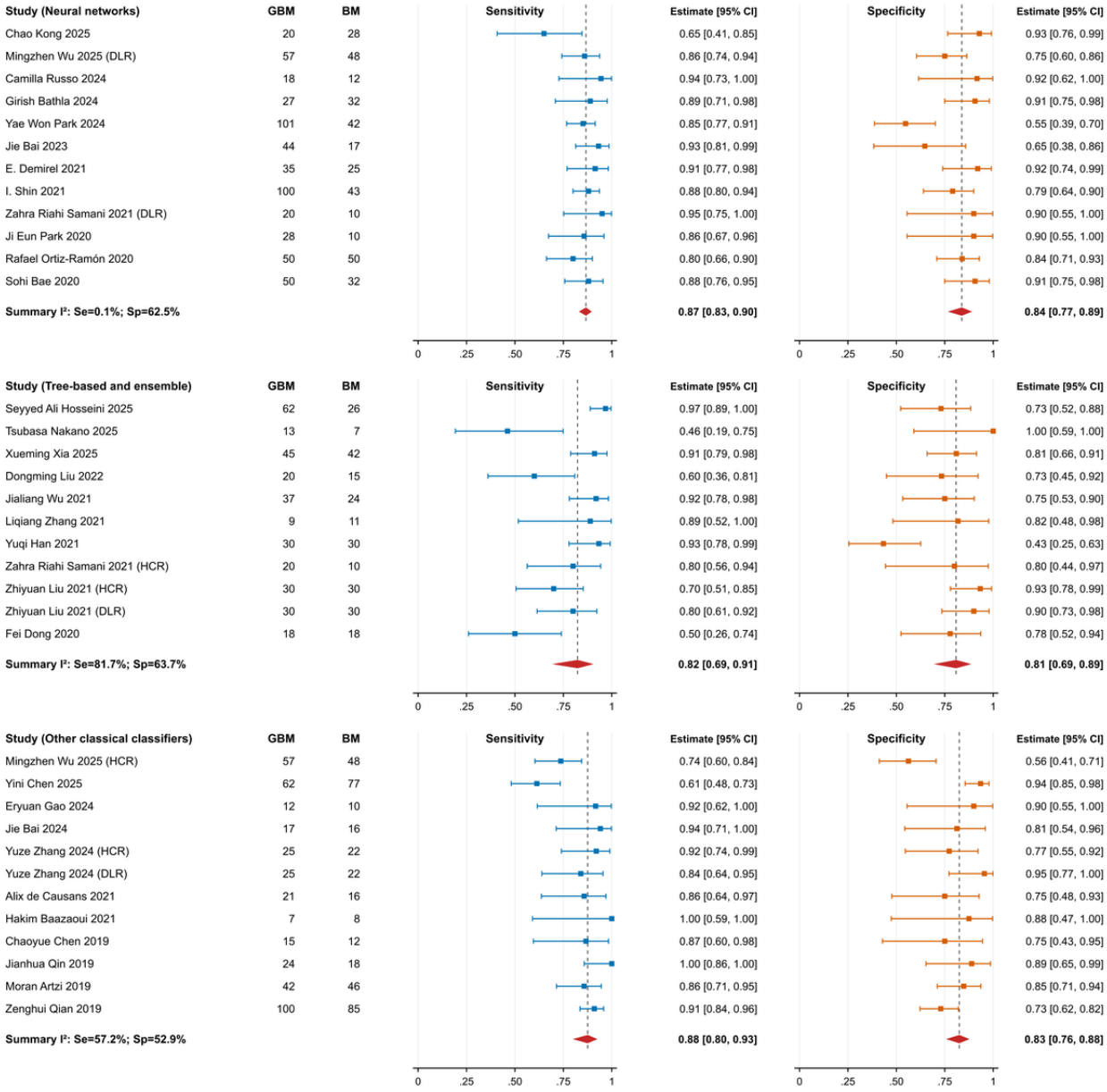
Subgroup Analyses

Within the internal validation data, exploratory subgroup analyses included results from cross-validation (n=5) and holdout validation (n=20; Figure S5 in Multimedia Appendix 2) [15,20,23-25,27-29,32,35,39,41,43,44,47-53,55,57]. Cross-validation yielded a pooled sensitivity of 93.5% (95% CI 83.2%-97.7%) and a specificity of 85.1% (95% CI 77.8%-90.4%), with an AUC of 0.955 (95% CI 0.895-0.981). Holdout validation yielded a pooled sensitivity of 87.1% (95% CI 82.7%-90.6%) and a specificity of 81.7% (95% CI 75.1%-86.8%), with an AUC of 0.911 (95% CI 0.873-0.939; Figure S6 in Multimedia Appendix 2). Bivariate metaregression did not detect a difference in sensitivity and specificity between cross-validation and holdout validation (P=.29).

Model results were grouped into neural networks (n=12), tree-based and ensemble models (n=11), and other classical classifiers (n=12; Figure 5) [15,20,21,23-29,32,34,35,39-44,46-57]. Neural networks yielded a pooled sensitivity of 86.7% (95% CI 83.0%-89.7%) and a specificity of 83.8% (95% CI 76.6%-89.1%), with an AUC of 0.870 (95% CI 0.801-0.918). Tree-based and ensemble models yielded a pooled sensitivity of 82.3% (95% CI 69.1%-90.6%) and a specificity of 80.8% (95% CI 69.3%-88.7%), with an AUC of 0.881 (95% CI 0.809-0.928). Other classical classifiers yielded a pooled sensitivity of 87.6% (95% CI 79.8%-92.6%) and a specificity of 82.5% (95% CI 75.6%-87.8%), with an AUC of 0.917 (95% CI 0.878-0.945; Figure S7 in Multimedia Appendix 2).

