



Get Clarity On Generics

Cost-Effective CT & MRI Contrast Agents



FRESENIUS
KABI

WATCH VIDEO

AJNR

Ten Years of VASARI Glioma Features: Systematic Review and Meta-Analysis of Their Impact and Performance

Aynur Azizova, Yeva Prysiazhniuk, Ivar J.H.G. Wamelink,
Jan Petr, Frederik Barkhof and Vera C. Keil

This information is current as
of August 3, 2025.

AJNR Am J Neuroradiol published online 27 June 2024
<http://www.ajnr.org/content/early/2024/06/27/ajnr.A8274>

Ten Years of VASARI Glioma Features: Systematic Review and Meta-Analysis of Their Impact and Performance

 Aynur Azizova,  Yeva Prisyazhniuk, Ivar J.H.G. Wamelink,  Jan Petr, Frederik Barkhof, and Vera C. Keil



ABSTRACT

BACKGROUND: Visually Accessible Rembrandt (Repository for Molecular Brain Neoplasia Data) Images (VASARI) features, a vocabulary to establish reproducible terminology for glioma reporting, have been applied for a decade, but a systematic performance evaluation is lacking.

PURPOSE: Our aim was to conduct a systematic review and meta-analysis of the performance of the VASARI features set for glioma assessment.

DATA SOURCES: MEDLINE, Web of Science, EMBASE, and the Cochrane Library were systematically searched until September 26, 2023.

STUDY SELECTION: Original articles predicting diagnosis, progression, and survival in patients with glioma were included.

DATA ANALYSIS: The modified Quality Assessment of Diagnostic Accuracy Studies-2 (QUADAS-2) tool was applied to evaluate the risk-of-bias. The meta-analysis used a random effects model and forest plot visualizations, if ≥ 5 comparable studies with a low or medium risk of bias were provided.

DATA SYNTHESIS: Thirty-five studies (3304 patients) were included. Risk-of-bias scores were medium ($n=33$) and low ($n=2$). Recurring objectives were overall survival ($n=18$) and isocitrate dehydrogenase mutation (*IDH*; $n=12$) prediction. Progression-free survival was examined in 7 studies. In 4 studies (glioblastoma $n=2$, grade 2/3 glioma $n=1$, grade 3 glioma $n=1$), a significant association was found between progression-free survival and single VASARI features. The single features predicting overall survival with the highest pooled hazard ratios were multifocality (hazard ratio = 1.80; 95%-CI, 1.21–2.67; $I^2 = 53\%$), ependymal invasion (hazard ratio = 1.73; 95% CI, 1.45–2.05; $I^2 = 0\%$), and enhancing tumor crossing the midline (hazard ratio = 2.08; 95% CI, 1.35–3.18; $I^2 = 52\%$). *IDH* mutation-predicting models combining VASARI features rendered a pooled area under the receiver operating characteristic curve of 0.82 (95% CI, 0.76–0.88) at considerable heterogeneity ($I^2 = 100\%$). Combined input models using VASARI plus clinical and/or radiomics features outperformed single data-type models in all relevant studies ($n=17$).

LIMITATIONS: Studies were heterogeneously designed and often with a small sample size. Several studies used The Cancer Imaging Archive database, with likely overlapping cohorts. The meta-analysis for *IDH* was limited due to a high study heterogeneity.

CONCLUSIONS: Some VASARI features perform well in predicting overall survival and *IDH* mutation status, but combined models outperform single features. More studies with less heterogeneity are needed to increase the evidence level.

ABBREVIATIONS: AUC = area under the receiver operating characteristic curve; HR = hazard ratio; *IDH* = isocitrate dehydrogenase; OS = overall survival; VASARI = Visually Accessible Rembrandt [Repository for Molecular Brain Neoplasia Data] Images

MRI is the essential pillar of preoperative glioma diagnosis and later therapy assessment. While there is consensus

regarding a standardized neuro-oncologic imaging protocol,¹ reporting of image aspects of glioma (radiophenotype) is less

Received February 4, 2024; accepted after revision March 1.

From the Radiology and Nuclear Medicine Department (A.A., I.J.H.G.W., J.P., F.B., V.C.K.), Amsterdam UMC location Vrije Universiteit Amsterdam, Amsterdam, the Netherlands; Imaging and Biomarkers (A.A., I.J.H.G.W., V.C.K.), Cancer Center Amsterdam, Amsterdam, the Netherlands; The Second Faculty of Medicine (Y.P.), Department of Pathophysiology, Charles University, Prague, Czech Republic; Institute of Radiopharmaceutical Cancer Research (J.P.), Helmholtz-Zentrum Dresden-Rossendorf, Dresden, Germany; Brain Imaging (F.B., V.C.K.), Amsterdam Neuroscience, Amsterdam, the Netherlands; and Queen Square Institute of Neurology and Center for Medical Image Computing (F.B.), University College London, London, United Kingdom.

This study received funding from the Hanarth Foundation.

Professor Barkhof is supported by the National Institute for Health and Care Research Biomedical Research Center at University College London Hospitals.

Ms Azizova is supported by the European Society of Neuroradiology Research Fellowship Grant.

Please address correspondence to Vera C. Keil, MD, PhD, Radiology and Nuclear Medicine Department, Amsterdam UMC location Vrije Universiteit Amsterdam, Amsterdam, the Netherlands; e-mail: v.c.w.keil@amsterdamumc.nl; @JanHonzaPetr; @FBarkhof



Indicates article with online supplemental data.

<http://dx.doi.org/10.3174/ajnr.A8274>

SUMMARY

PREVIOUS LITERATURE: The VASARI feature set for standardized glioma reporting has been applied by numerous studies to address different research questions. Applications range from human radiologic evaluations by using single features to multivariable machine-learning approaches. Clinical research questions encompass glioma subtype discrimination and noninvasive image-based survival prediction. Prediction of overall survival and IDH mutation status was among the most explored research questions. Several studies have identified multifocality, ependymal invasion, and enhancing tumor crossing the midline as unfavorable overall survival predictors. Enhancement quality, proportion of enhancing tumor, proportion of necrosis, and proportion of edema were identified as the main indicators of IDH mutation status.

KEY FINDINGS: This meta-analysis identified 3 robust VASARI features (multifocality, ependymal invasion, enhancing tumor crossing midline) to predict overall survival. Meta-analysis for IDH-predicting models showed a pooled AUC of 0.82 with considerable heterogeneity between studies. Combined models, including VASARI features next to clinical, genomics, and radiomics features, usually outperformed VASARI-only models.