XSL•FO
RenderX

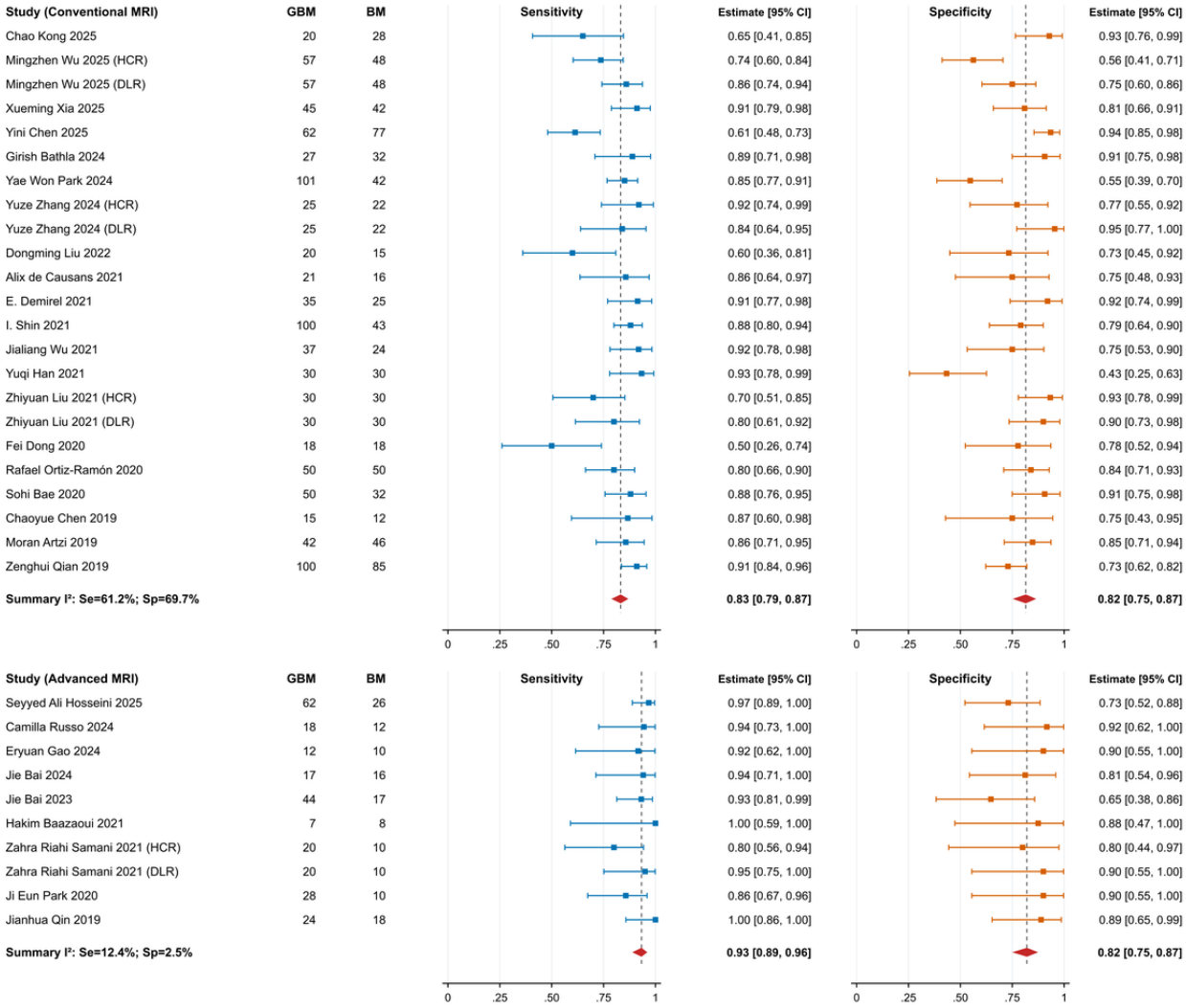
Figure 5. Forest plots of sensitivity and specificity by model type: neural networks, tree-based and ensemble models, and other classical classifiers [15,20,21,23-29,32,34,35,39-44,46-57]. BM: brain metastasis; DLR: deep learning radiomics; GBM: glioblastoma; HCR: handcrafted radiomics.



Imaging inputs were grouped into conventional MRI (n=23) and advanced MRI (n=10; Figure 6) [15,20,21,24-29,32,34,35,39,40,42-44,46-57]. Conventional MRI yielded a pooled sensitivity of 83.2% (95% CI 78.8%-86.9%) and a specificity of 81.5% (95% CI 75.2%-86.5%), with an AUC of 0.893 (95% CI 0.863-0.917). Advanced MRI yielded a pooled sensitivity of 93.3% (95% CI 88.8%-96.0%) and a specificity of 82.1% (95% CI 75.0%-87.5%), with an AUC of 0.911 (95% CI 0.580-0.987; Figure S8 in Multimedia Appendix 2).

XSL•FO
RenderX

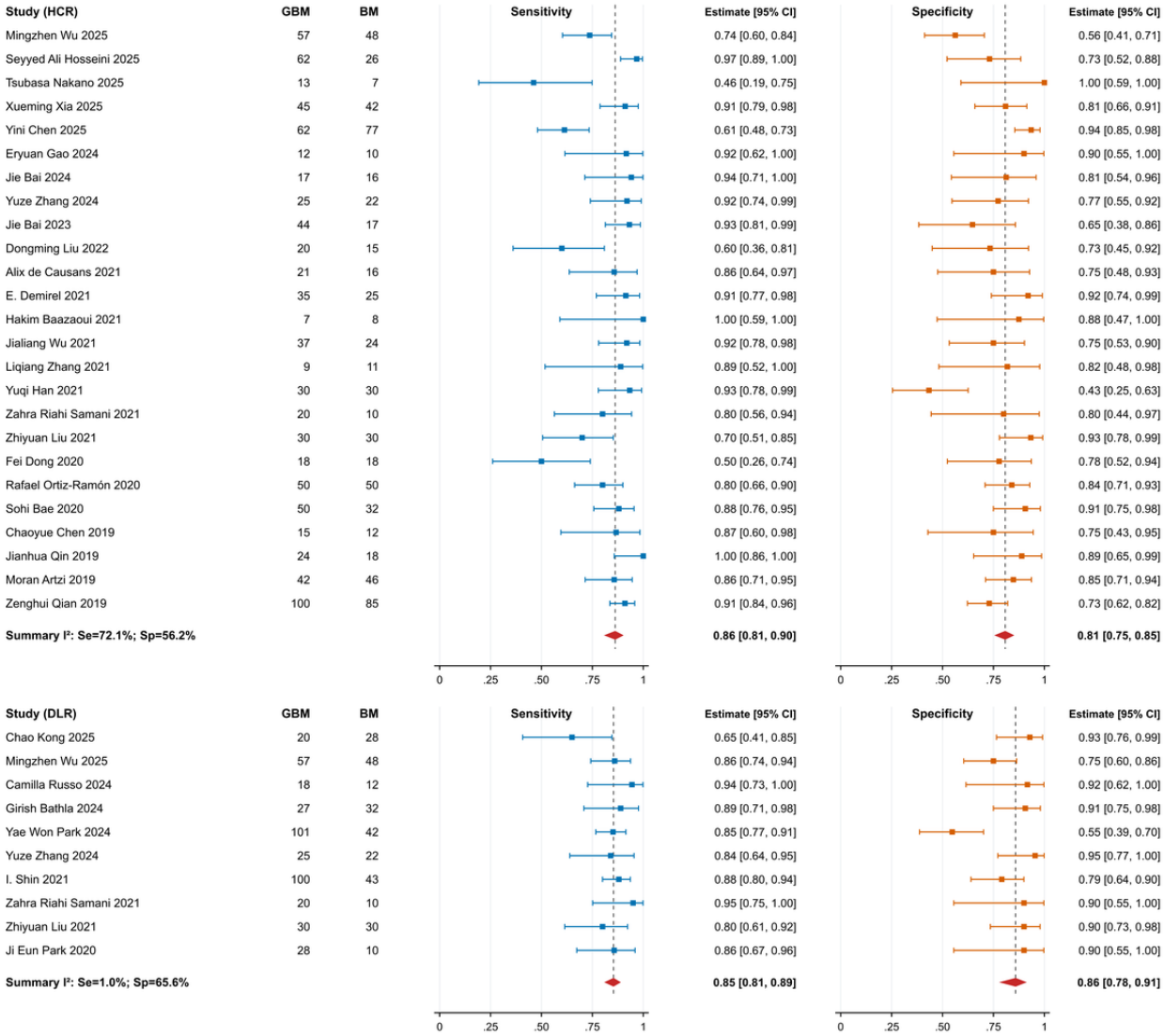
Figure 6. Forest plots of sensitivity and specificity by magnetic resonance imaging (MRI) input: conventional MRI and advanced MRI [15,20,21,24-29,32,34,35,39,40,42-44,46-57]. BM: brain metastasis; DLR: deep learning radiomics; GBM: glioblastoma; HCR: handcrafted radiomics.



For the subgroup analysis by feature extraction approach, 25 HCR and 10 DLR model results were summarized separately (Figure 7) [15,20,21,23-29,32,34,35,39-44,46-57]. HCR models yielded a pooled sensitivity of 86.2% (95% CI 80.7%-90.4%) and a specificity of 80.8% (95% CI 75.4%-85.2%), with an AUC of 0.899 (95% CI 0.867-0.924). DLR models yielded a pooled sensitivity of 85.4% (95% CI 80.8%-89.0%) and a specificity of 85.9% (95% CI 77.7%-91.4%), with an AUC of 0.869 (95% CI 0.825-0.903; Figure S9 in Multimedia Appendix 2).

XSL•FO
RenderX

Figure 7. Forest plots of sensitivity and specificity by feature extraction approach: handcrafted radiomics (HCR) and deep learning radiomics (DLR) [15,20,21,23-29,32,34,35,39-44,46-57]. BM: brain metastasis; GBM: glioblastoma.



Human Expert Performance

There were 29 physician assessments: 12 in the >10-year group and 17 in the ≤10-year group (Figure S10 in Multimedia Appendix 2) [21,26,34,42,43,49,54,56]. Pooled sensitivity and specificity were 84.3% (95% CI 71.6%-91.9%) and 79.7% (95% CI 70.8%-86.4%) in the >10-year group, and 70.8% (95% CI 61.1%-78.9%) and 68.3% (95% CI 56.3%-78.2%) in the ≤10-year group. The AUCs were 0.885 (95% CI 0.752-0.951) and 0.753 (95% CI 0.660-0.828), respectively (Figure S11 in Multimedia Appendix 2).

Multiclass Classification (GBM, BM, and PCNSL)

Multiclass analyses included 8 studies (Figure S12 in Multimedia Appendix 2) [22,30,31,33,36-38,45]. The pooled sensitivity was 84.4% (95% CI 72.4%-91.8%) for GBM, 70.6% (95% CI 54.4%-82.8%) for BM, and 75.0% (95% CI 62.7%-84.3%) for PCNSL. The corresponding specificities were 82.5% (95% CI 70.1%-90.4%), 88.8% (95% CI 81.9%-93.3%), and 94.4% (95% CI 88.4%-97.4%), respectively. Sensitivity differences were 13.8 percentage points for GBM-BM, 9.4 for GBM-PCNSL, and −4.4 for BM-PCNSL (Holm-adjusted P=.18 for GBM-BM, P=.45 for GBM-PCNSL, and P=.51 for BM-PCNSL). Specificity differences were −6.3, −11.9, and −5.6 percentage points for GBM-BM, GBM-PCNSL, and BM-PCNSL, respectively (Holm-adjusted P=.16 for GBM-BM, P=.15 for GBM-PCNSL, and P=.15 for BM-PCNSL).

Discussion

Summary of Main Findings

Radiomics-based ML models could discriminate between GBM and BM in the included cohorts, although pooled sensitivity and specificity were lower in external validation than in internal validation. In multiclass tasks, pooled class-specific sensitivity and specificity differed, although the adjusted comparisons were not statistically significant.

Subgroup analyses showed a higher pooled sensitivity for advanced MRI models than for conventional MRI models, with a similar specificity. Advanced MRI may provide additional microstructural information relevant to tumor infiltration and peritumoral edema [39,44]. The biological interpretation and reproducibility of these features require further validation. HCR

XSL•FO
RenderX

and DLR showed a similar pooled sensitivity, while DLR had a higher pooled specificity. Neural networks and other classical classifiers showed a higher pooled sensitivity than tree-based and ensemble models. These subgroup comparisons were descriptive.

The internal and external validation estimates were not fully paired, so the observed differences cannot be attributed solely to the validation setting. Differences in patient populations, imaging protocols, and model development may also contribute. Most included studies lacked independent external validation; therefore, their generalizability remains unclear. Further evaluation through multicenter prospective cohorts is needed to determine model performance across clinical settings.