KNOWLEDGE ADVANCEMENT: Future studies should adhere to the original VASARI scoring definitions to minimize between-study heterogeneity. Given the time-consuming nature of manual extraction, it is crucial to develop automatic extraction technology. Reducing the feature set to the most promising ones can decrease the workload when radiologist input is required.

standardized. Criteria, as defined by Response Assessment in Neuro-Oncology (RANO 2.0),² comprise only limited standards relying on quantitative measurements dedicated to the follow-up setting. In a study setting, however, a controlled reporting vocabulary is needed to identify reproducible imaging glioma biomarkers or generate data input suitable for artificial intelligence approaches.

Approximately a decade ago, The Cancer Genome Atlas project of the National Cancer Institute addressed this problem by suggesting a controlled vocabulary for glioma imaging called Visually Accessible Rembrandt (Repository for Molecular Brain Neoplasia Data) Images (VASARI).³ The VASARI set combines different MRI features (Online Supplemental Data), such as enhancement pattern and tumor location. A consensus group defined the features on the basis of expert opinion and literature. The VASARI project incorporated standard MRI sequences, including DWI, but excluded advanced imaging techniques like PWI. The current set comprises 30 semantic features, 3 of which apply to postoperative situations. All features are rated on the basis of scoring systems (Online Supplemental Data).

Numerous studies have applied the VASARI set since its proposal.⁴⁻³⁸ Applications range from human radiologic evaluations using single features⁵⁻⁸ to multivariable machine learning approaches.^{21,27-29} Clinical research questions cover glioma subtype discrimination^{13,14,16,17} and noninvasive image-based survival prediction.^{7,10,13,22} Various VASARI features were identified as prognostic factors, including tumor location, involvement of eloquent brain areas, ependymal or pial invasion, as well as diagnostic indicators, such as the definition of the enhancing margin and the proportion of necrotic or enhancing tumor, in these studies. However, these applications have never been systematically evaluated, perhaps explaining why using the VASARI features set is not recommended by any clinical or scientific guideline. Because complete feature rating is time-consuming, a critical analysis identifying the most powerful features and models is pivotal for future study designs and potential clinical use. Recent, however not VASARI-centered, publications indicated pial and

subependymal invasion (features 18 and 19) as negative for survival.³⁹ A meta-analysis may confirm the hypothesis of features 18 and 19 as particularly promising predictive VASARI features.

This systematic review and meta-analysis aimed to gauge the performance of the VASARI set for glioma evaluation to identify a subset of the most diagnostic and prognostic predictive features to warrant use in trials or even clinically.

MATERIALS AND METHODS

This study was registered in PROSPERO (ID for the published protocol: CRD42023392548) and was conducted according to Preferred Reporting Items for Systematic Review and Meta-Analysis (PRISMA).⁴⁰ The research question was “What is the VASARI features’ performance in predicting diagnosis, progression, and survival in patients with glial tumors?” using the participants, index tests, and target conditions criteria.⁴¹

Data Sources and Screening Step

The search string had 2 components: VASARI and tumor types, including high-grade and low-grade oligodendroglioma, astrocytoma, and glioblastoma. Inclusion criteria were the following: 1) participants: adult or pediatric patients with glial tumors; 2) intervention/index test: human or automated methods using multiparametric conventional MRI as a source for VASARI; 3) comparison: standard interpretation of images with or without VASARI; 4) outcomes: prediction of the diagnosis, including histologic tumor grades and certain well-known mutations (isocitrate dehydrogenase [*IDH*], 1p/19q codeletion, telomerase reverse transcriptase promoter, and O[6]-methylguanine-DNA methyltransferase methylation status), progression, or survival; and 5) study design: original articles with a retrospective or prospective design.

Exclusion criteria were the following: 1) studies not performed on living humans; 2) studies with <10 participants (considered case series); 3) studies exclusively predicting other genetic/molecular alterations than indicated in the inclusion

criteria; 4) non-peer-reviewed journals, conference abstracts, review papers, preprints; and 5) publications not in English.

A database search was performed on September 26, 2023, using MEDLINE (PubMed), the Web of Science, EMBASE, and the Cochrane Library. The search protocol is presented as Online Supplemental Data. One author screened the titles and abstracts (A.A., 7 years of radiology experience) for inclusion criteria after excluding duplicates using rayyan software (<https://rayyan.ai/reviews>). Additionally, the reference lists of included articles were screened by title and abstract (hand search). Uncertain cases were evaluated in consultation with another author (V.C.K., 11 years of neuroradiology experience).

Data Extraction

To guarantee identical rating standards, we used 5 studies to pilot the extraction process by 2 authors (A.A., V.C.K.). Two authors then performed data extraction for the remaining studies (A.A., for radiologic content, and Y.P., a second-year PhD student in neuroscience with expertise in statistics). The Online Supplemental Data list the data-extraction components.

Systematic Review Quality Assessment

The quality of each study was evaluated by 2 authors (A.A., V.C.K.) using the modified Quality Assessment of Diagnostic Accuracy Studies-2 (QUADAS-2; <https://www.bristol.ac.uk/population-health-sciences/projects/quadas/quadas-2/>) tool,⁴² encompassing 5 domains for assessing the risk of bias and 3 domains for evaluating applicability concerns (Online Supplemental Data). Discrepancies were resolved by consensus. Five studies, also used to pilot data extraction, were piloted for quality control assessment to identify systematic discrepancies in QUADAS-2 tool use.

Meta-Analysis and Statistics

Studies examining patient cohorts with a similar tumor type, statistical models with comparable inputs, and reported identical end points of interest were grouped. The availability of ≥ 5 studies with most QUADAS-2 categories scoring low or medium risk of bias was the liberal minimum for a meta-analysis. Otherwise, a narrative synthesis summarized the findings. If studies with overlapping cohorts were available, the total sample size was determined by including the largest cohort from overlapping studies (Online Supplemental Data). In case of missing data, other reported metrics (standard errors, sample sizes)⁴³ served to derive parameters when feasible. Alternatively, corresponding authors were contacted and requested to provide supplementary information. Meta-analyses were conducted in R (Version 4.3.0.; <http://www.r-project.org/>) using generic inverse-variance and random effect models to account for methodologic between-study heterogeneity. To mitigate interpretability concerns arising from potential collinearity within VASARI data sets, the meta-analysis of hazard ratios (HRs) used metrics derived solely from univariable Cox proportional models. Heterogeneity was assessed with Higgins I^2 statistics, considering values above 50% as significant heterogeneity. All tests were 2-sided. Because this is a meta-analysis, no correction was done for multiple testing. Forest plots were used to visualize results, complemented by pooled statistics with corresponding confidence and prediction intervals. The best-

performing VASARI model was chosen if >1 model had been tested. The criteria for performance of meta-regression, sensitivity analysis, and nonreporting bias analysis were the availability of at least 10 studies to ensure the reliability of conclusions drawn from these analyses according to the Cochrane Handbook for Systematic Reviews of Interventions.⁴⁴