Comparison With Prior Work

Building on previous reviews of ML-based discrimination between GBM and BM and the broader classification of brain tumors [16,17], this review further expanded the research scope by updating the literature search time frame to January 2026 and conducting separate analyses for internal and independent external validation. Furthermore, we examined model type, MRI input, and feature extraction approach in subgroup analyses and summarized physician performance as an exploratory benchmark. These subgroup analyses provide additional information for interpreting pooled diagnostic performance and its generalizability across clinical settings.

Habibi et al [16] used QUADAS-2 (Quality Assessment of Diagnostic Accuracy Studies 2) to assess the risk of bias in diagnostic accuracy studies [75]. In contrast, we used the RQS 2.0 assessment for radiomics-specific methodological quality and clinical readiness [58]. Its criteria addressed reproducibility, robustness, and explainability in both HCR and DLR. RRLs helped distinguish model development from clinical translation, with low clinical translation scores indicating limited evidence for deployment readiness.

Limitations

Our meta-analysis is constrained by the methodological quality of the included literature. The included studies primarily used retrospective designs, and nearly half were conducted in China (19/39, 48.7%), which may limit geographic representativeness. The scarcity of independent external validation limited the precision of our analysis and restricted further subgroup analyses.

Although external validation was prioritized, validation types remained unevenly distributed across subgroups. Differences in pooled diagnostic performance may therefore be partly attributed to validation design, model type, MRI input, and feature extraction approach. Physician assessments provided an exploratory benchmark, with interpretation dependent on reader experience, image quality, and available clinical information. Differences in patient populations and imaging protocols also affect comparability with model estimates. Future comparisons using the same patients and comparable diagnostic information would more clearly assess the performance of models and physicians. Clinical variables, including age and systemic malignancy history, may complement imaging features, although their incremental value requires further evaluation [27,42,57].

Some 2×2 tables were reconstructed from reported diagnostic metrics and sample sizes, which may introduce additional uncertainty into the pooled estimates. Selecting the best reported model may also lead to overly optimistic estimates. Reference standards also varied. Some BM diagnoses were based on clinical and imaging reference standards rather than histopathology, which may result in misclassifications. Furthermore, heterogeneity in imaging protocols, scanners, feature extraction approaches, and model architectures limits comparability.

Conclusions

Radiomics-based ML models could distinguish between GBM and BM in the included cohorts, with lower pooled sensitivity and specificity in external validation. Consequently, the generalizability and clinical utility of such models remain uncertain, owing to methodological limitations and scarce prospective evidence. Further multicenter prospective studies should conduct independent external validation in consecutive patients.

Data Availability

The full study characteristics and complete Radiomics Quality Score 2.0 scoring are provided in Multimedia Appendix 2. The extracted data used in the meta-analyses and the analysis code are available from the corresponding author upon reasonable request.

Funding

This work was supported by the National Natural Science Foundation of China (grant 82304952), the Beijing Natural Science Foundation (grant L2510078), and the Young Scientist Project of the National Key Research and Development Program of China (grant 2025YFC3510000). The funders provided financial support only and had no role in the study design, data collection, analysis, interpretation, or preparation of the manuscript.

Generative AI Use

Codex (GPT-6 Astra; OpenAI) was used solely to assist with figure preparation. All authors take full responsibility for the accuracy of the data, the rigor of the analyses, and the integrity of the manuscript.

XSL•FO
RenderX

Authors' Contributions

All authors contributed to the study conception and design. CH contributed to writing—original draft preparation. CH, GL, and ZB contributed to writing—review and editing. CH, PL, and SL contributed to conceptualization. YG contributed to methodology. YD and YH contributed to formal analysis. CH, ZB, YD, and YH contributed to investigation. CH, PL, and SL contributed to resources. PL and SL contributed to supervision. All authors commented on previous versions of the manuscript. All authors read and approved the final manuscript. SL is a co-corresponding author and can be reached at shiweili@mail.ccmu.edu.cn.

Conflicts of Interest

None declared.

Multimedia Appendix 1

PRISMA 2020 checklist.
[DOCX File , 21 KB-Multimedia Appendix 1]

Multimedia Appendix 2

Search strategies, study characteristics, quality assessments, and supplementary analyses.
[DOCX File , 4025 KB-Multimedia Appendix 2]

References

1. Schaff LR, Mellinghoff IK. Glioblastoma and other primary brain malignancies in adults: a review. JAMA. Feb 21, 2023;329(7):574-587. [doi: 10.1001/jama.2023.0023] [Medline: 36809318]
2. Obrador E, Moreno-Murciano P, Oriol-Caballo M, López-Blanch R, Pineda B, Gutiérrez-Arroyo JL, et al. Glioblastoma therapy: past, present and future. Int J Mol Sci. Feb 21, 2024;25(5):2529. [FREE Full text] [doi: 10.3390/ijms25052529] [Medline: 38473776]
3. Gritsch S, Batchelor TT, Gonzalez Castro LN. Diagnostic, therapeutic, and prognostic implications of the 2021 World Health Organization classification of tumors of the central nervous system. Cancer. Jan 01, 2022;128(1):47-58. [FREE Full text] [doi: 10.1002/cncr.33918] [Medline: 34633681]
4. Achrol AS, Rennert RC, Anders C, Soffietti R, Ahluwalia MS, Nayak L, et al. Brain metastases. Nat Rev Dis Primers. Jan 17, 2019;5(1):5. [doi: 10.1038/s41572-018-0055-y] [Medline: 30655533]
5. Pellerino A, Garbossa D, Rudà R, Soffietti R. The role of the neurologist in the diagnosis and treatment of brain metastases and carcinomatous meningitis. Rev Neurol (Paris). Jun 2023;179(5):464-474. [doi: 10.1016/j.neurol.2023.03.005] [Medline: 36990824]
6. Malone H, Yang J, Hershman DL, Wright JD, Bruce JN, Neugut AI. Complications following stereotactic needle biopsy of intracranial tumors. World Neurosurg. Oct 2015;84(4):1084-1089. [doi: 10.1016/j.wneu.2015.05.025] [Medline: 26008141]
7. Louis DN, Perry A, Wesseling P, Brat DJ, Cree IA, Figarella-Branger D, et al. The 2021 WHO Classification of Tumors of the Central Nervous System: a summary. Neuro Oncol. Aug 02, 2021;23(8):1231-1251. [FREE Full text] [doi: 10.1093/neuonc/noab106] [Medline: 34185076]
8. Chiang IC, Kuo YT, Lu CY, Yeung KW, Lin WC, Sheu FO, et al. Distinction between high-grade gliomas and solitary metastases using peritumoral 3-T magnetic resonance spectroscopy, diffusion, and perfusion imagings. Neuroradiology. Aug 2004;46(8):619-627. [doi: 10.1007/s00234-004-1246-7] [Medline: 15243726]
9. Law M, Cha S, Knopp EA, Johnson G, Arnett J, Litt AW. High-grade gliomas and solitary metastases: differentiation by using perfusion and proton spectroscopic MR imaging. Radiology. Mar 2002;222(3):715-721. [doi: 10.1148/radiol.2223010558] [Medline: 11867790]
10. Neska-Matuszewska M, Bladowska J, Sąsiadek M, Zimny A. Differentiation of glioblastoma multiforme, metastases and primary central nervous system lymphomas using multiparametric perfusion and diffusion MR imaging of a tumor core and a peritumoral zone-searching for a practical approach. PLoS One. Jan 17, 2018;13(1):e0191341. [FREE Full text] [doi: 10.1371/journal.pone.0191341] [Medline: 29342201]
11. Lambin P, Leijenaar RT, Deist TM, Peerlings J, de Jong EE, van Timmeren J, et al. Radiomics: the bridge between medical imaging and personalized medicine. Nat Rev Clin Oncol. Dec 2017;14(12):749-762. [doi: 10.1038/nrclinonc.2017.141] [Medline: 28975929]
12. Aerts HJ, Velazquez ER, Leijenaar RT, Parmar C, Grossmann P, Carvalho S, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. Nat Commun. Jun 03, 2014;5:4006. [FREE Full text] [doi: 10.1038/ncomms5006] [Medline: 24892406]
13. Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, van Stiphout RG, Granton P, et al. Radiomics: extracting more information from medical images using advanced feature analysis. Eur J Cancer. Mar 2012;48(4):441-446. [FREE Full text] [doi: 10.1016/j.ejca.2011.11.036] [Medline: 22257792]

XSL•FO
RenderX

14. Rudie JD, Rauschecker AM, Bryan RN, Davatzikos C, Mohan S. Emerging applications of artificial intelligence in neuro-oncology. Radiology. Mar 2019;290(3):607-618. [FREE Full text] [doi: 10.1148/radiol.2018181928] [Medline: 30667332]
15. Chen C, Ou X, Wang J, Guo W, Ma X. Radiomics-based machine learning in differentiation between glioblastoma and metastatic brain tumors. Front Oncol. Aug 22, 2019;9:806. [FREE Full text] [doi: 10.3389/fonc.2019.00806] [Medline: 31508366]
16. Habibi MA, Omid R, Asgarzade S, Derakhshandeh S, Soltani Farsani A, Tajabadi Z. Distinguishing glioblastoma from brain metastasis; a systematic review and meta-analysis on the performance of machine learning. Egypt J Neurosurg. Feb 06, 2025;40:26. [doi: 10.1186/s41984-025-00386-w]
17. Silva Santana L, Borges Camargo Diniz J, Mothé Glioche Gasparri L, Buccaran Canto A, Batista Dos Reis S, Santana Neville Ribeiro I, et al. Application of machine learning for classification of brain tumors: a systematic review and meta-analysis. World Neurosurg. Jun 2024;186:204-18.e2. [doi: 10.1016/j.wneu.2024.03.152] [Medline: 38580093]
18. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. Mar 29, 2021;372:n71. [FREE Full text] [doi: 10.1136/bmj.n71] [Medline: 33782057]
19. McInnes MD, Moher D, Thombs BD, McGrath TA, Bossuyt PM, the PRISMA-DTA Group, et al. Preferred reporting items for a systematic review and meta-analysis of diagnostic test accuracy studies: the PRISMA-DTA statement. JAMA. Jan 23, 2018;319(4):388-396. [FREE Full text] [doi: 10.1001/jama.2017.19163] [Medline: 29362800]
20. Xia X, Wu W, Tan Q, Gou Q. Interpretable machine learning models for differentiating glioblastoma from solitary brain metastasis using radiomics. Acad Radiol. Sep 2025;32(9):5388-5400. [FREE Full text] [doi: 10.1016/j.acra.2025.05.016] [Medline: 40436711]
21. Wu M, Luan J, Ma R, Zhang X, Chen J, Mu F, et al. Automatic segmentation and classification of glioblastoma and solitary brain metastasis using a deep learning model on multiparametric MRI. J Imaging Inform Med. Aug 2026;39(4):3563-3572. [doi: 10.1007/s10278-025-01695-0] [Medline: 41326876]
22. Sun Y, Xu J, Kong D, Zhang Y, Wu Q, Wei L, et al. Development and validation of a preoperative magnetic resonance imaging-based and machine learning model for the noninvasive differentiation of intracranial glioblastoma, primary central nervous system lymphoma and brain metastases: a retrospective analysis. Front Oncol. Apr 22, 2025;15:1541350. [FREE Full text] [doi: 10.3389/fonc.2025.1541350] [Medline: 40330833]
23. Nakano T, Hirahara D, Hasegawa T, Kamimura K, Nakajo M, Kamizono J, et al. Electron density and effective atomic number as quantitative biomarkers for differentiating malignant brain tumors: an exploratory study with machine learning. Tomography. Oct 29, 2025;11(11):120. [FREE Full text] [doi: 10.3390/tomography11110120] [Medline: 41295554]
24. Kong C, Yan D, Liu K, Yin Y, Ma C. Multiple deep learning models based on MRI images in discriminating glioblastoma from solitary brain metastases: a multicentre study. BMC Med Imaging. May 19, 2025;25(1):171. [FREE Full text] [doi: 10.1186/s12880-025-01703-3] [Medline: 40389875]
25. Hosseini SA, Servaes S, Hall B, Bhaduri S, Rajan A, Rosa-Neto P, et al. Quantitative physiologic MRI combined with feature engineering for developing machine learning-based prediction models to distinguish glioblastomas from single brain metastases. Diagnostics (Basel). Dec 27, 2024;15(1):38. [FREE Full text] [doi: 10.3390/diagnostics15010038] [Medline: 39795566]
26. Chen Y, Guo W, Li Y, Lin H, Dong D, Qi Y, et al. Differentiation of glioblastoma and solitary brain metastasis using brain-tumor interface radiomics features based on MR images: a multicenter study. Acad Radiol. Jul 2025;32(7):4164-4176. [FREE Full text] [doi: 10.1016/j.acra.2025.04.008] [Medline: 40280830]
27. Zhang Y, Zhang H, Zhang H, Ouyang Y, Su R, Yang W, et al. Glioblastoma and solitary brain metastasis: differentiation by integrating demographic-MRI and deep-learning radiomics signatures. J Magn Reson Imaging. Sep 2024;60(3):909-920. [doi: 10.1002/jmri.29123] [Medline: 37955154]
28. Russo C, Maresca P, Marinelli A. A cognitive computing tool for identification and classification of brain tumors. In: 2024 IEEE 8th Forum on Research and Technologies for Society and Industry Innovation (RTSI). New York, NY. IEEE; 2024:402-407.
29. Park YW, Eom S, Kim S, Lim S, Park JE, Kim HS, et al. Differentiation of glioblastoma from solitary brain metastasis using deep ensembles: empirical estimation of uncertainty for clinical reliability. Comput Methods Programs Biomed. Sep 2024;254:108288. [doi: 10.1016/j.cmpb.2024.108288] [Medline: 38941861]
30. Liu X, Liu J. Aided diagnosis model based on deep learning for glioblastoma, solitary brain metastases, and primary central nervous system lymphoma with multi-modal MRI. Biology (Basel). Feb 05, 2024;13(2):99. [FREE Full text] [doi: 10.3390/biology13020099] [Medline: 38392317]
31. Garcia-Ruiz A, Pons-Escoda A, Grussu F, Naval-Baudin P, Monreal-Aguero C, Hermann G, et al. An accessible deep learning tool for voxel-wise classification of brain malignancies from perfusion MRI. Cell Rep Med. Mar 19, 2024;5(3):101464. [FREE Full text] [doi: 10.1016/j.xcrm.2024.101464] [Medline: 38471504]
32. Gao E, Wang P, Bai J, Ma X, Gao Y, Qi J, et al. Radiomics analysis of diffusion kurtosis imaging: distinguishing between glioblastoma and single brain metastasis. Acad Radiol. Mar 2024;31(3):1036-1043. [doi: 10.1016/j.acra.2023.07.023] [Medline: 37690885]