RESULTS

Overview

Figure 1 visualizes the literature selection. Thirty-five studies (29–335 patients) fulfilled the inclusion criteria (Online Supplemental Data). Eleven studies used The Cancer Imaging Archive cohort, possibly covering overlapping cohorts. The overlap-corrected number of recruited patients was 3304 (Online Supplemental Data). The overall risk-of-bias scores were medium ($n = 33$) and low ($n = 2$; Online Supplemental Data). Only 2 studies applied the 2021 World Health Organization Classification of Central Nervous System Tumors (WHO CNS 5)³⁸ or reported *IDH* wild-type glioblastoma only.¹⁶ The remaining 33 studies explicitly used the 2016 or 2007 WHO CNS classifications ($n = 5^{19,22,26,27,30}$) or did not mention it ($n = 28^{4-15,17,18,20,21,23-25,28,29,31-37}$), resulting in applicability uncertainties regarding patient selection. One study³⁸ was vague in the use of the reference standard.

Survival Prediction

Fifteen^{4,5,7-11,15,18,21,24,27,29,34,35} of 18 overall survival (OS)-predicting studies exclusively involved glioblastomas. Results for univariable analyses indicating the per-feature performance of the VASARI set are shown in Fig 2A. Eleven glioblastoma studies found a significant association between OS and different VASARI features.^{5,7-10,15,18,24,29,34,35} Comparing single-data-type models with only 1 input datum type (e.g., clinical, pathologic, FET-PET/CT, or genomics data), VASARI-based models outperformed others in 2 studies.^{21,27} However, a clinical model (age and Karnofsky performance status) slightly outperformed a VASARI-only model in the study by Peeken et al¹⁸ (C-indexes = 0.64; 95% CI, 0.55–0.72 versus 0.66; 95% CI, 0.58–0.73). Compared with VASARI-only models, several studies^{9,15,18,21,27} stressed the superiority of predictive models trained on combined features, including VASARI plus clinical, genomics, treatment, and/or radiomics features. Peeken et al¹⁸ concluded that a combined model (VASARI, clinical, age, and Karnofsky performance status) and pathologic features (O[6]-methylguanine-DNA methyltransferase methylation) performed better than the VASARI-only model (C-indexes = 0.72; 95% CI, 0.61–0.80 versus 0.64; 95% CI, 0.55–0.72). Mazurowski et al⁴ showed the added value of VASARI features (VASARI+clinical: area under the receiver operating characteristic curve [AUC] 0.81; 95% CI, 0.71–0.90; C-index = 0.69; 95% CI, 0.63–0.75) compared with clinical-only, including age, sex, and Karnofsky performance status (AUC = 0.62; 95% CI, 0.49–0.74; C-index 0.58; 95% CI, 0.50–0.66; both $P < .01$).

Three OS-predicting studies recruited grade 2 and 3 gliomas (Fig 2A). Zhou et al¹³ found a significant association between OS and 10 VASARI features in a univariable analysis of which only 1 feature, definition of the nonenhancing margin, was significant in multivariable analyses. They demonstrated that a well-defined

nonenhancing margin correlated with longer OS than an ill-defined nonenhancing margin. Lee et al²² showed that the performance of clinical features (age, Karnofsky performance status, and extent of resection) plus molecular subtype (*IDH* mutation and 1p/19q codeletion status) model increased when extending it with VASARI features (C-indices = 0.84; 95% CI, 0.75–0.90 versus 0.91; 95% CI, 0.86–0.96). Similar results were found by Park et al,²⁶ improving the integral AUC (initially 0.74; 95% CI, 0.69–0.81) of the clinical model (age, Karnofsky performance status, extent of resection, and histologic grade) by adding VASARI features (LASSO model 0.77; <https://tullo.ch/articles/lasso-estimator>; 95% CI, 0.74–0.85; elastic net model = 0.78; 95% CI, 0.74–0.85).

Glioblastoma was included in 5 of 7 progression-free survival predicting studies, with 2 studies^{10,18} showing a significant association with single VASARI features (Fig 2B). Peeken et al²¹ evaluated the performance of a combined model (VASARI plus clinical age, sex and Karnofsky performance status) plus pathologic (O[6]-methylguanine-DNA methyltransferase methylation, *IDH* mutation status, and Ki 67 proliferation index) showing superior performance compared with single-data-type models including their VASARI-only model (C-index validation set = 0.68; 95% CI, 0.57–0.78 versus 0.61; 95% CI, 0.50–0.72, $P = .014$, respectively). In that study, VASARI features showed the highest performance in the validation set among single-data-type models (C-index = 0.61; 95% CI, 0.50–0.72). They were also the dominant variable in the combined model (VASARI plus clinical plus pathologic features).

Two of 7 progression-free survival-predicting studies involved grade 2 and 3¹³ or grade 3 gliomas²² (Fig 2B). Lee et al²² showed that adding VASARI features increased the performance of the combined model (clinical plus molecular) with C-indices = 0.79 (95% CI, 0.71–0.85) versus 0.84 (95% CI, 0.79–0.91).

IDH Mutation Status Prediction

Reported tumor entities were glioblastoma^{5,14,29,30} and grade 2 to 4 gliomas^{13,17,20,25,28,32,36,37} (Online Supplemental Data). Individually evaluated VASARI features are shown in Fig 2C. Five studies^{5,13,28,29,32} found no association between *IDH* mutation and individual VASARI features. Among other studies, the main features that were more consistently identified as significant predictors of *IDH* mutation status were enhancement quality (feature 4),^{17,20,30,36,37} proportion of enhancing tumor (feature 5),^{17,20,30,37} proportion of necrosis (feature 7),^{17,30,36,37} and proportion of edema (feature 14).^{14,17,20,36} There was a negative correlation between these features and the presence of an *IDH* mutation.

A higher performance of combined models was shown in several studies.^{13,20,28,32} Su et al²⁰ showed that a VASARI model (using feature 6: proportion of nonenhancing tumor) outperformed the diffusion texture analysis model with an AUC of 0.92 (95% CI, 0.80–0.98 versus 0.72; 95% CI, 0.57–0.85). The proportion of nonenhancing tumor was significantly higher in *IDH*-mutant high-grade gliomas compared with *IDH* wild-type high-grade gliomas. However, the combined model (proportion of nonenhancing tumor plus ADC entropy) was the best-performing model (AUC = 0.95; 95% CI, 0.85–0.99). Sun et al³² determined

an imaging model (VASARI features plus the T2 FLAIR mismatch sign) as the best single data-type model (AUC = 0.75; 95% CI, 0.60–0.89) in the test set.