XSL•FO
RenderX

33. Bathla G, Dhruba DD, Soni N, Liu Y, Larson NB, Kassmeyer BA, et al. AI-based classification of three common malignant tumors in neuro-oncology: a multi-institutional comparison of machine learning and deep learning methods. J Neuroradiol. May 2024;51(3):258-264. [doi: 10.1016/j.neurad.2023.08.007] [Medline: 37652263]
34. Bathla G, Dhruba DD, Liu Y, Le NH, Soni N, Zhang H, et al. Differentiation between glioblastoma and metastatic disease on conventional MRI imaging using 3D-convolutional neural networks: model development and validation. Acad Radiol. May 2024;31(5):2041-2049. [doi: 10.1016/j.acra.2023.10.044] [Medline: 37977889]
35. Bai J, He M, Gao E, Yang G, Zhang C, Yang H, et al. High-performance presurgical differentiation of glioblastoma and metastasis by means of multiparametric neurite orientation dispersion and density imaging (NODDI) radiomics. Eur Radiol. Oct 2024;34(10):6616-6628. [doi: 10.1007/s00330-024-10686-8] [Medline: 38485749]
36. Tariciotti L, Caccavella VM, Fiore G, Schisano L, Carrabba G, Borsa S, et al. A deep learning model for preoperative differentiation of glioblastoma, brain metastasis and primary central nervous system lymphoma: a pilot study. Front Oncol. Feb 24, 2022;12:816638. [FREE Full text] [doi: 10.3389/fonc.2022.816638] [Medline: 35280801]
37. Joo B, Ahn SS, An C, Han K, Choi D, Kim H, et al. Fully automated radiomics-based machine learning models for multiclass classification of single brain tumors: glioblastoma, lymphoma, and metastasis. J Neuroradiol. Jun 2023;50(4):388-395. [doi: 10.1016/j.neurad.2022.11.001] [Medline: 36370829]
38. Biggs ML, Wang Y, Soni N, Priya S, Bathla G, Canahuate G. Evaluating autoencoders for dimensionality reduction of MRI-derived radiomics and classification of malignant brain tumors. Sci Stat Database Manag. Jul 2023;2023:1. [FREE Full text] [doi: 10.1145/3603719.3603737] [Medline: 38344216]
39. Bai J, He M, Gao E, Yang G, Yang H, Dong J, et al. Radiomic texture analysis based on neurite orientation dispersion and density imaging to differentiate glioblastoma from solitary brain metastasis. BMC Cancer. Dec 14, 2023;23(1):1231. [FREE Full text] [doi: 10.1186/s12885-023-11718-0] [Medline: 38098041]
40. Liu D, Chen J, Ge H, Hu X, Yang K, Liu Y, et al. Differentiation of malignant brain tumor types using intratumoral and peritumoral radiomic features. Front Oncol. Jul 28, 2022;12:848846. [FREE Full text] [doi: 10.3389/fonc.2022.848846] [Medline: 35965511]
41. Zhang L, Yao R, Gao J, Tan D, Yang X, Wen M, et al. An integrated radiomics model incorporating diffusion-weighted imaging and 18F-FDG PET imaging improves the performance of differentiating glioblastoma from solitary brain metastases. Front Oncol. Aug 30, 2021;11:732704. [FREE Full text] [doi: 10.3389/fonc.2021.732704] [Medline: 34527594]
42. Wu J, Liang F, Wei R, Lai S, Lv X, Luo S, et al. A multiparametric MR-based radiofusionomics model with robust capabilities of differentiating glioblastoma multiforme from solitary brain metastasis. Cancers (Basel). Nov 18, 2021;13(22):5793. [FREE Full text] [doi: 10.3390/cancers13225793] [Medline: 34830943]
43. Shin I, Kim H, Ahn SS, Sohn B, Bae S, Park JE, et al. Development and validation of a deep learning-based model to distinguish glioblastoma from solitary brain metastasis using conventional MR images. AJNR Am J Neuroradiol. May 2021;42(5):838-844. [FREE Full text] [doi: 10.3174/ajnr.A7003] [Medline: 33737268]
44. Samani ZR, Parker D, Wolf R, Hodges W, Brem S, Verma R. Distinct tumor signatures using deep learning-based characterization of the peritumoral microenvironment in glioblastomas and brain metastases. Sci Rep. Jul 14, 2021;11(1):14469. [FREE Full text] [doi: 10.1038/s41598-021-93804-6] [Medline: 34262079]
45. Priya S, Liu Y, Ward C, Le NH, Soni N, Pillenahalli Maheshwarappa R, et al. Radiomic based machine learning performance for a three class problem in neuro-oncology: time to test the waters? Cancers (Basel). May 24, 2021;13(11):2568. [FREE Full text] [doi: 10.3390/cancers13112568] [Medline: 34073840]
46. Liu Z, Jiang Z, Meng L, Yang J, Liu Y, Zhang Y, et al. Handcrafted and deep learning-based radiomic models can distinguish GBM from brain metastasis. J Oncol. Jun 03, 2021;2021:5518717. [FREE Full text] [doi: 10.1155/2021/5518717] [Medline: 34188680]
47. Han Y, Zhang L, Niu S, Chen S, Yang B, Chen H, et al. Differentiation between glioblastoma multiforme and metastasis from the lungs and other sites using combined clinical/routine MRI radiomics. Front Cell Dev Biol. Aug 26, 2021;9:710461. [FREE Full text] [doi: 10.3389/fcell.2021.710461] [Medline: 34513840]
48. Demirel E, Gokaslan CO, Dilek O, Ozdemir C, Boyaci MG, Korkmaz S. Differential diagnosis of glioblastoma and solitary brain metastasis – the success of artificial intelligence models created with radiomics data obtained by automatic segmentation from conventional MRI sequences. Cesk Slov Neurol N. 2021;84(6):541-546. [doi: 10.48095/cccsnn2021541]
49. de Causans A, Carré A, Roux A, Tauziède-Espariat A, Ammari S, Dezamis E, et al. Development of a machine learning classifier based on radiomic features extracted from post-contrast 3D T1-weighted MR images to distinguish glioblastoma from solitary brain metastasis. Front Oncol. Jul 13, 2021;11:638262. [FREE Full text] [doi: 10.3389/fonc.2021.638262] [Medline: 34327133]
50. Baazaoui H, Hubertus S, Maros ME, Mohamed SA, Förster A, Schad LR, et al. Artificial neural network-derived cerebral metabolic rate of oxygen for differentiating glioblastoma and brain metastasis in MRI: a feasibility study. Appl Sci. Oct 24, 2021;11(21):9928. [doi: 10.3390/app11219928]
51. Park JE, Kim HS, Lee J, Cheong EN, Shin I, Ahn SS, et al. Deep-learned time-signal intensity pattern analysis using an autoencoder captures magnetic resonance perfusion heterogeneity for brain tumor differentiation. Sci Rep. Dec 08, 2020;10(1):21485. [FREE Full text] [doi: 10.1038/s41598-020-78485-x] [Medline: 33293590]

XSL•FO
RenderX

52. Ortiz-Ramón R, Ruiz-España S, Mollá-Olmos E, Moratal D. Glioblastomas and brain metastases differentiation following an MRI texture analysis-based radiomics approach. Phys Med. Aug 2020;76:44-54. [doi: 10.1016/j.ejmp.2020.06.016] [Medline: 32593138]
53. Dong F, Li Q, Jiang B, Zhu X, Zeng Q, Huang P, et al. Differentiation of supratentorial single brain metastasis and glioblastoma by using peri-enhancing oedema region-derived radiomic features and multiple classifiers. Eur Radiol. May 2020;30(5):3015-3022. [doi: 10.1007/s00330-019-06460-w] [Medline: 32006166]
54. Bae S, An C, Ahn SS, Kim H, Han K, Kim SW, et al. Robust performance of deep learning for distinguishing glioblastoma from single brain metastasis using radiomic features: model development and validation. Sci Rep. Jul 21, 2020;10(1):12110. [FREE Full text] [doi: 10.1038/s41598-020-68980-6] [Medline: 32694637]
55. Qin J, Li Y, Liang D, Zhang Y, Yao W. Histogram analysis of absolute cerebral blood volume map can distinguish glioblastoma from solitary brain metastasis. Medicine (Baltimore). Oct 2019;98(42):e17515. [FREE Full text] [doi: 10.1097/MD.0000000000017515] [Medline: 31626111]
56. Qian Z, Li Y, Wang Y, Li L, Li R, Wang K, et al. Differentiation of glioblastoma from solitary brain metastases using radiomic machine-learning classifiers. Cancer Lett. Jun 01, 2019;451:128-135. [doi: 10.1016/j.canlet.2019.02.054] [Medline: 30878526]
57. Artzi M, Bressler I, Ben Bashat D. Differentiation between glioblastoma, brain metastasis and subtypes using radiomics analysis. J Magn Reson Imaging. Aug 2019;50(2):519-528. [doi: 10.1002/jmri.26643] [Medline: 30635952]
58. Lambin P, Woodruff HC, Mali SA, Zhong X, Kuang S, Lavrova E, et al. Radiomics Quality Score 2.0: towards radiomics readiness levels and clinical translation for personalized medicine. Nat Rev Clin Oncol. Nov 2025;22(11):831-846. [doi: 10.1038/s41571-025-01067-1] [Medline: 40903523]
59. Reitsma JB, Glas AS, Rutjes AW, Scholten RJ, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. J Clin Epidemiol. Oct 2005;58(10):982-990. [doi: 10.1016/j.jclinepi.2005.02.022] [Medline: 16168343]
60. Fagan TJ. Letter: nomogram for Bayes's theorem. N Engl J Med. Jul 31, 1975;293(5):257. [doi: 10.1056/NEJM197507312930513] [Medline: 1143310]
61. Deeks JJ, Macaskill P, Irwig L. The performance of tests of publication bias and other sample size effects in systematic reviews of diagnostic test accuracy was assessed. J Clin Epidemiol. Sep 2005;58(9):882-893. [doi: 10.1016/j.jclinepi.2005.01.016] [Medline: 16085191]
62. Lin L, Chu H. Meta-analysis of proportions using generalized linear mixed models. Epidemiology. Sep 2020;31(5):713-717. [FREE Full text] [doi: 10.1097/EDE.0000000000001232] [Medline: 32657954]
63. Cepeda S, García-García S, Arrese I, Fernández-Pérez G, Velasco-Casares M, Fajardo-Puentes M, et al. Comparison of intraoperative ultrasound B-mode and strain elastography for the differentiation of glioblastomas from solitary brain metastases. An automated deep learning approach for image analysis. Front Oncol. 2020;10:590756. [FREE Full text] [doi: 10.3389/fonc.2020.590756] [Medline: 33604286]
64. Meier R, Pahud de Mortanges A, Wiest R, Knecht U. Exploratory analysis of qualitative MR imaging features for the differentiation of glioblastoma and brain metastases. Front Oncol. 2020;10:581037. [FREE Full text] [doi: 10.3389/fonc.2020.581037] [Medline: 33425734]
65. Pomohaci D, Marciuc EA, Dobrovă BI, Popescu MR, Istrate AC, Onicescu Oniciuc OM, et al. Artificial intelligence-based MRI segmentation for the differential diagnosis of single brain metastasis and glioblastoma. Diagnostics (Basel). Sep 05, 2025;15(17):2248. [FREE Full text] [doi: 10.3390/diagnostics15172248] [Medline: 40941735]
66. Bijari S, Jahanbakhshi A, Hajishafiezahramini P, Abdolmaleki P. Differentiating glioblastoma multiforme from brain metastases using multidimensional radiomics features derived from MRI and multiple machine learning models. Biomed Res Int. 2022;2022:2016006. [FREE Full text] [doi: 10.1155/2022/2016006] [Medline: 36212721]