Other Studies

Six studies predicted histologic tumor grade^{12,13,19,33,37,38} or further mutations (1p/19q codeletion, $n = 3$;^{13,17,32} O[6]-methylguanine-DNA methyltransferase methylation, $n = 2$;^{16,29} telomerase reverse transcriptase promoter, $n = 3$ ^{16,23,30}) or the presence of true progression ($n = 3$),^{6,13,31} respectively. The characteristics of these studies are summarized in the Online Supplemental Data.

Meta-Analysis

Survival Prediction. The meta-analysis of individual VASARI features to predict OS in glioblastoma studies identified 8 features with a HR above 1.25 (Fig 3; HR range = 1.32–2.08) and 7 equal to or below it (Fig 4; HR range = 0.89–1.25). The HR cutoff of 1.25 was arbitrarily defined for visual representation. Enhancing tumor crossing the midline (feature 23, pooled HR = 2.08; 95% CI, 1.35–3.18), multifocality (feature 9, pooled HR = 1.80; 95% CI, 1.21–2.67), and ependymal invasion (feature 19, pooled HR = 1.73; 95% CI, 1.45–2.05) were the strongest predictors.

Figure 5 gives a detailed overview of HRs for individual VASARI features, including those features that were not selected for the meta-analysis, lacking the mandatory minimum number of 5 articles. Meta-regressions were not possible due to a lack of homogeneous articles.

Meta-analysis of progression-free survival-predicting studies could not be conducted because only 2 studies^{10,18} met the necessary prerequisites for this analysis. An overview of HRs for individual VASARI features used in these 2 studies is provided in the Online Supplemental Data.

IDH Status Prediction

IDH status studies displayed a very heterogeneous design regarding prediction models and VASARI features (Fig 6). None of these studies used the same feature combination, allowing only a general approach (Are multivariable VASARI feature models powerful in predicting *IDH* status?). These studies^{13,17,20,25,28,32,37} included grade 1–4 gliomas. The AUC of these models was good, ranging from 0.73 to 0.92 (Fig 6). Meta-regressions were not possible due to a lack of homogeneous articles.

DISCUSSION

This study shows that VASARI features have primarily been used to predict OS and *IDH* mutation status. A meta-analysis of OS-predicting studies revealed that the 3 most robust single features in determining OS were multifocality (pooled HR = 1.80), ependymal invasion (pooled HR = 1.73), and enhancing tumor crossing the midline (pooled HR = 2.08), confirming our hypothesis that some features are stronger predictors than others. In a meta-analysis of *IDH* mutation-predicting VASARI models that incorporated different combinations of single VASARI features, the pooled AUC was 0.82, with considerable variability among single studies. Combined models incorporating different nonimaging data types outperformed single-data-type models, including

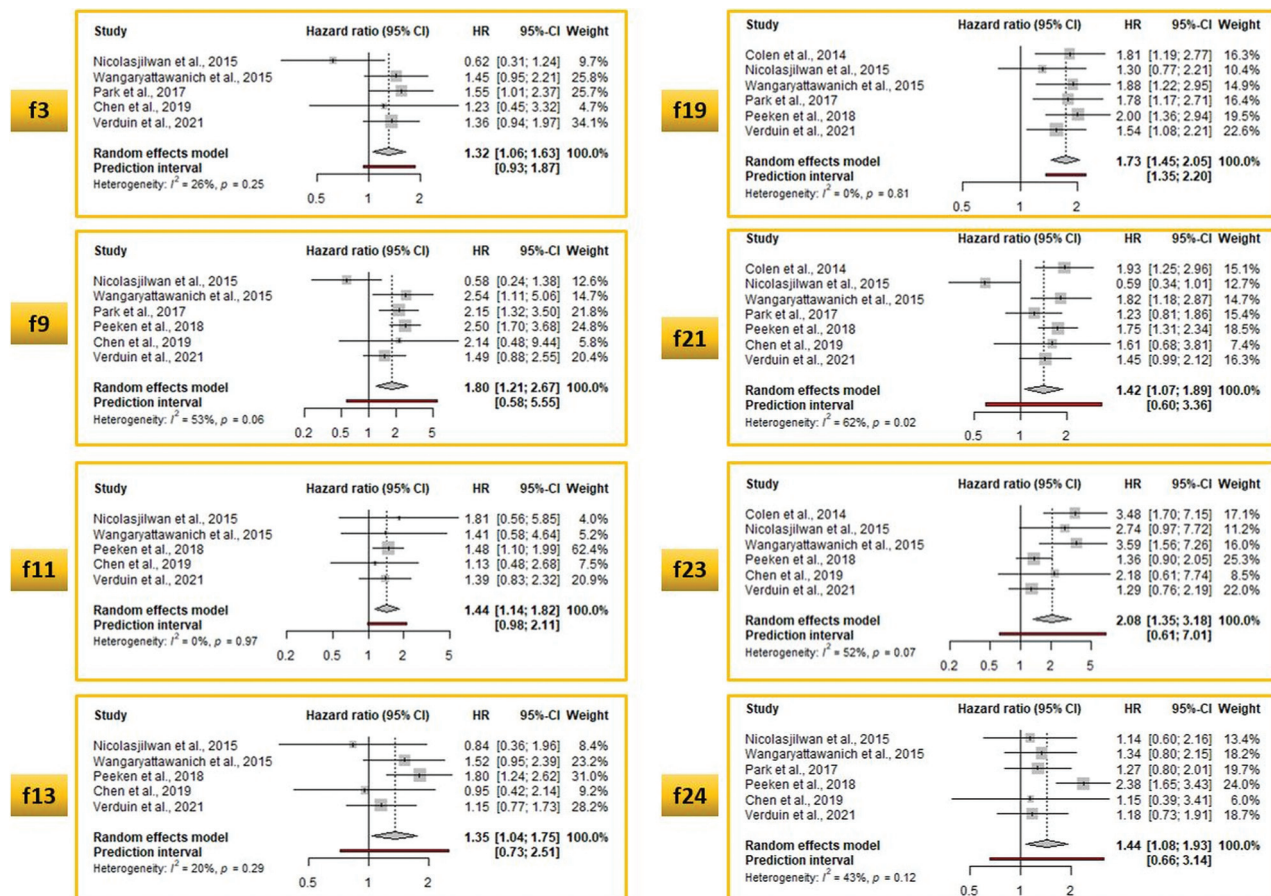


FIG 3. Meta-analysis of OS-predicting studies with a pooled HR >1.25. Note that the list of VASARI features and their detailed descriptions, along with respective scoring systems for each feature are provided in Online Supplemental Data. f indicates VASARI feature; f3, eloquent brain; f9, multifocality; f11, thickness of enhancing margin; f13, definition of nonenhancing margin; f19, ependymal invasion; f21, deep white matter invasion; f23, enhancing tumor crosses the midline; f24, satellites.

VASARI-only models, in determining survival, mutation status, or grades of glial tumors.

The survival rates of patients with gliomas remain low despite an aggressive treatment strategy. This outcome is likely the main reason why prognosis prediction was found to be the primary objective of the included studies. Multifocal tumor distribution (feature 9) and ependymal invasion (feature 19) were among the most unfavorable OS predictors,^{10,15,18,29} which could be prioritized in studies when evaluating all 30 features is impractical. Thomas et al⁴⁵ also demonstrated a significant correlation between multiple lesions and other negative prognostic indicators, such as a low Karnofsky performance score and resection volume. Lim et al⁴⁶ found that newly diagnosed glioblastoma with ependymal invasion and cortical involvement were more likely to have multifocal distribution and noncontiguous tumor recurrence with the initial lesion. The midline crossing enhancing tumor (feature 23) had the highest pooled HR of 2.08 and was the strongest OS-limiting predictor. Wangaryattawanich et al¹⁰ observed a 9.2-month OS difference (4.8 versus 14 months, $P = .001$) and a 4.2-month progression-free survival difference (2.4 versus 6.6 months, $P < .001$) in cases separated by whether the enhancing tumor was crossing the midline (feature 23); similar results were shown by Colen et al⁸ with mean OS of 5.9 versus 14.3 months ($P < .001$).

According to the 2021 WHO CNS classification, the *IDH* mutation status is crucial to classifying adult-type diffuse gliomas.⁴⁷ The term “*IDH*-mutant glioblastoma” has been changed to “*IDH*-mutant astrocytoma” requiring *IDH* wild-type status for glioblastomas. Our study identified 4 VASARI features (enhancement quality,^{17,20,30,36,37} proportion of enhancing tumor,^{17,20,30,37} proportion of necrosis,^{17,30,36,37} and proportion of edema^{14,17,20,36}) as the main indicators of *IDH* mutation status. However, it was not possible to quantitatively evaluate the impact of these features on *IDH* status prediction due to the limited number of studies and their methodologic differences.

Combined feature models, including various variables, such as clinical, imaging/VASARI, radiomics, genomics, or pathologic features, predicting OS, or other objectives, outperformed single-data-type models, showing the importance of a multidisciplinary approach in decision-making. The input used for these models varied among studies. However, certain factors were frequently chosen, such as age and Karnofsky performance score for clinical models,^{4,18,22,26} *IDH* mutation, 1p/19q codeletion, and O(6)-methylguanine-DNA methyltransferase or histologic grade^{18,21,22} for pathologic/genomics models, and shape or texture features for radiomics models.^{24,31,32,34} Nevertheless, the methodologic heterogeneity among those studies, such as different sample sizes

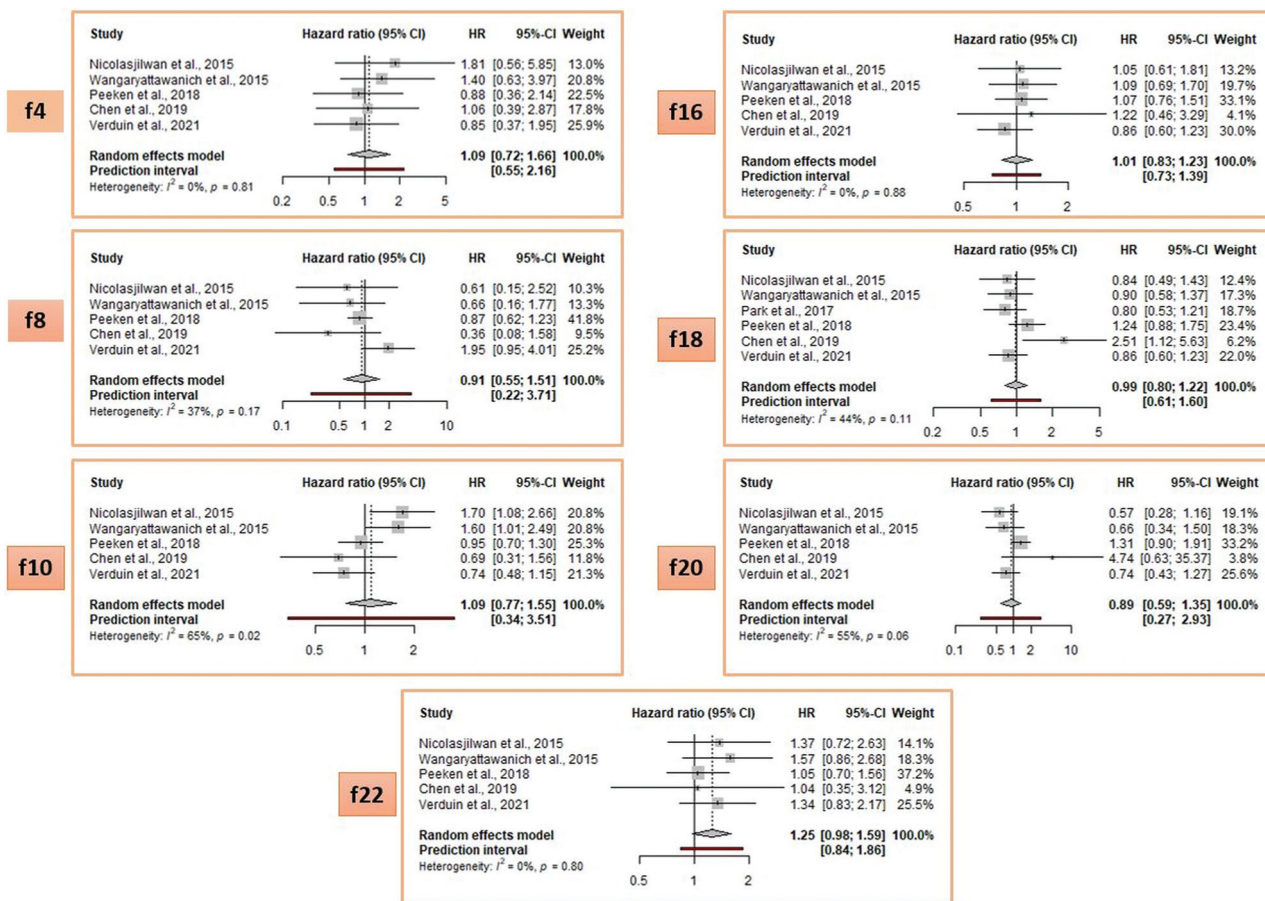


FIG 4. Meta-analysis of OS-predicting studies with a pooled HR ≤ 1.25 . Note that the list of VASARI features and their detailed descriptions, along with respective scoring systems for each feature are provided in the Online Supplemental Data. f indicates VASARI feature; f4, enhancement quality; f8, cysts; f10, TI/FLAIR ratio; f16, hemorrhage; f18, pial invasion; f20, cortical involvement; f22, nonenhancing tumor crosses midline.

and VASARI features or differently structured multivariable models, lowers evidence and precludes translation from the research field to clinical application.

On the basis of the results of our study, several recommendations can be made to improve the generalizability of glioma research using VASARI features. The number and choice of applied VASARI features differed among studies, and some VASARI features were independently modified from their original definition in different studies. The postoperative features 26–28 and feature 15 (edema crossing the midline) were rarely used, making their significance challenging to estimate. Some studies modified VASARI features (using features 4–7, 14, and 17) from the original scoring system, hindering generalizability. Future studies should adhere to the original VASARI scoring system and, if feasible, evaluate all VASARI features to find relevant features for a particular objective. Because the manual extraction of VASARI features is time-consuming. Therefore, automatic extraction technology should be developed, which may also help minimize the heterogeneity described. When a radiologist's input is needed, reducing the entire feature set to only the most promising features, such as multifocality, ependymal invasion, and enhancing tumor crossing the midline for overall survival prediction, as identified in this meta-analysis, may alleviate the

workload. It lies in the nature of the exhaustive VASARI feature set to include collinearity among several features and features that leave a higher chance of interrater disagreement. A reduction of features for research purposes according to reproducibility and predictive value appears to be the proximate consequence.

Our study has several limitations. Studies were heterogeneous in design, and some had a small sample size, with a mean of 124 cases. Some studies used The Cancer Imaging Archive database, leading to partially overlapping cohorts, which we attempted to correct. The meta-analysis was limited due to a high study heterogeneity. Metaregression, sensitivity analysis, and nonreporting bias analysis could not be performed due to the limited number of studies and the diversity in reported predictive models and used feature sets. Studies reporting the performance of VASARI-based OS-predicting models had substantial methodologic heterogeneity, because each study used a different set of VASARI features. Model end points were also heterogeneous. Most of the studies reported a regression model predicting a continuous OS parameter (in days or months); some used classification with different thresholds to predict a more-or-less favorable OS outcome. Other reasons for heterogeneity were using different statistical models (random survival forest, Cox proportional regression, or clustering analysis) and inconsistent use of VASARI feature-

Study	Colen et al. 2014	Nicolasjilwan et al. 2015	Wangaryattawanich et al. 2015	Park et al. 2017	Peeken et al. 2018	Chen et al. 2019	Verduin et al. 2021
Tumor type	glioblastoma	glioblastoma	glioblastoma	glioblastoma	glioblastoma	glioblastoma	glioblastoma
Sample size	104	102	94	108	189	127	188
feature 1		0.54	2.28	0.99	frontal: 1.00, temporal: 0.92, insular: 2.09, parietal: 0.67, occipital: 1.27, brainstem: 0.40	0.63	
feature 2		0.59	0.75		1.52	0.56	
feature 3		0.62	1.45	1.55	speech motor: 0.85, speech: 1.16, receptive: 0.73, motol/vision: 0.70	1.23	1.36
feature 4		1.81	1.40		0.88	1.06	0.85
feature 5		1.89			0.80	1.32	
feature 6		0.86		(34-67%) 1.15 (≥68%) 1.41	0.87	0.56	
feature 7		1.19		(34-67%) 0.83 (≥68%) 0.72	0.94	0.94	
feature 8		0.61	0.66		0.87	0.36	1.95
feature 9		0.58	2.54	2.15	2.50	2.14	1.49
feature 10		1.70	1.60		0.95	0.69	0.75
feature 11		1.82	1.41		1.48	1.13	1.39
feature 12			1.15		1.14	1.13	1.00
feature 13		0.84	1.52		1.80	0.95	1.15
feature 14		0.79		(34-67%) 0.90 (≥68%) 0.68	1.48	2.02	
feature 15					1.21	1.04	
feature 16		1.05	1.09		1.07	1.22	0.86
feature 17		0.76	1.07		facilitated: 0.68, restricted: 0.96, neither/equal: 0.77		
feature 18		0.84	0.90	0.80	1.24	2.51	0.86
feature 19	1.81	1.30	1.88	1.78	2.00		1.54
feature 20		0.57	0.66		1.31	4.74	0.74
feature 21	1.93	0.59	1.82	1.23	1.75	1.61	1.45
feature 22		1.37	1.57		1.05	1.04	1.34
feature 23	3.48	2.74	3.59	1.59	1.36	2.18	1.29
feature 24		1.14	1.34	1.27	2.38	1.15	1.18
feature 25		0.87	0.89		0.92	2.78	
feature 26					0.66		
feature 27					0.73		
feature 28					0.64		
feature 29		1.01	1.02		1.03	1.82	1.07
feature 30		1.00	1.01		1.03	0.65	1.04
FWER correction	Bonferroni	No	No	No	No	No	No

FIG 5. Color table describing HRs of single VASARI features in OS-predicting studies using univariable Cox proportional models. Thresholds of single HRs were defined arbitrarily, and each color represents a different range: blue, HR <0.75; green, HR 0.75 to <0.85; light red, HR 0.85–1.25; orange, HR >1.25–2.5; dark red, HR >2.5. Values highlighted in bold indicate HRs with statistical significance (P value < .05). The list of VASARI features and their detailed descriptions, along with respective scoring systems for each feature, are provided in the Online Supplemental Data.

selection methods for single-data-type models. Statistical data, needed for a meta-analysis, such as confidence intervals or standard errors, were inconsistently reported. Although we attempted a meta-analysis for *IDH*-predictive VASARI models, these models used different combinations of VASARI features, making it difficult to establish an added value of the VASARI model. Additionally, most studies did not provide molecular diagnostics for glioma grading. Although expected for studies published before 2021, this issue was one of the main reasons

most studies were assigned the “medium” QUADAS-2 category for risk-of-bias assessment with applicability concerns in patient selection.

CONCLUSIONS

This meta-analysis reveals that certain features in the VASARI set have promise in predicting OS and *IDH* mutation status. However, the added value of VASARI for predicting tumor grade, true progression, and the status of other mutations remains

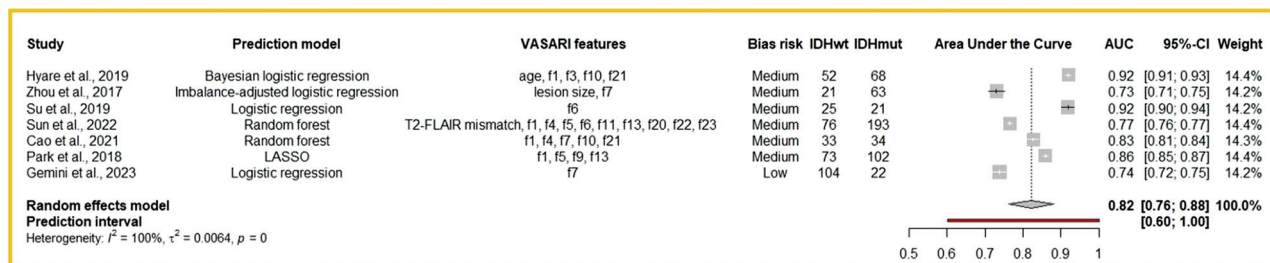


FIG 6. Meta-analysis of glioma IDH mutation status–predicting studies. The list of VASARI features and their detailed descriptions, along with respective scoring systems for each feature, are provided in the Online Supplemental Data. IDHwt indicates IDH wild-type; IDHmut, IDH-mutant.

uncertain, mainly due to insufficiently comparable studies. The discriminatory power of individual VASARI features differs considerably. A core set of promising features, as emerged from this meta-analysis, may be worth prioritizing for scientific evaluation and considered in clinical use, to avoid using the exhaustive list of VASARI features, which is too time-consuming for daily practice.

Disclosure forms provided by the authors are available with the full text and PDF of this article at www.ajnr.org.

REFERENCES

- Ellingson BM, Bendszus M, Boxerman J, et al. Consensus recommendations for a standardized brain tumor imaging protocol in clinical trials. *Neuro Oncol* 2015;17:1188–98 [CrossRef](#)
- Wen PY, van den Bent M, Youssef G, et al. RANO 2.0: update to the Response Assessment in Neuro-Oncology criteria for high- and low-grade gliomas in adults. *J Clin Oncol* 2023;41:5187–99 [CrossRef](#) [Medline](#)
- The Cancer Imaging Archive. VASARI Research Project. <https://wiki.cancerimagingarchive.net/display/Public/VASARI+Research+Project>. Accessed September 26, 2023
- Mazurowski MA, Desjardins A, Malof JM. Imaging descriptors improve the predictive power of survival models for glioblastoma patients. *Neuro Oncol* 2013;15:1389–94 [CrossRef](#) [Medline](#)
- Gutman DA, Cooper LA, Hwang SN, et al. MR imaging predictors of molecular profile and survival: multi-institutional study of the TCGA glioblastoma data set. *Radiology* 2013;267:560–69 [CrossRef](#) [Medline](#)
- Agarwal A, Kumar S, Narang J, et al. Morphologic MRI features, diffusion tensor imaging and radiation dosimetric analysis to differentiate pseudo-progression from early tumor progression. *J Neurooncol* 2013;112:413–20 [CrossRef](#) [Medline](#)
- Jain R, Poisson LM, Gutman D, et al. Outcome prediction in patients with glioblastoma by using imaging, clinical, and genomic biomarkers: focus on the nonenhancing component of the tumor. *Radiology* 2014;272:484–93 [CrossRef](#) [Medline](#)
- Colen RR, Vangel M, Wang J, et al; TCGA Glioma Phenotype Research Group. Imaging genomic mapping of an invasive MRI phenotype predicts patient outcome and metabolic dysfunction: a TCGA glioma phenotype research group project. *BMC Med Genomics* 2014;7:30 [CrossRef](#) [Medline](#)
- Nicolasjlwan M, Hu Y, Yan C, et al; TCGA Glioma Phenotype Research Group. Addition of MR imaging features and genetic biomarkers strengthens glioblastoma survival prediction in TCGA patients. *J Neuroimaging* 2015;42:212–21 [CrossRef](#) [Medline](#)
- Wangaryattawanich P, Hatami M, Wang J, et al. Multicenter imaging outcomes study of The Cancer Genome Atlas glioblastoma patient cohort: imaging predictors of overall and progression-free survival. *Neuro Oncol* 2015;17:1525–37 [CrossRef](#) [Medline](#)
- Rao A, Rao G, Gutman DA, et al; TCGA Glioma Phenotype Research Group. A combinatorial radiographic phenotype may stratify patient survival and be associated with invasion and proliferation characteristics in glioblastoma. *J Neurosurg* 2016;124:1008–17 [CrossRef](#) [Medline](#)
- Yu J, Wang M, Song J, et al. Potential utility of Visually Accessible Rembrandt Images assessment in brain astrocytoma grading. *J Comput Assist Tomogr* 2016;40:301–06 [CrossRef](#) [Medline](#)
- Zhou H, Vallières M, Bai HX, et al. MRI features predict survival and molecular markers in diffuse lower-grade gliomas. *Neuro Oncol* 2017;19:862–70 [CrossRef](#) [Medline](#)
- Lasocki A, Tsui A, Gaillard F, et al. Reliability of noncontrast-enhancing tumor as a biomarker of IDH1 mutation status in glioblastoma. *J Clin Neurosci* 2017;39:170–75 [CrossRef](#) [Medline](#)
- Park M, Lee SK, Chang JH, et al. Elderly patients with newly diagnosed glioblastoma: can preoperative imaging descriptors improve the predictive power of a survival model? *J Neurooncol* 2017;134:423–31 [CrossRef](#) [Medline](#)
- Ersay TF, Keil VC, Hadizadeh DR, et al. New prognostic factor telomerase reverse transcriptase promoter mutation presents without MR imaging biomarkers in primary glioblastoma. *Neuroradiology* 2017;59:1223–31 [CrossRef](#) [Medline](#)
- Park YW, Han K, Ahn SS, et al. Prediction of IDH1-mutation and 1p/19q-codeletion status using preoperative MR imaging phenotypes in lower grade gliomas. *AJNR Am J Neuroradiol* 2018;39:37–42 [CrossRef](#) [Medline](#)
- Peeken JC, Hesse J, Haller B, et al. Semantic imaging features predict disease progression and survival in glioblastoma multiforme patients. *Strahlenther Onkol* 2018;194:580–90 [CrossRef](#) [Medline](#)
- Su CQ, Lu SS, Han QY, et al. Integrating conventional MRI, texture analysis of dynamic contrast-enhanced MRI, and susceptibility weighted imaging for glioma grading. *Acta Radiol* 2019;60:777–87 [CrossRef](#) [Medline](#)
- Su CQ, Lu SS, Zhou MD, et al. Combined texture analysis of diffusion-weighted imaging with conventional MRI for non-invasive assessment of IDH1 mutation in anaplastic gliomas. *Clin Radiol* 2019;74:154–60 [CrossRef](#) [Medline](#)
- Peeken JC, Goldberg T, Pyka T, et al. Combining multimodal imaging and treatment features improves machine learning-based prognostic assessment in patients with glioblastoma multiforme. *Cancer Med* 2019;8:128–36 [CrossRef](#) [Medline](#)
- Lee M, Han K, Ahn SS, et al. The added prognostic value of radiological phenotype combined with clinical features and molecular subtype in anaplastic gliomas. *J Neurooncol* 2019;142:129–38 [CrossRef](#) [Medline](#)
- Ivanidze J, Lum M, Pisapia D, et al. MRI features associated with TERT promoter mutation status in glioblastoma. *J Neuroimaging* 2019;29:357–63 [CrossRef](#) [Medline](#)
- Chen X, Fang M, Dong D, et al. Development and validation of a MRI-based radiomics prognostic classifier in patients with primary glioblastoma multiforme. *Acad Radiol* 2019;26:1292–300 [CrossRef](#) [Medline](#)

25. Hyare H, Rice L, Thust S, et al. **Modelling MR and clinical features in grade II/III astrocytomas to predict IDH mutation status.** *Eur J Radiol* 2019;114:120–27 [CrossRef Medline](#)
26. Park CJ, Han K, Shin H, et al. **MR image phenotypes may add prognostic value to clinical features in IDH wild-type lower-grade gliomas.** *Eur Radiol* 2020;30:3035–45 [CrossRef Medline](#)
27. Lu Y, Patel M, Natarajan K, et al. **Machine learning-based radiomic, clinical and semantic feature analysis for predicting overall survival and MGMT promoter methylation status in patients with glioblastoma.** *Magn Reson Imaging* 2020;74:161–70 [CrossRef Medline](#)
28. Cao M, Suo S, Zhang X, et al. **Qualitative and quantitative MRI analysis in IDH1 genotype prediction of lower-grade gliomas: a machine learning approach.** *Biomed Res Int* 2021;2021:1235314 [CrossRef Medline](#)
29. Verduin M, Primakov S, Compter I, et al. **Prognostic and predictive value of integrated qualitative and quantitative magnetic resonance imaging analysis in glioblastoma.** *Cancers (Basel)* 2021;13:722 [CrossRef Medline](#)
30. Ahn SS, An C, Park YW, et al. **Identification of magnetic resonance imaging features for the prediction of molecular profiles of newly diagnosed glioblastoma.** *J Neurooncol* 2021;154:83–92 [CrossRef Medline](#)
31. Wang J, Yi X, Fu Y, et al. **Preoperative magnetic resonance imaging radiomics for predicting early recurrence of glioblastoma.** *Front Oncol* 2021;11:769188 [CrossRef Medline](#)
32. Sun C, Fan L, Wang W, et al. **Radiomics and qualitative features from multiparametric MRI predict molecular subtypes in patients with lower-grade glioma.** *Front Oncol* 2021;11:756828 [CrossRef Medline](#)
33. Ge X, Wang M, Ma H, et al. **Investigated diagnostic value of synthetic relaxometry, three-dimensional pseudo-continuous arterial spin labelling and diffusion-weighted imaging in the grading of glioma.** *Magn Reson Imaging* 2022;86:20–27 [CrossRef Medline](#)
34. Ruan Z, Mei N, Lu Y, et al. **A comparative and summative study of radiomics-based overall survival prediction in glioblastoma patients.** *J Comput Assist Tomogr* 2022;46:470–79 [CrossRef Medline](#)
35. Sacli-Bilmez B, Firat Z, Topcuoglu OM, et al. **Identifying overall survival in 98 glioblastomas using VASARI features at 3T.** *Clin Imaging* 2023;93:86–92 [CrossRef Medline](#)
36. Sahu A, Patnam NG, Goda JS, et al. **Multiparametric magnetic resonance imaging correlates of isocitrate dehydrogenase mutation in WHO high-grade astrocytomas.** *J Pers Med* 2022;13:72 [CrossRef Medline](#)
37. Gemini L, Tortora M, Giordano P, et al. **Vasari scoring system in discerning between different degrees of glioma and IDH status prediction: a possible machine learning application?** *J Imaging* 2023;9:75 [CrossRef](#)
38. You W, Mao Y, Jiao X, et al. **The combination of radiomics features and VASARI standard to predict glioma grade.** *Front Oncol* 2023;13:1083216 [CrossRef Medline](#)
39. Andersen BM, Miranda C, Hatzoglou V, et al. **Leptomeningeal metastases in glioma: the Memorial Sloan Kettering Cancer Center experience.** *Neurology* 2019;92:e2483–91 [CrossRef Medline](#)
40. Page MJ, McKenzie JE, Bossuyt PM, et al. **The PRISMA 2020 statement: an updated guideline for reporting systematic reviews.** *Int J Surg* 2021;88:105906 [CrossRef Medline](#)
41. Salameh J-P, Bossuyt PM, McGrath TA, et al. **Preferred reporting items for systematic review and meta-analysis of diagnostic test accuracy studies (PRISMA-DTA): explanation, elaboration, and checklist.** *BMJ* 2020;370:m2632 [CrossRef Medline](#)
42. Whiting PF, Rutjes AW, Westwood ME, et al; QUADAS-2 Group. **QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies.** *Ann Intern Med* 2011;155:529–36 [CrossRef Medline](#)
43. Cortes C, Mohri M. **Confidence Intervals for the Area Under the ROC Curve.** In: Saul L, Weiss Y, Bottou L, eds. *Advances in Neural Information Processing Systems*. Vol 17. MIT Press; 2004
44. Higgins J, Thomas J, Miranda JC, et al. **Cochrane Handbook for Systematic Reviews of Interventions.** <https://training.cochrane.org/handbook#how-to-access>. Accessed December 28, 2023
45. Thomas RP, Xu LW, Lober RM, et al. **The incidence and significance of multiple lesions in glioblastoma.** *J Neurooncol* 2013;112:91–97 [CrossRef Medline](#)
46. Lim DA, Cha S, Mayo MC, et al. **Relationship of glioblastoma multiforme to neural stem cell regions predicts invasive and multifocal tumor phenotype.** *Neuro Oncol* 2007;9:424–29 [CrossRef Medline](#)
47. Kurokawa R, Kurokawa M, Baba A, et al. **Major changes in 2021 World Health Organization Classification of Central Nervous System Tumors.** *Radiographics* 2022;42:1474–93 [CrossRef Medline](#)