67. Tateishi M, Nakaura T, Kitajima M, Uetani H, Nakagawa M, Inoue T, et al. An initial experience of machine learning based on multi-sequence texture parameters in magnetic resonance imaging to differentiate glioblastoma from brain metastases. J Neurol Sci. Mar 15, 2020;410:116514. [doi: 10.1016/j.jns.2019.116514] [Medline: 31869660]
68. An C, Park YW, Ahn SS, Han K, Kim H, Lee SK. Radiomics machine learning study with a small sample size: single random training-test set split may lead to unreliable results. PLoS One. 2021;16(8):e0256152. [FREE Full text] [doi: 10.1371/journal.pone.0256152] [Medline: 34383858]
69. Bathla G, Zamboni CG, Larson N, Liu Y, Zhang H, Lee NH, et al. Radiomics-based differentiation of glioblastoma and metastatic disease: impact of different T1-contrast-enhanced sequences on radiomics features and model performance. AJNR Am J Neuroradiol. Feb 03, 2025;46(2):321-329. [doi: 10.3174/ajnr.A8470] [Medline: 39179298]
70. Parvaze PS, Bhattacharjee R, Verma YK, Singh RK, Yadav V, Singh A, et al. Quantification of radiomics features of peritumoral vasogenic edema extracted from fluid-attenuated inversion recovery images in glioblastoma and isolated brain metastasis, using T1-dynamic contrast-enhanced perfusion analysis. NMR Biomed. May 2023;36(5):e4884. [doi: 10.1002/nbm.4884] [Medline: 36453877]
71. Csutak C, tefan PA, Lenghel LM, Moro anu CO, Lupean RA, imonca L, et al. Differentiating high-grade gliomas from brain metastases at magnetic resonance: the role of texture analysis of the peritumoral zone. Brain Sci. Sep 16, 2020;10(9):638. [FREE Full text] [doi: 10.3390/brainsci10090638] [Medline: 32947822]

XSL•FO
RenderX

72. Kopřivová T, Dostál M, Jůza T, Vybíhal V, Ovesná P, Kozubek M, et al. Differentiating brain metastasis and high-grade glioma using multi-b value diffusion MRI and tumor volumetry. J Neuroimaging. 2025;35(6):e70103. [doi: 10.1111/jon.70103] [Medline: 41229073]
73. Hong C, Wang H, Wu Z, Liu Z, Lv J. FoTNet: an effective deep learning model for preoperative differentiation of common malignant brain tumors. In: 2024 IEEE International Conference on Medical Artificial Intelligence (MedAI). New York, NY. IEEE; 2024.
74. Huang Y, Huang S, Liu Z. Multi-task learning-based feature selection and classification models for glioblastoma and solitary brain metastases. Front Oncol. 2022;12:1000471. [FREE Full text] [doi: 10.3389/fonc.2022.1000471] [Medline: 36212457]
75. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med. Oct 18, 2011;155(8):529-536. [FREE Full text] [doi: 10.7326/0003-4819-155-8-201110180-00009] [Medline: 22007046]

Abbreviations

AUC: area under the curve
BM: brain metastasis
DL: deep learning
DLR: deep learning radiomics
GBM: glioblastoma
HCR: handcrafted radiomics
ML: machine learning
MRI: magnetic resonance imaging
PCNSL: primary central nervous system lymphoma
PIRD: population, index test, reference standard, and diagnosis of interest
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
QUADAS-2: Quality Assessment of Diagnostic Accuracy Studies 2
RQS: Radiomics Quality Score
RRL: radiomics readiness level

Edited by M Balcarras; submitted 09.Jun.2026; peer-reviewed by P Iranmanesh, I Ramireddy; comments to author 28.Sep.2026; revised version received 15.Oct.2026; accepted 22.Oct.2026; published 19.Nov.2026

Please cite as:
Hong C, Liu G, Bie Z, Guo Y, Du Y, Hu Y, Li S, Liu P
Diagnostic Performance of Radiomics-Based Machine Learning in Differentiating Glioblastoma From Brain Metastasis: Systematic Review and Meta-Analysis
J Med Internet Res 2026;28:e104060
URL: https://www.jmir.org/2026/1/e104060
doi: 10.2196/104060
PMID:

©Chen Hong, Gemingtian Liu, Zhixu Bie, Yuanhang Guo, Yi Du, Yanan Hu, Shiwei Li, Pinan Liu. Originally published in the Journal of Medical Internet Research (https://www.jmir.org), 19.Nov.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research (ISSN 1438-8871), is properly cited. The complete bibliographic information, a link to the original publication on https://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX