Neuro-Oncology Advances

4(1), 1–16, 2022 | https://doi.org/10.1093/noajnl/vdac027 | Advance Access date 01 March 2022

Diagnostic performance of DSC perfusion MRI to distinguish tumor progression and treatment-related changes: a systematic review and meta-analysis

Rongwei Fu, Laszlo Szidonya, Ramon F. Barajas Jr., Prakash Ambady, Csanad Varallyay, and Edward A. Neuwelt

Oregon Health & Science University-Portland State University, School of Public Health, Portland, Oregon, USA (R.F.); Department of Medical Informatics & Clinical Epidemiology, Oregon Health & Science University, Portland, Oregon, USA (R.F.); Department of Radiology, Oregon Health & Science University, Portland, Oregon, USA (L.S., R.F.B.); Neuro-Oncology Program, Oregon Health & Science University, Portland, Oregon, USA (L.S., P.A., E.A.N.); Heart and Vascular Center, Diagnostic Radiology, Semmelweis University, Budapest, Hungary (L.S.); Advanced Imaging Research Center, Oregon Health & Science University, Portland, Oregon, USA (R.F.B.); Knight Cancer Institute Translational Oncology Program, Oregon Health & Science University, Portland, Oregon, USA (R.F.B.); Department of Radiology, EPIC Imaging, Portland, Oregon, USA (C.V.); Department of Neurosurgery, Oregon Health and Sciences University, Portland, Oregon, USA (E.A.N.); Office of Research and Development, Department of Veterans Affairs Medical Center, Portland, Oregon, USA (E.A.N.)

Corresponding Author: Edward A. Neuwelt, MD, Oregon Health and Science University, Mail code L 603, 3181 SW Sam Jackson Park Road, Portland, OR 97239, USA (neuwelte@ohsu.edu).

Abstract

Background. In patients with high-grade glioma (HGG), true disease progression and treatment-related changes often appear similar on magnetic resonance imaging (MRI), making it challenging to evaluate therapeutic response. Dynamic susceptibility contrast (DSC) MRI has been extensively studied to differentiate between disease progression and treatment-related changes. This systematic review evaluated and synthesized the evidence for using DSC MRI to distinguish true progression from treatment-related changes.

Methods. We searched Ovid MEDLINE and the Ovid MEDLINE in-process file (January 2005–October 2019) and the reference lists. Studies on test performance of DSC MRI using relative cerebral blood volume in HGG patients were included. One investigator abstracted data, and a second investigator confirmed them; two investigators independently assessed study quality. Meta-analyses were conducted to quantitatively synthesize area under the receiver operating curve (AUROC), sensitivity, and specificity.

Results. We screened 1177 citations and included 28 studies with 638 patients with true tumor progression, and 430 patients with treatment-related changes. Nineteen studies reported AUROC and the combined AUROC is 0.85 (95% CI, 0.81–0.90). All studies contributed data for sensitivity and specificity, and the pooled sensitivity and specificity are 0.84 (95% CI, 0.80–0.88), and 0.78 (95% CI, 0.72–0.83). Extensive subgroup analyses based on study, treatment, and imaging characteristics generally showed similar results.

Conclusions. There is moderate strength of evidence that relative cerebral blood volume obtained from DSC imaging demonstrated "excellent" ability to discriminate true tumor progression from treatment-related changes, with robust sensitivity and specificity.

Keywords:

diagnostic performance | dynamic susceptibility contrast (DSC) MRI | high-grade glioma | meta-analysis | treatment-related changes

In patients with high-grade glioma (HGG), imaging evaluation of therapeutic response following chemoradiation therapy based on the Stupp protocol[1] can be challenging. True disease progression and treatment-related changes often appear similar on T1-weighted gadolinium-enhanced magnetic resonance imaging (MRI).[2,3] The development of treatment-related changes is a hallmark of effective therapy. Therefore, its misdiagnosis as true progression and subsequent cancelation of effective therapy may have adverse implications for patient outcomes. On the other hand, in the case of true progression, the patient should be offered a second-line therapy. It follows that the reliable differentiation between true disease progression and treatment-related changes is of the utmost clinical importance. In the literature, treatment-related changes are often described as radiation necrosis, radiation injury, or pseudoprogression. Radiation necrosis is histopathologically characterized as chemoradiotherapy-induced injury to the central nervous system with fibrinoid necrosis of blood vessel walls and adjacent perivascular parenchymal coagulative necrosis, and pseudoprogression refers to new or increasing contrast enhancement that eventually subsides without any change in therapy. However, such terms are not well distinguished and often used interchangeably in the published HGG articles. Based on a recent systematic review and meta-analysis, the incidence of pseudoprogression is 36% in patients with HGG after standard of care chemoradiation.[4] An increase in the incidence of pseudoprogression has been observed with the introduction of immunotherapy. This diagnostic uncertainty is not only emotionally draining for patients and their families, but also can halt administration of effective therapy or delay enrollment into potentially beneficial clinical trials.

Tumor angiogenesis mediated though various genetic and pro-angiogenic growth factors, such as vascular endothelial growth factors, is a characteristic feature of HGGs. This poorly regulated tumor neo-angiogenesis results in disrupted endothelial tight junctions with increased fenestrations. This results in a leaky blood brain barrier and is associated with the heterogeneous pattern of contrast enhancement seen on MRI. The presence or absence of florid microvascular proliferation is an integral part of the WHO grading system used in HGG.[5] However, currently used Response Assessment in Neuro-Oncology (RANO) criteria, updated to include brain metastases (RANO-BM)[6] and immunotherapy (iRANO) focus on changes in apparent tumor size over time and do not include functional changes such as vascularity or inflammation.

Although histopathology remains an important tool, a reliable technique to assess tumor vascularity can help distinguish progressive tumors that are expected to have a higher degree of vascularity from treatment-related changes. Noninvasive evaluation of tumor vascularity based on perfusion MRI techniques has been utilized as an approach to improve the clinical diagnostic utility of clinical response assessment of HGG. Dynamic Susceptibility Contrast Enhanced (DSC) T2*-weighted MRI is the most extensively studied perfusion technique.[7] DSC MRI allows for the estimation of relative cerebral blood volume within brain tumors; and it has demonstrated potential for determining prognosis, predicting therapeutic response, and assessing early treatment response of gliomas.[8] Several recent systematic reviews and meta-analyses have demonstrated the diagnostic performance of DSC MRI.[9–12] However, each reviews only included a small number of studies, and a paucity of reporting critically important imaging sequence parameters such as flip angle, the use of "contrast preload" or leakage correction in these pooled analyses hinders interoperability and clinical adoption. To this end, the objective of this review is to systematically evaluate and synthesize the evidence for using DSC MRI to distinguish true progression from treatment-related changes (including pseudoprogression (PsP) and radiation necrosis (RN)) in patients with HGGs after tumor resection and standard of care therapy, with specific considerations on how study and clinical characteristics, and imaging parameters impact the diagnostic performance of DSC MRI.

Materials and Methods

This review followed the systematic review methodology and procedures developed by the Agency for Healthcare Research and Quality (https://www.ncbi.nlm.nih.gov/books/NBK98241/) that were in accordance with current guidance for systematic reviews.[13] The reporting of this review followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses statement.[14]

Literature Search

A research librarian searched Ovid MEDLINE and the Ovid MEDLINE in-process file. The dates of coverage for the searches were January 2015 through October 31, 2019 to cover the published literature after the searches conducted for the publication of the systematic review and meta-analysis by Patel et al. (2017).[10] Published literature before 2015 was retrieved based on the search done by Patel et al. and other identified systematic reviews and meta-analyses. Search details are provided in the online Supplementary Materials. Additional references were obtained by searching reference lists.

Eligibility Criteria and Study Selection

The inclusion and exclusion criteria and selection of studies followed the PICOTS format:

- **Participants/Population:** Eligible Patients were those with HGG (grade III and IV) that had been treated with surgery and chemotherapy or chemoradiotherapy, or other accepted treatments during the time of the study, and received DSC MRI perfusion imaging. If a study enrolled both high- and low-grade gliomas and grade II glioma patients accounted for < 25% of the patient population, it remained eligible to be included.
- **Intervention/Exposure:** The index test was based on DSC perfusion MRI. Relative cerebral blood volume (rCBV), i.e., CBV normalized to the normal appearing white matter, was needed to be quantitatively measured at the time of lesion enhancement.

- **Comparator/Control:** Reference standard were histopathology or repeated image/clinical follow-ups.
- **Outcomes:** Disease progression (vs. treatment-related changes).
- **Timing:** All follow-up intervals were considered.
- **Setting:** All settings were considered.

We considered all studies that provided valid data to evaluate sensitivity, specificity, and the area under the receiver operating curve (AUROC) regardless of study design. Systematic reviews were considered to search for relevant studies.

Not considered were animal studies, studies that did not report or did not report adequate data to calculate sensitivity, specificity or AUROC, or studies where rCBV was only visually or qualitatively assessed. Results published *only* in abstract form were not included because inadequate details were available for quality and risk of bias assessment and it is not clear whether the results would be valid. Only English language articles were included.

Two reviewers (authors RF and LS) assessed titles and abstracts for citations identified through literature searches, and full-text articles of potentially relevant citations were retrieved and assessed for inclusion. We also retrieved and assessed full-text papers that evaluated DSC perfusion MRI from the prior systematic reviews for inclusion. All full-text papers were dually assessed and disagreements were resolved by consensus.

Data Abstraction

The data were abstracted from included studies by one reviewer and checked by a second reviewer. Disagreements were resolved by consensus. Evidence table was constructed to show the population and study characteristics, including study design, patient demographics, initial tumor histology, treatment regimens, reference standard, time to diagnosis of postsurgery enhancing lesion, MRI acquisition details and parameters, sample size, and diagnostic performance data.

Risk of Bias/Quality Assessment

All studies were independently assessed by two reviewers (authors RF and LS). Risk of bias and applicability were assessed using the Quality Assessment of Diagnostic Accuracy Studies -2 (QUADAS-2) tool and the approach recommended in the AHRQ Methods Guide for Medical Test Reviews.[13,15]

Quantitative Data Synthesis

We assessed the clinical and methodological diversity of the included studies (Table 1) in addition to statistical heterogeneity in order to determine appropriateness of pooling data. We conducted meta-analysis only if the included studies were deemed similar enough to produce a meaningful combined estimate and there were adequate data from included studies.

We performed random effects meta-analysis to calculate the combined AUROC using the profile likelihood (PL) method.[44] When a study only reported the point estimate of AUROC without providing a 95% CI or a standard error, we calculated the standard error using the numbers of patients with true tumor progression and treatment-related changes.[45] The following classification was used to guide the interpretation of AUROC: ≤0.50 no discriminative ability; 0.70–0.80 acceptable discriminative ability; 0.80–0.90 excellent discriminative ability; ≥0.90 outstanding discriminative ability.[46] To obtain combined sensitivity and specificity, we used a bivariate logistic mixed effects model, incorporating the correlation between sensitivity and specificity. The model assumed random effects across studies for sensitivity and specificity, and heterogeneity among the studies was measured based on the random effect variance. We also assessed statistical heterogeneity using the standard χ^2 test and I^2 statistic.[47] We calculated positive likelihood ratio (LR+) and negative likelihood ratio (LR–) using the summarized sensitivity and specificity.[48,49]

Included papers used different measures of relative cerebral blood volume (rCBV). We grouped the rCBV measures into two categories:

1) Mean rCBV: the studies generally used the mean rCBV of the sampled or entire enhancing lesion.
2) Maximum rCBV: the studies used the maximum, or 90th percentile of rCBV of the enhancing lesion, or the mean of the hot-spot areas.

In the primary analysis, we combined studies using either Mean or Maximum rCBVs given the similar performance of the two measures. If a study reported results using both measures, results from the Mean rCBV were used in the primary analysis and results from the Maximum rCBV were used in the sensitivity analysis. Subgroup analyses were conducted by combining studies using Mean rCBV or Maximum rCBV only. We combined all eligible studies in the primary analysis regardless of the variation in acquisition and postprocessing parameters since they were all reasonable to use in practice, and such analysis produced one overall combined estimate. Nevertheless, we conducted extensive subgroup analyses to qualitatively evaluate how the study characteristics (e.g., study design, country, funding source, reference standard) and clinical and imaging characteristics including acquisition and postprocessing parameters (e.g., time of image evaluation, histology, reference standards, sequence and flip angle, leakage correction) impact the diagnostic performance. Time of image evaluation was categorized as early if rCBV data were acquired within 6 months of chemoradiation therapy, late if after 6 months, and both if rCBV data were acquired within 6 months in some patients and after 6 months in others. Treatment-related changes observed during early time period were more likely associated with pseudoprogression and late treatment-related changes were more likely associated with radiation necrosis. For gradient echo sequence, flip angle was categorized as low if ≤40, intermediate if 60; and high if 90. If the flip angle was between 35 and 90, it was categorized as "mixed". In terms of chemoradiation

Table 1. Study Characteristics of the Included Studies

Author, Year	Country	Study Design	Mean Age (Yrs, SD or range)	Gender (M/F)	Histology	Chemoradiation Therapy	Time of Image Evaluation (range, or mean±SD)**	Reference Standard	Imaging Field Strength	Sequence Flip Angle	Preload	Leakage Correction (Software)	Imaging Parameter	Cutoff	Optimal Performance	*N* of Patients /Lesions***	Industry Funding
Alexiou GA, 2014[16]	Greece	P	61.5 (11.1)	21/9	HGG (III, 3; IV, 27)	CCRT	Both (3–24 mos)	HP or CRFU	1.5	GRE 40	NR	Yes (LUPE)	Max rCBV	2.2	Yes	30	No, NF
Baek HJ, 2012[17]	SK	R	50.6 (19–83)	46/33	GBM	CCRT	Early (4 wks)	HP or MRI FU	3	GRE 35	NR	Yes (NI, Matlab)	Max rCBV	3.1	Yes	79	No, GF
Barajas RF, 2009[18]	USA	R	54.2 (10.2)	33/24	GBM	CRT	Late (12.8 ± 7.4 mos)	HP or CRFU	1.5	GRE 35	No	No (FuncTool)	Mean rCBV	1.68	Yes	57/66	No, GF
Blasel S, 2016[19]	Germany	R	56.3 (33–82)	23/21	GBM	CRT	Both (1.2–136.9 mos)	HP or MRI FU	3	GRE 90	Yes	NR (Syngo)	Mean/ Max rCBV	2.2/2.6	NR	44	NR
Bobek-Billewicz B, 2010[20]	Poland	R	39.7 (23–68)	3/4	HGG (III, 5; IV, 2)	CCRT (2) RT (5)	Both (3–70 mos)	HP or RFU	1.5 or 3	SE (1.5T) GRE 7 (3T)	NR	NR (VS)	Mean/ Max rCBV	1.2/2	Yes	7/10	NR
Boxerman JL. 2017[21]	USA	R	58 (38–70)	5/4	H (III, 3; IV 6)	CCRT	NR	HP or CRFU	1.5 or 3	GRE 90	Yes	Yes (IB Neuro)	Mean rCBV	2.4	NR	9/19	Yes, Partly
Cha J, 2014[22]	SK	R	49 (24–70)	18/17	GBM	CCRT	Early (< 6 mos)	HP or RFU	3	GRE 40	No	Yes (NI)	Mean rCBV	1.8	NR	35	NR
D'Souza MM, 2014*[23]	India	R	44.9 (22–61)	14/4	HGG (III, 8; IV,10)	CCRT (13) CRT (5)	Late (7–19 mos)	HP or CRFU	3	GRE 90	NR	NR (NR)	Max rCBV	1.8	Yes	18	No, No COI
Dandois V, 2010*[24]	Belgium	R	51 (25–74)	16/12	HGG (III, 9; IV,7)	CCRT (18) RT (10)	Both (3–40 mos)	HP	1.5	GRE NR	Yes	No (FuncTool)	Max rCBV	3.78	Yes	16	NR, No COI
Di Costanzo A, 2014[25]	Italy	R	62.5 (38–74)	18/11	GBM	CCRT	Both (1–24 mos)	RFU	3	GRE 90	Yes	No (FuncTool)	Mean rCBV	NR	Yes	29	NR, No COI
Gasparetto EL, 2009[26]	USA	R	53.8 (42–71)	8/4	GBM	CRT	Both (3–91 mos)	HP	1.5	SE	Yes	No (FuncTool)	Mean rCBV	1.8	Yes	12	NR, NO FR
Hu LS, 2010[27]	USA	R	46.9 (31–62)	10/1	HGG (III, 3; IV 8)	CRT	Both (2–28.5 mos)	HP	3	GRE 60	Yes	Yes (Matlab)	Mean rCBV	1.15	Yes	11/36	No, GF
Hu XT, 2011[28]	USA	R	NR	NR	GBM	RT	NR	MRI Follow-up	NR	SE	NR	NR (NR)	Mean CBV	1.14	Yes	16	No, GF
Kim HS, 2010[29]	SK	R	48.2 (18–78)	22/17	GBM	CRT/RT	NR	HP	3	GRE 35	NR	Yes (NI)	Max rCBV	2.6	Yes	39	NR, NO FR
Kim YH, 2010[30]	SK	R	46.1 (31–66)	8/2	HGG (III 5, IV 5)	CRT (not TMZ)	Both (3–91 mos)	HP or CRFU	1.5 or 3	GRE NR	NR	NR (NR)	Mean CBV	3.69	Yes	10	NR, No COI
Kim TH, 2017[31]	SK	R	52.9 (11.6)	30/21	HGG (III, 29; IV, 22)	RT	Both (23.5 ± 22.4 mos)	HP or RFU	3	GRE 35-90	NR	Yes (NI)	Mean rCBV/ rCBV90	1.07/ 2.07	Yes	51	No, GF

Table 1. Continued

Author, Year	Country	Study Design	Mean Age (Yrs, SD or range)	Gender (M/F)	Histology	Chemoradiation Therapy	Time of Image Evaluation (range, or mean±SD)**	Reference Standard	Imaging Field Strength	Sequence Flip Angle	Preload	Leakage Correction (Software)	Imaging Parameter	Cutoff	Optimal Performance	*N* of Patients /Lesions***	Industry Funding
Kim JY, 2019[32]	SK	R	58 (34–83)	38/23	GBM	CCRT	Early (≤ 12 wks)	HP or CRFU	3	GRE 35	Yes	Yes (NI)	Mean/ Max rCBV	NR/NR	Yes	61	No, GF
Kong DS, 2011*[33]	SK	P	50 (25–74)	49/41	GBM	CCRT	Late (6.2–40 mos)	RFU	3	GRE 40	NR	No (Viewforum)	Max rCBV	1.49	Yes	59	No, GF
Matsusue E, 2010[34]	USA	R	50.4 (30–64)	5/3	HGG (III, 3; IV 5)	CRT	Late (4–156 mos)	HP or CRFU	3	3D GRE7	NR	No (EWS)	Max rCBV	2.1	Yes	8	NR, COI
Ozsunar Y, 2010*[35]	USA	R	42 (11)	8/22	Glioma (II 7, III 9 and IV 19)	CT/PBT	Late (>6 mos)	HP or RFU	1.5	SE	Yes	Yes (NR)	Max rCBV	1.3	Yes	NR/19	No, GF
Park JE, 2015[36]	SK	R	45.5 (9.7)	80/82	GBM	CCRT	Early (≤ 12 wks)	HP or CRFU	3	GRE 35	Yes	Yes (NI)	rCBV90	3.38	Yes	162	No, GF
Prager AJ, 2015*[37]	USA	R	54.9 (22.6–79.4)	51/17	HGG (III 13, IV 55)	CCRT	Both (0.4–40 mos)	HP	1.5 and 3	GRE 60	No	Yes (FuncTool)	Mean/ Max rCBV	1.27/ 1.74	Yes	41	No, GF
Seeger A, 2013[38]	Germany	R	53.6 (13.6)	24/16	HGG (III NR, IV NR)	CCRT	NR	CRFU	1.5	GRE 90	Yes	Yes (Leonardo)	Mean CBV	2.15	Yes	40	NR
Shin KE, 2014[39]	SK	R	54.5 (11.9)	17/14	Gliomas (II 4; III, 8; IV 20)	CCRT/RT	Both (1–20 mos)	HP or CRFU	3	GRE 90	Yes	Yes (Siemens)	Max rCBV	2.33	Yes	31	NR
Sugahara T, 2000[40]	Japan	P	54.3 (38–78)	7/4	HGG (III, 6; IV 5)	CRT (not TMZ)	Late (6–94mos)	HP or CRFU	1.5	GRE NR	Yes	No (NR)	Mean CBV	1.6	Yes	11	NR
Wang S, 2016[41]	USA	R	55.7 (23–80)	27/14	GBM	CCRT	Early (≤6 mos)	HP	3	GRE NR	Yes	Yes (NI)	rCBV90	2.77	Yes	41	No, GF
Young RJ, 2013[42]	USA	R	58 (9–84)	14/6	GBM	CCRT	Early (0.5–6.3 mos)	HP or CRFU	1.5 and 3	GRE 60	No	No (FuncTool)	Mean CBV	2.4	Yes	20	NR
Zakhari N, 2019*[43]	Canada	P	54.1 (13.3)	43/23	HGG (III, 16; IV 47)	CCRT	Late (RN, 12.2 mos TR, 7.2mos)	HP or CRFU	3	GRE 90	Yes	Yes (Oleasphere)	Max rCBV	2.74	NR	63	No, GF

*Paper reported data for all patients. For our analysis, only a subset of patients were included.
**Time of imaging evaluation: denotes the time of rCBV acquisition from completion of CRT. Early: Data acquired up to 6 months from completion of CRT; Late: Data acquired 6 months after completion of CRT; Both: Data acquired during early time period in some patients and during late time period in others.
***The numbers of the patients and lesions are the same if only one number is reported.
CCRT, Concurrent chemoradiation therapy; COI, Conflict of interest; CRFU, Clinicoradiologic follow-up; CRT, Chemoradiation therapy; EWS, Philips Extended MR Workstation; FR, Financial relationship; GBM, Glioblastoma multiforme; GF, Grant funded; GRE, Gradient echo; HGG, High-grade glioma; HP, Histopathology; IB Neuro, Imaging Biometrics Neuro perfusion; LUPE, The Lund Perfusion Program; NF, Not funded; NI, Nordic Neurolabs Nordic Ice; NR, Not reported; P, Prospective; R, Retrospective; rCBV, relative cerebral blood volume; RFU, Radiologic follow-up; RN, radiation necrosis; RT, Radiation therapy; TR, Tumor recurrence; SE, Spin echo; SK, South Korean; VS, Vendor software.

therapy, if all or the majority of patients in a study received concurrent chemoradiation therapy (CCRT), it was categorized as CCRT; otherwise, non-CCRT.

All analyses were conducted using Stata/IC 16.1 (StataCorp LP, College Station, TX).

Qualitative Synthesis and Grading the Evidence

Qualitative synthesis considered the critical appraisal of the body of evidence, putting the results of any quantitative synthesis into perspective. We used a GRADE approach to grading the strength of evidence, which considers risk of bias, consistency of study results, directness, and precision of estimates.[50]

Results

Included Studies and Study Characteristics

We screened 1177 citations, 67 full-text articles, and identified five new studies[21,32,37,41,43] and six systematic reviews.[9–12,51,52] From the six systematic reviews and reference searches, we further retrieved and screened 45

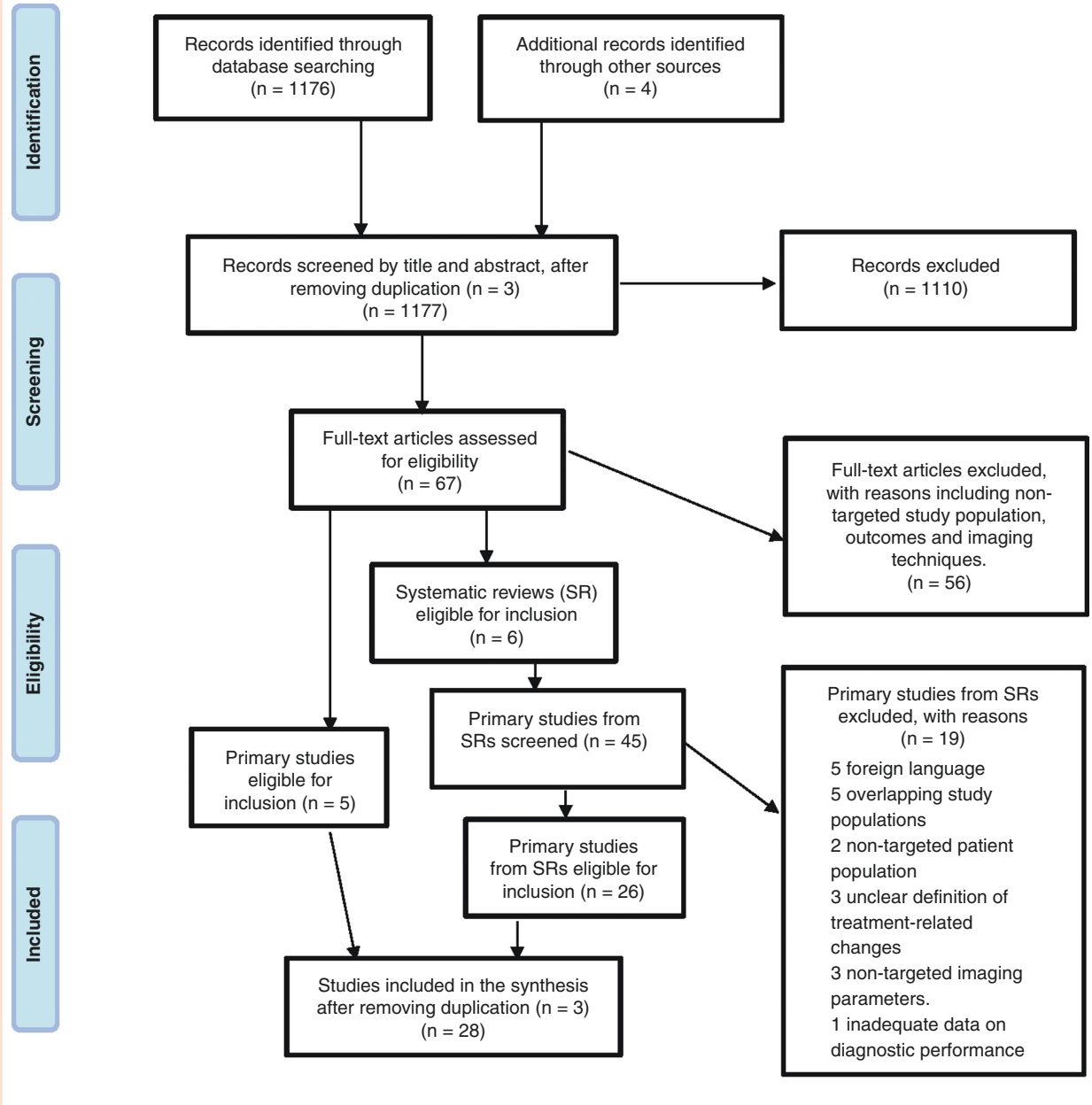

Figure 1. PRISMA FLOW diagram of the article selection process.

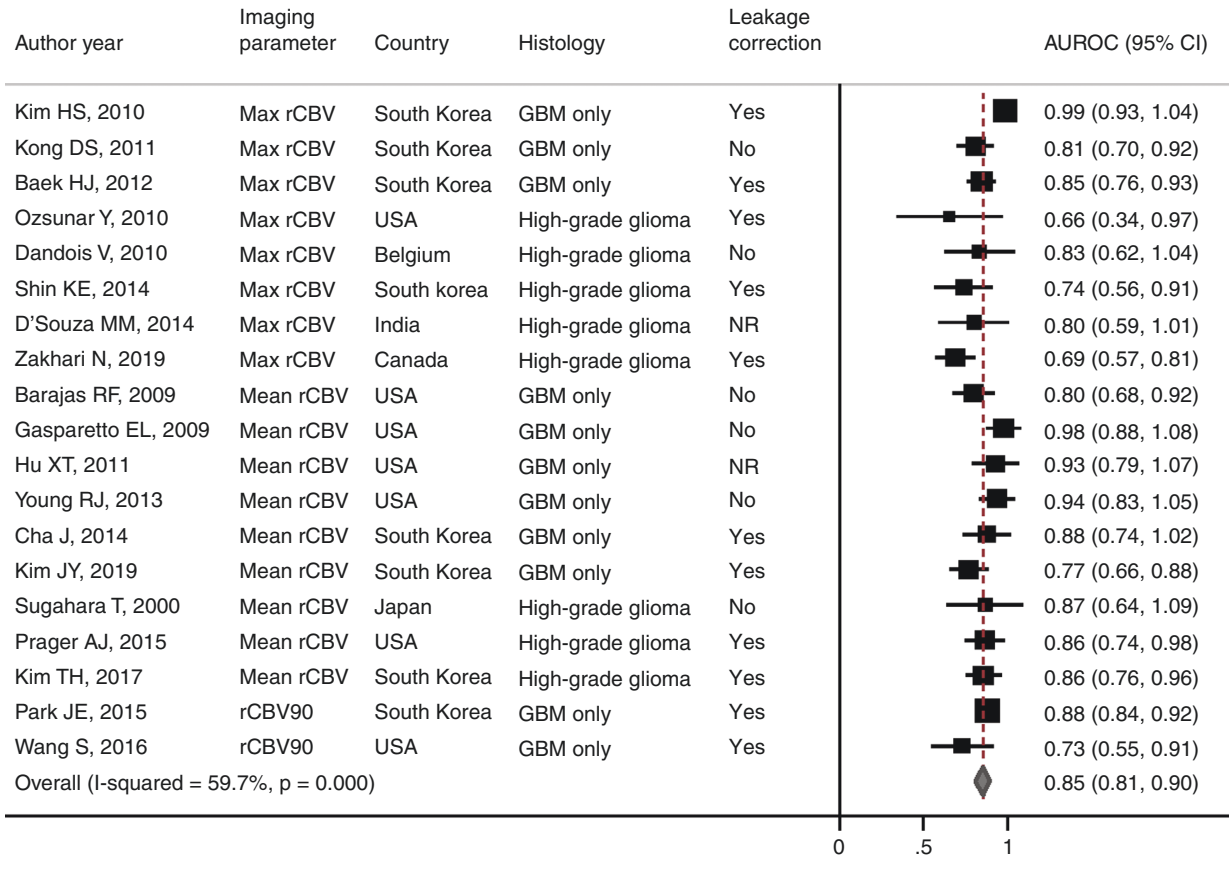

Figure 2. Combined area under ROC across included studies.

full-text papers with 26 meeting our inclusion criteria. After excluding duplicated papers, we included 28 studies (Figure 1).

The included studies are predominantly retrospective (n = 24), grant-funded or no conflict of interest reported (n = 19) with ten studies conducted in the US. Fifteen studies included high-grade glioma patients and thirteen studies only included GBM patients. All studies included newly diagnosed patients, except for one study,[19] which included 20 recurrent patients with second, third, or fourth line treatments. Five studies reported sensitivity and specificity based on both Mean and Maximum rCBV. Details about the included studies are provided in Table 1.

Pooled AUROC

Nineteen studies (n = 842 subjects) reported AUROC and the combined AUROC is 0.85 (95% CI, 0.81–0.90), indicating "excellent" ability to discriminate true tumor progression from treatment-related changes (Figure 2). There is moderate heterogeneity among the included studies (I^2 = 59.7%, P < .001), with results relatively consistent across studies.

Nine studies reported AUROC based on Mean rCBV with a combined AUROC of 0.88 (95% CI, 0.83–0.93) (Supplementary Figure 1), while the combined AUROC from the thirteen studies based on Max rCBV is 0.83 (95% CI, 0.77–0.88) (Supplementary Figure 2). Both estimates remain in the range of "excellent" discriminative ability. Subgroup analyses by country, industry funding, histology, reference standard, and imaging parameters do not reveal important differences in AUROC by these factors (Table 2). Among the 19 studies, 11 studies reported to use 3T scanners, 9 studies reported to use contrast pre-load bolus and leakage correction and no studies reported to use a high flip angle. A notable difference was observed between studies of early image evaluation (6 studies; AUROC 0.87, 95% CI 0.81–0.90) and late image evaluation (6 studies, AUROC 0.77, 95% CI 0.71–0.84), which suggests that rCBV might have "excellent" and higher discriminative ability in distinguishing early treatment-related changes (pseudoprogression) from true progression than distinguishing late treatment-related changes (radiation necrosis) from true progression. However, the number of studies was small in each category and precluded any definite conclusion. A similar difference was also observed between prospective (3 studies; AUROC 0.77, 95% CI 0.66–0.89) and retrospective

Table 2. Subgroup Analysis of AUROC, by Study and Imaging Characteristics

	N of studies	*N* of patients	Pooled AUROC (95% CI)	I^2, *P*-value
All studies	19	842	0.85 (0.81, 0.90)	59.7%, $P < .01$
All studies, using Max rCBV first	19	842	0.85 (0.81, 0.90)	59.9%, $P < .01$
Imaging parameter: Mean rCBV	9	648	0.88 (0.83, 0.93)	23.1%, $P = .23$
Imaging parameter: Max rCBV	13	774	0.83 (0.77, 0.88)	66.1%, $P < .01$
Country				
USA	7	215	0.88 (0.79, 0.95)	32.4%, $P = .08$
Non-USA	12	627	0.84 (0.78, 0.89)	65.3%, $P < .01$
Study Design				
Prospective	3	135	0.77 (0.66, 0.89)	0.8%, $P = .23$
Retrospective	16	707	0.87 (0.82, 0.91)	53.3%, $P < .01$
Industry Funded				
No, or No COI	15	745	0.85 (0.80, 0.90)	65.5%, $P < .01$
NR	4	97	0.88 (0.77, 0.96)	0.0%, $P = .31$
Histology				
GBM only	11	590	0.88 (0.83, 0.93)	63.0%, $P < .01$
High-Grade Glioma	8	252	0.80 (0.73, 0.87)	13.0%, $P = .40$
Timing of Image Evaluation				
Early	6	398	0.87 (0.81, 0.90)	0.0%, $P = .20$
Late	6	238	0.77 (0.71, 0.84)	0.0%, $P = .59$
Both	5	151	0.88 (0.78, 0.95)	24.8%, $P = .16$
NR	2	55	0.98 (0.89, 1.00)	0.0%, $P = .45$
Reference Standard				
Histopathology in < 75% of pts	11	575	0.85(0.80, 0.89)	27.6%, $P = .12$
Histopathology in ≥ 75% of pts	8	267	0.86(0.75, 0.94)	64.8%, $P < .01$
Risk in Patient Selection				
Low	7	502	0.83(0.77, 0.88)	39.6%, $P = .06$
Unclear	12	340	0.88 (0.81, 0.93)	51.2%, $P < .01$
Risk in Index Test				
Low	7	385	0.86(0.76, 0.93)	67.2%, $P < .01$
Unclear	12	457	0.85(0.79, 0.91)	49.7%, $P < .01$
Chemoradiation Therapy				
CCRT	12	628	0.83(0.78, 0.87)	36.4%, $P = .08$
Non-CCRT	7	214	0.91(0.82, 0.97)	48.2%, $P = .03$
Imaging Field Strength				
1.5T	5	124	0.87 (0.74, 0.96)	33.8%, $P = .12$
3.0T	11	641	0.84 (0.77, 0.89)	70.1%, $P < .01$
1.5T or 3.0T	2	61	0.90 (0.80, 1.00)	0.0%, $P = .36$
NR	1	16	0.93 (0.79, 1.00)	NA
Preload				
Yes	9	418	0.81 (0.73, 0.89)	61.7%, $P < .01$
No	4	162	0.87 (0.80, 0.94)	0.0%, $P = .43$
NR	6	262	0.89 (0.81, 0.95)	55.6%, $P < .01$
Leakage Correction				
Yes	9	624	0.84 (0.77, 0.89)	70.7%, $P < .01$
No	4	184	0.88 (0.80, 0.95)	33.2%, $P = .15$
NR	6	34	0.89 (0.71, 1.00)	0.0%, $P = .31$
Sequence & Flip Angle				

Table 2. Continued

	N of studies	*N* of patients	Pooled AUROC (95% CI)	I^2, *P*-value
GRE				
Low	7	501	0.87 (0.80, 0.92)	69.1%, $P < .01$
Intermediate or Mixed	3	112	0.89 (0.82, 0.96)	0.0%, $P = .52$
Low or Intermediate	9	562	0.87 (0.82, 0.92)	61.5%, $P < .01$

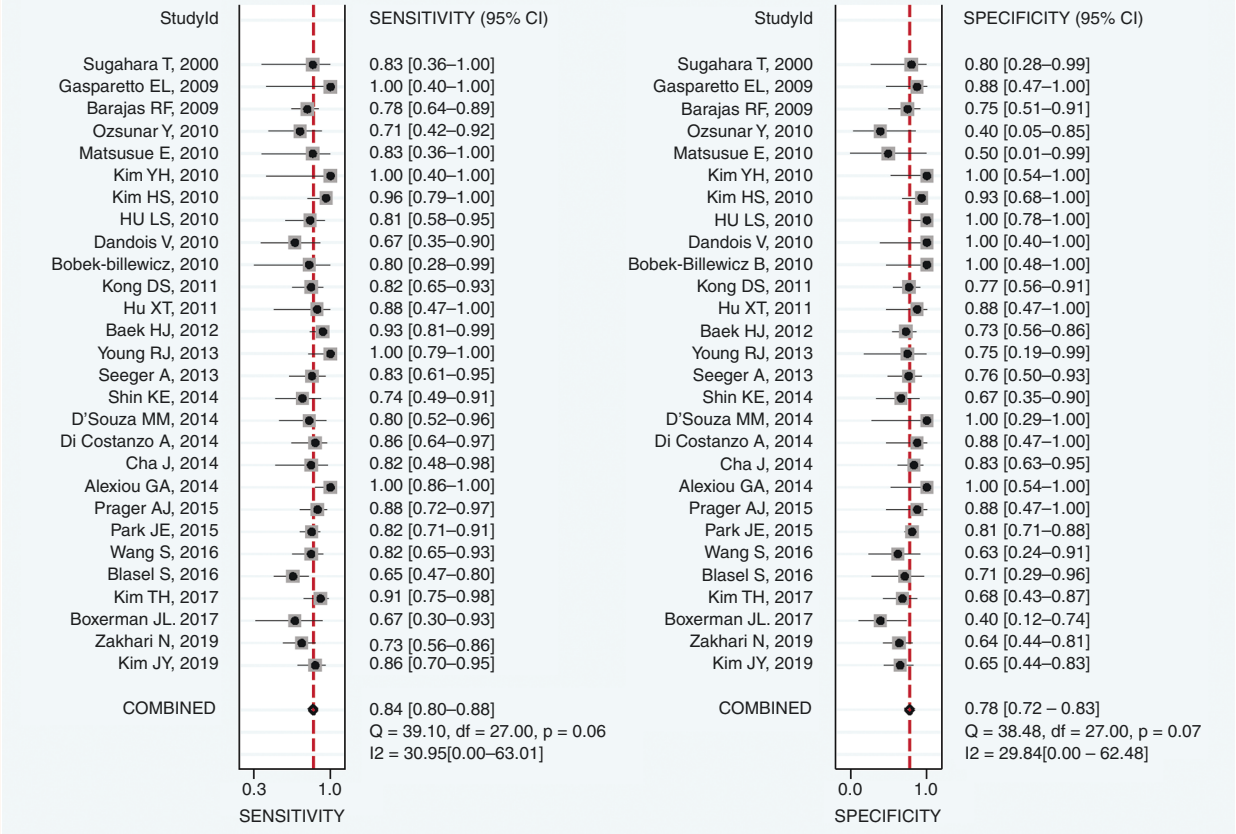

Figure 3. Combined Sensitivity and Specificity across included studies.

studies (16 studies, AUROC 0.87, 95% CI 0.82–0.91), though the small number of prospective studies limited any reliable interpretation.

Pooled Sensitivity and Specificity

To calculate sensitivity and specificity, the included studies used different cutoff points and most studies sought to optimize the diagnostic performance. Given that 1) the choice of the region of interest for DSC imaging and other imaging parameters impacted the rCBV acquired in each study, therefore the cutoffs from the different studies were expected to be different; 2) there is no commonly accepted cutoff points, we deemed it reasonable to combine the studies with different cutoff points, and the combined estimates provided a clinical meaningful estimate to demonstrate the general performance of rCBV to distinguish true tumor progression from treatment-related changes.

Twenty-eight studies reported sensitivity and specificity and these studies included 638 patients with true tumor progression and 430 patients with treatment-related changes. The pooled sensitivity and specificity are 0.84 (95% CI, 0.80–0.88), and 0.78 (95% CI, 0.72–0.83), indicating robust performance to diagnose true tumor progression (from treatment-related changes) (Figure 3). Results were almost identical when excluding the one study that included 20 recurrent patients[19] (sensitivity, 0.84 (95% CI, 0.81–0.88); specificity, 0.78 (95% CI, 0.72–0.83)). When only

combining studies that sought optimal diagnostic performance, the pooled sensitivity and specificity are 0.86 (0.82–0.89), and 0.78 (0.74–0.82).

Results based on Mean rCBV and Max rCBV were similar. Eighteen studies reported sensitivity and specificity based on Mean rCBV. The pooled sensitivity and specificity are 0.81 (95% CI, 0.75–0.86), and 0.78 (95% CI, 0.71–0.84) (Supplementary Figure 3). Seventeen studies reported sensitivity and specificity based on Max rCBV. The pooled sensitivity and specificity are 0.85 (95% CI, 0.80–0.88), and 0.75 (95% CI, 0.69–0.80) (Supplementary Figure 4). Again, subgroup analyses by country, study design, industry funding, histology, reference standard, and imaging parameters generally do not reveal important differences in sensitivity and specificity by these factors (Table 3). Some notable differences were observed between studies of early vs. late image evaluation, and among studies using different flip angles with gradient echo sequence. Studies of early image evaluation yielded higher sensitivity (0.87, 95% CI 0.80–0.91) and specificity (0.76, 95% CI 0.69–0.82) compared to studies of late image evaluation (sensitivity, 0.78, 95% CI 0.71–0.84; sensitivity, 0.71, 95% CI 0.61–0.79). In terms of differences among studies using different flip angles, studies using a high flip angle (90°) demonstrated lower sensitivity (0.75, 95% CI 0.67–0.81) and specificity (0.69, 95% CI 0.57–0.79) compared to studies using low flip angle (≤40°; sensitivity, 0.87, 95% CI 0.81–0.92; sensitivity, 0.79, 95% CI 0.73–0.84), and intermediate or mixed flip angle (60°, one study used 35°–90°; sensitivity, 0.90, 95% CI 0.80–0.95; sensitivity, 0.86, 95% CI 0.59–0.96), though the number of studies using intermediate or mixed flip angle was small (4 studies). A small number of studies (4) used both a high flip angle (90°) and leakage correction, yielding a combined sensitivity of 0.75 (n = 88; 95% CI 0.65–0.83) and specificity of 0.64 (n = 67; 95% CI 0.52–0.75). When combining studies using a low or intermediate angle, the combined sensitivity was 0.87 (95% CI 0.82–0.91) and the combined specificity was 0.80 (95% CI 0.74–0.85).

Risk of Bias and Applicability

Figure 4 summarizes the assessment for risk of bias and applicability based on QUADAS-2. The risk of bias for patient selection is generally rated as low (n = 9), or unclear (n = 18) due to unclear patient sampling method or exclusion criteria. For index test (rCBV obtained from DSC), over half of the studies (n = 16) were rated as unclear due to unclear blinding of the radiologists to the results of the reference standards. The cutoff used to calculate sensitivity and specificity was to optimize the diagnostic performance in most studies (n = 24) and rarely pre-specified. This was not downrated since there was no widely accepted threshold for rCBV. The reference standard was generally histopathology and/or clinic-radiologic follow-up, rated as low risk since it is the current diagnosis standard. The included patient populations and reference standards apply well to the targeted patient population and the research question in this review. The applicability of index test was rated unclear in some studies due to unclear blinding and no use of the leakage correction.

Discussion

We provide a comprehensive review comprising of 28 studies to synthesize evidence on the diagnostic performance of rCBV from DSC imaging to distinguish disease progression and treatment-related changes. Compared to the earlier reviews[9–12,51,52] this review included the largest number of eligible studies with 1068 patients and provided combined estimates for AUROC for the first time in addition to combined estimates for sensitivity and specificity. Our results suggest that rCBV provides robust performance differentiating true tumor progression from otherwise similar appearing treatment-related changes. Subgroup analysis, which included relevant DSC sequence parameters, generally did not reveal important differences in AUROC. Based on risk of bias, consistency of study results, directness, precision of estimates, the strength of evidence was graded to be moderate. That is, while the included studies have potential for some bias, the deficiencies are not likely to invalidate results or introduce significant bias.[13,53] Taken together, the evidence presented herein suggests that DSC MRI derived rCBV values may provide additional diagnostic utility in the clinical response assessment of patients with HGG.

Nineteen studies produced a combined AUROC of 0.85 (95% CI, 0.81–0.90) for rCBV and demonstrated "excellent" ability to discriminate true tumor progression to treatment-related changes. Our combined estimates of sensitivity and specificity (84% and 78%) were obtained from 28 eligible studies, and indicated robust diagnostic performance. While they are slightly lower than those reported in the earlier reviews,[10–12] our inclusion of studies is more thorough and rigorous and we explicitly focused on quantitatively measured rCBV. The earlier reviews included 11 to 18 studies with inclusion of non-rCBV studies or abstracts in the meta-analysis.

Our analysis consisted of patients with high-grade glioma (grade 3 or 4) with reference standard either using histopathology or follow-up imaging to determine disease status. The studies also varied by country of conduct, use of concurrent chemoradiation therapy, time of image evaluation, and choice of DSC sequence parameters. About half of the included studies applied leakage correction (15 studies) using various software (most commonly, 6 studies used Nordic Ice software) and preload of contrast agents; the other studies either did not apply, or did not report whether or not they were used. Fortunately given the relatively large number of included studies, we were able to conduct extensive subgroup analyses to evaluate the impact of single study and imaging characteristics, on AUROC, sensitivity, and specificity. These diagnostic measures were largely consistent across these characteristics, except for notable differences in a couple of cases. For example, studies of early image evaluation showed higher sensitivity (0.87) and AUROC (0.87) than studies of late image evaluation (sensitivity 0.78; AUROC 0.77). Nevertheless, the number of studies was not large enough to evaluate the impact of multiple study characteristics. For example, preload, flip angle, and type of leakage correction may interact to affect imaging quality and diagnostic performance, but the number of studies did not support subgroup analyses based on the combination of these characteristics.

Table 3. Subgroup Analysis of Sensitivity and Specificity, by Study and Imaging Characteristics

		Sensitivity			Specificity			LR+	LR-
	N of studies	*N* of patients	Pooled Sensitivity (95% CI)	I^2, *P*-value	*N* of patients	Pooled Specificity (95% CI)	I^2, *P*-value	Pooled LR+ (95% CI)	Pooled LR- (95% CI)
All studies	28	638	0.84 (0.80, 0.88)	31.0%, *P* = .06	430	0.78 (0.72, 0.83)	29.8%, *P* = .07	3.8(2.9, 5.0)	0.20 (0.15, 0.27)
All studies, using Max rCBV first	28	638	0.85 (0.81, 0.88)	16.5%, *P* = .22	430	0.78 (0.72, 0.83)	31.4%, *P* = .06	3.8(2.9, 4.9)	0.20 (0.16, 0.25)
All studies, excluding Blasel, 2016	27	601	0.85 (0.81, 0.88)	14.4%, *P* = .25	423	0.78 (0.72, 0.84)	32.6%, *P* = .05	3.9(3.0, 5.2)	0.20 (0.15, 0.25)
Optimal diagnostic performance	24	544	0.86 (0.82, 0.89)	11.4%, *P* = .30	361	0.78 (0.74, 0.82)	19.0%, *P* = .20	4.0(3.2, 4.8)	0.18 (0.15, 0.23)
Imaging parameter: Mean rCBV	19	413	0.82 (0.75, 0.86)	41.0%, *P* = .03	235	0.79 (0.71, 0.85)	23.8%, *P* = .17	3.9(2.8, 5.4)	0.23 (0.17, 0.32)
Imaging parameter: Max rCBV	17	469	0.85 (0.80, 0.88)	32.3%, *P* = .10	305	0.75 (0.69, 0.80)	24.0%, *P* = .18	3.4(2.7, 4.4)	0.20 (0.15, 0.28)
Country									
USA	10	190	0.83 (0.76, 0.88)	0.0%, *P* = .44	88	0.77 (0.60, 0.88)	52.3%, *P* = .03	3.6(1.9, 6.9)	0.22 (0.14, 0.35)
Non-USA	18	448	0.84 (0.79, 0.89)	43.7%, *P* = .03	342	0.78 (0.73, 0.83)	9.1%, *P* = .35	3.8(3.0, 4.9)	0.20 (0.14, 0.28)
Study Design									
Prospective	4	100	0.88 (0.66, 0.96)	76.6%, *P* = .01	65	0.81(0.56, 0.93)	33.0%, *P* = .21	4.6(1.6,13.5)	0.15(0.04, 0.55)
Retrospective	24	538	0.84 (0.79, 0.87)	23.3%, *P* = .15	365	0.78 (0.72, 0.83)	31.4%, *P* = .07	3.8(2.9, 5.0)	0.21 (0.16, 0.28)
Industry Funded									
No, or No COI	19	506	0.85 (0.81, 0.89)	25.0%, *P* = .16	344	0.80 (0.72, 0.85)	34.7%, *P* = .07	4.2(3.0, 5.8)	0.19 (0.14, 0.25)
NR	7	117	0.81 (0.68, 0.89)	38.2%, *P* = .14	74	0.78 (0.67, 0.87)	0.0%, *P* = .81	3.7(2.3, 6.0)	0.25 (0.14, 0.44)
Yes or COI	2	15	0.73 (0.47, 0.90)	0.0%, *P* = .40	12	0.42 (0.18, 0.69)	0.0%, *P* = .69	1.3 (0.7, 2.2)	0.64 (0.22, 1.87)
Histology									
GBM only	13	378	0.85 (0.79, 0.90)	43.0%, *P* = .05	285	0.78 (0.73, 0.83)	0.0%, *P* = .73	3.9(3.1, 4.9)	0.19 (0.13, 0.28)
High-Grade Glioma	15	260	0.83 (0.76, 0.88)	18.4%, *P* = .25	145	0.81 (0.66, 0.91)	53.0%, *P* = .01	4.4(2.2, 8.7)	0.21 (0.14, 0.32)
Timing of Image Evaluation									
Early	6	205	0.87 (0.80, 0.91)	12.8%, *P* = .33	193	0.76 (0.69, 0.82)	0.0%, *P* = .47	3.6 (2.7, 4.8)	0.17 (0.12, 0.26)
Late	7	157	0.78 (0.71, 0.84)	0.0%, *P* = .97	89	0.71 (0.61, 0.79)	0.0%, *P* = .50	2.7 (1.9, 3.7)	0.31 (0.23, 0.43)
Both	11	212	0.85 (0.76, 0.91)	60.3%, *P* = .01	98	0.89 (0.74, 0.96)	39.4%, *P* = .09	7.5 (3.0, 19.0)	0.17 (0.10, 0.29)
NR	4	64	0.86 (0.68, 0.95)	41.4%, *P* = .16	50	0.79 (0.51, 0.93)	72.4%, *P* = .01	4.0 (1.4, 11.9)	0.18 (0.06, 0.52)
Reference Standard									
Histopathology in ≥75% of pts	10	214	0.83(0.76, 0.88)	0.0%, *P* = .55	99	0.84 (0.70, 0.93)	48.9%, *P* = .04	5.3(2.5, 11.1)	0.20 (0.14, 0.30)
Histopathology in <75% of pts	18	424	0.85(0.79, 0.89)	45.5%, *P* = .02	331	0.76 (0.70, 0.81)	18.5%, *P* = .23	3.5(2.7, 4.6)	0.20 (0.14, 0.29)
Risk in Patient Selection									
Low	9	301	0.84(0.79, 0.88)	0.0%, *P* = .58	266	0.78 (0.72, 0.83)	32.7%, *P* = .16	3.8(2.9, 5.0)	0.21(0.16, 0.27)
Unclear	18	333	0.84(0.77, 0.89)	41.5%, *P* = .03	158	0.77 (0.67, 0.84)	20.2%, *P* = .21	3.6(2.4, 5.2)	0.21(0.14, 0.31)
High	1	4	1.0 (0.40, 1.0)	Not applicable	6	1.0 (0.54, 1.0)	Not applicable	Not applicable	

Table 3. Continued

		Sensitivity			Specificity			LR+	LR-
	N of studies	*N* of patients	Pooled Sensitivity (95% CI)	I^2, *P*-value	*N* of patients	Pooled Specificity (95% CI)	I^2, *P*-value	Pooled LR+ (95% CI)	Pooled LR- (95% CI)
Risk in IndexTest									
Low	11	295	0.84 (0.76, 0.90)	44.2%, $P = .06$	215	0.81 (0.67,0.90)	60.7%, $P < .01$	4.3(2.3, 8.2)	0.20(0.12, 0.34)
Unclear	17	343	0.84 (0.78, 0.88)	25.2%, $P = .16$	215	0.77 (0.71, 0.83)	0.0%, $P = .52$	3.7(2.8, 4.8)	0.21(0.15, 0.29)
ChemoradiationTherapy									
CCRT	16	431	0.84 (0.79, 0.88)	27.4%, $P = .15$	315	0.76 (0.69, 0.81)	22.6%, $P = .20$	3.5 (2.7, 4.5)	0.21(0.16, 0.28)
Non-CCRT	12	207	0.84 (0.75, 0.90)	41.2%, $P = .07$	115	0.83 (0.71,0.91)	41.1%, $P = .07$	5.0 (2.7, 9.6)	0.19(0.12, 0.32)
Imaging Field Strength									
1.5T	7	129	0.84 (0.71, 0.92)	45.5%, $P = .09$	65	0.80 (0.64, 0.90)	29.7%, $P = .20$	4.3 (2.1, 8.9)	0.20 (0.09, 0.41)
3.0T	15	434	0.83 (0.78, 0.87)	28.7%, $P = .14$	324	0.77 (0.71, 0.81)	20.2%, $P = .23$	3.5 (2.8, 4.5)	0.22 (0.17, 0.30)
1.5T or 3.0T	5	67	0.89 (0.73, 0.96)	44.9%, $P = .12$	33	0.86 (0.51, 0.97)	63.1%, $P = .03$	6.3 (1.4, 28.9)	0.12 (0.04, 0.37)
Preload									
Yes	14	339	0.78 (0.73, 0.83)	0.0%, $P = .61$	247	0.74 (0.75, 0.81)	43.1%, $P = .04$	3.0 (2.1, 4.1)	0.30 (0.23, 0.39)
No	4	106	0.86 (0.74, 0.93)	37.5%, $P = .19$	56	0.81 (0.68, 0.89)	0.0%, $P = .84$	4.4 (2.5, 7.9)	0.18 (0.09, 0.34)
NR	10	193	0.91 (0.85, 0.95)	13.7%, $P = .32$	127	0.82(0.69, 0.91)	25.3%, $P = .21$	5.1 (2.7, 9.6)	0.11 (0.06, 0.20)
Leakage Correction									
Yes	15	425	0.85 (0.80,0.89)	33.3%, $P = .10$	324	0.76 (0.67, 0.83)	50.1%, $P = .01$	3.5 (2.5, 5.1)	0.20 (0.14, 0.29)
No	8	144	0.83 (0.76,0.88)	1.5%, $P = .42$	77	0.79 (0.69, 0.87)	0.0%, $P = .88$	4.0 (2.6, 6.2)	0.22 (0.15, 0.32)
NR	5	69	0.77 (0.59, 0.89)	0.0%, $P = .87$	29	0.92 (0.66, 0.98)	0.0%, $P = .83$	9.2 (1.8, 47.5)	0.25 (0.12, 0.50)
Sequence & Flip Angle									
GRE	24	607	0.84 (0.80, 0.88)	36.8%, $P = .04$	404	0.78 (0.72, 0.83)	27.2%, $P = .11$	3.7 (2.9, 4.9)	0.21 (0.15, 0.27)
Low	9	289	0.87 (0.81, 0.92)	30.4%, $P = .18$	250	0.79 (0.73, 0.84)	8.9%, $P = .36$	4.1 (3.1, 5.3)	0.16 (0.11, 0.25)
Intermediate or Mixed	4	102	0.90 (0.80, 0.95)	15.5%, $P = .31$	46	0.86 (0.59, 0.96)	50.9%, $P = .11$	6.4 (1.9, 21.6)	0.12 (0.06, 0.23)
Low or Intermediate	12	359	0.87 (0.82, 0.91)	28.4%, $P = .17$	277	0.80 (0.74,0.85)	22.3%, $P = .23$	4.4 (3.3, 5.8)	0.16 (0.11, 0.23)
High	7	161	0.75 (0.67, 0.81)	0.0%, $P = .59$	85	0.69 (0.57, 0.79)	17.3%, $P = .30$	2.4 (1.6, 3.5)	0.37 (0.26, 0.52)
NR	4	55	0.79 (0.65, 0.89)	0.0%, $P = .49$	23	0.85 (0.51, 0.97)	31.5%, $P = .22$	5.5 (1.3, 23.8)	0.24 (0.13, 0.43)
SE or SE/GRE	4	43	0.83 (0.61, 0.94)	0.0%, $P = .44$	30	0.83 (0.52, 0.96)	58.3%, $P = .07$	4.9(1.3, 18.3)	0.20 (0.07, 0.57)

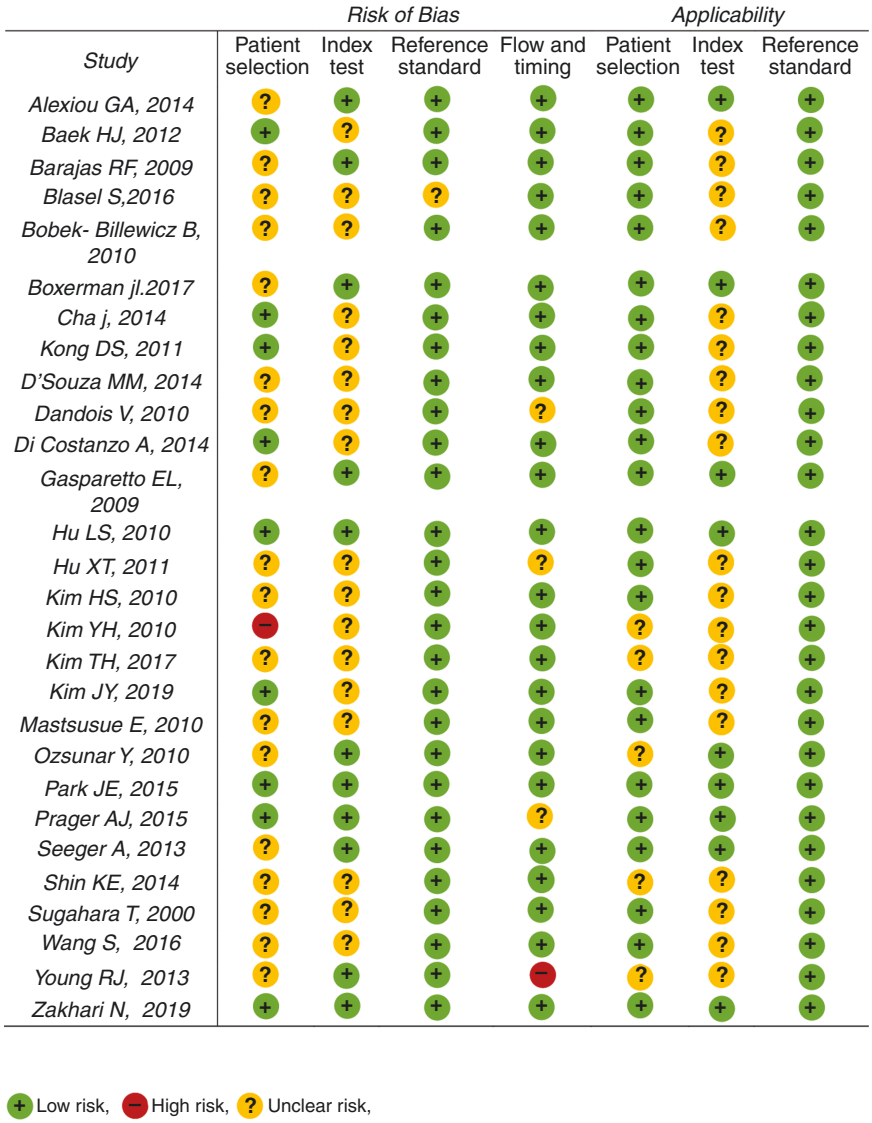

Figure 4. Summary of Risk of Bias and Applicability based on QUADAS-2.

Current consensus recommendations suggest a full dose gadolinium contrast preload (0.1 mmol/kg) followed by a full dose contrast bolus administered in the setting of intermediate 60° flip angle.[54] However, there is growing clinical desire to provide only one full dose of gadolinium contrast. Therefore, consensus recommendations provide for the use of a low 30° flip angle in this context; obviating the use of preload technique. Our results suggest that studies with high flip angle may impact the diagnostic performance of DSC rCBV with lower sensitivity and specificity. Studies with low and intermediate flip angles indicated similar AUROC, sensitivity and specificity, and better performance than high flip angle. This finding is consistent with current clinical consensus guideline for the use of DSC perfusion MRI.[27,54] On the other hand, the number of patients included in these studies are still small to clearly delineate the impact of study and imaging characteristics, in particular, the impact of the combination of imaging characteristics, and no definitive conclusions on the impact of these characteristics could be made. Our subgroup analysis highlights the critical need for reporting of relevant MRI sequence parameters and timing of contrast administration in the medical literature. Variation in MRI sequence parameters, as well as timing and dosage of MRI contrast administration, may affect quantitative metrics, and therefore the reliability of imaging measurements for tumor response assessment. The adoption of widely reproducible imaging approaches recommended through consensus guidelines may help to further improve the medical literature and subsequent systemic reviews.[27,54,55]

The lack of histopathological consensus guidelines for treatment-related change (vs. true tumor progression) and prognostics of patient outcomes remains a limitation for patients with changing imaging features

following Stupp protocol chemoradiotherapy.[56] Indeed, our study found that utilization of a histopathological reference standard in 75% of the cases did not influence the sensitivity and specificity of rCBV (Table 3). There are numerous possible etiologies for this observation, with likely contributing factors including variation in thresholds for diagnosing tumor progression used by pathologists and high inter-interpreter discrepancy in pathological interpretation,[57] and the use of gadolinium enhancement as the only metric to target tissue sampling. Furthermore, the diagnosis is complicated by the fact that some tissue samples have an admixture of treatment-related change and recurrent tumor, so it is not always a dichotomy and the disease may span over a spectrum.

Other limitations of this review are that the included studies were generally small and retrospective, and many studies did not report clear patient sampling method or exclusion criteria, or clear blinding of the radiologists to the results of the reference standards. While the consistency in results across the studies might slightly mitigate this limitation, the literature lacks well-designed and prospective studies, which remains a major limitation. Although prospective studies showed somewhat worse diagnostic performance, in particular, for AUROC (0.77 for prospective studies vs. 0.87 for retrospective studies), no meaningful interpretation and conclusion could be drawn given the small number of prospective studies and small sample size. This highlights the critical need for large prospective studies to provide reliable evidence.

Similar to earlier reviews, we combined sensitivity and specificity using different cutoff points, since optimal cutoff points are expected to be different from included studies given the differences in DSC techniques and postprocessing methods. A common optimal cutoff point could not be obtained from the included studies, which limited meaningful comparison and direct application of the study results across settings. Standardization of imaging parameters and definitions of true tumor progression and treatment-related changes will be necessary to create a common optimal cutoff point.

In conclusion, in this review, we found that there is moderate strength of evidence that rCBV obtained from DSC imaging demonstrated "excellent" ability to discriminate true tumor progression to treatment-related changes, with robust sensitivity and specificity.

Supplementary material

Supplemental material is available at *Neuro-Oncology Advances* online.

Acknowledgments

The authors thank Andrew Hamilton, M.S., M.L.S, for helping with literature search, and Ms. Sofia Gallamore for administrative support and table formatting.

Funding

This work was supported in part by National Institutes of Health grants CA199111, the Jonathan D. Lewis Foundation, and by the Walter S. and Lucienne Driskill Foundation, all to EAN, and in part by National Institutes of Health/National Cancer Institute 1K08CA237809-01A1 to RFB.

Conflict of interest statement. The authors have no conflicts of interest to disclose.

Authorship statement. RF contributed significantly to the study design, conduct of the systematic review, analysis and interpretation of the data, as well as draft and revision of the manuscript. LS contributed significantly to the conduct of the systematic review, interpretation of the data, and revision of the manuscript. RFB, PA and CV contributed significantly to the interpretation of the data and revision of the manuscript. EAN contributed significantly to the revision of the manuscript. All authors have read and approved the final version.

References

1. Stupp R, Mason WP, van den Bent MJ, et al. Radiotherapy plus concomitant and adjuvant temozolomide for glioblastoma. *N Engl J Med.* 2005;352(10):987–996.
2. Hygino da Cruz LC, Jr., Rodriguez I, Domingues RC, Gasparetto EL, Sorensen AG. Pseudoprogression and pseudoresponse: imaging challenges in the assessment of posttreatment glioma. *AJNR Am J Neuroradiol.* 2011;32(11):1978–1985.
3. Mabray MC, Barajas RF, Jr., Cha S. Modern brain tumor imaging. *Brain Tumor Res Treat.* 2015; 3(1):8–23.
4. Abbasi AW, Westerlaan HE, Holtman GA, et al. Incidence of tumour progression and pseudoprogression in high-grade gliomas: a systematic review and meta-analysis. *Clin Neuroradiol.* 2018;28(3):401–411.
5. Louis DN, Perry A, Burger P, et al. International Society Of Neuropathology--Haarlem consensus guidelines for nervous system tumor classification and grading. *Brain Pathol.* 2014;24(5): 429–435.
6. Lin NU, Lee EQ, Aoyama H, et al. Response assessment criteria for brain metastases: proposal from the RANO group. *Lancet Oncol.* 2015;16(6):e270–e278.
7. Shiroishi MS, Castellazzi G, Boxerman JL, et al. Principles of T2 *-weighted dynamic susceptibility contrast MRI technique in brain tumor imaging. *J Magn Reson Imaging.* 2015;41(2):296–313.
8. Boxerman JL, Shiroishi MS, Ellingson BM, Pope WB. Dynamic susceptibility contrast mr imaging in glioma: review of current clinical practice. *Magn Reson Imaging Clin N Am.* 2016;24(4):649–670.
9. Dongas J, Asahina A, Bacchi S, Patel S. Magnetic resonance perfusion imaging in the diagnosis of high-grade glioma progression and treatment-related changes: a systematic review. *Open J Modern Neurosurg.* 2018;8(3):282–305.

10. Patel P, Baradaran H, Delgado D, et al. MR perfusion-weighted imaging in the evaluation of high-grade gliomas after treatment: a systematic review and meta-analysis. *Neuro Oncol.* 2017;19(1):118–127.
11. Wan B, Wang S, Tu M, et al. The diagnostic performance of perfusion MRI for differentiating glioma recurrence from pseudoprogression: a meta-analysis. *Medicine (Baltim).* 2017;96(11):e6333.
12. van Dijken BRJ, van Laar PJ, Holtman GA, van der Hoorn A. Diagnostic accuracy of magnetic resonance imaging techniques for treatment response evaluation in patients with high-grade glioma, a systematic review and meta-analysis. *Eur Radiol.* 2017;27(10):4129–4144.
13. Chang SM, Matchar DB, Smetana GW, et al., editors. Methods guide for medical test reviews [Internet]. Rockville (MD): Agency for Healthcare Research and Quality (US); 2012 Jun. https://www.ncbi.nlm.nih.gov/books/NBK98241/.
14. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *Bmj.* 2021;372:n71.
15. Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med.* 2011;155(8):529–536.
16. Alexiou GA, Zikou A, Tsiouris S, et al. Comparison of diffusion tensor, dynamic susceptibility contrast MRI and (99m)Tc-Tetrofosmin brain SPECT for the detection of recurrent high-grade glioma. *Magn Reson Imaging.* 2014;32(7):854–859.
17. Baek HJ, Kim HS, Kim N, Choi YJ, Kim YJ. Percent change of perfusion skewness and kurtosis: a potential imaging biomarker for early treatment response in patients with newly diagnosed glioblastomas. *Radiology.* 2012;264(3):834–843.
18. Barajas RF, Jr., Chang JS, Segal MR, et al. Differentiation of recurrent glioblastoma multiforme from radiation necrosis after external beam radiation therapy with dynamic susceptibility-weighted contrast-enhanced perfusion MR imaging. *Radiology.* 2009;253(2):486–496.
19. Blasel S, Zagorcic A, Jurcoane A, et al. Perfusion MRI in the evaluation of suspected glioblastoma recurrence. *J Neuroimaging.* 2016;26(1):116–123.
20. Bobek-Billewicz B, Stasik-Pres G, Majchrzak H, Zarudzki L. Differentiation between brain tumor recurrence and radiation injury using perfusion, diffusion-weighted imaging and MR spectroscopy. *Folia Neuropathol.* .2010;48(2):81–92.
21. Boxerman JL, Ellingson BM, Jeyapalan S, et al. Longitudinal DSC-MRI for Distinguishing tumor recurrence from pseudoprogression in patients with a high-grade glioma. *Am J Clin Oncol.* 2017;40(3):228–234.
22. Cha J, Kim ST, Kim HJ, et al. Differentiation of tumor progression from pseudoprogression in patients with posttreatment glioblastoma using multiparametric histogram analysis. *Am J Neuroradiol.* 2014;35(7):1309–1317.
23. D'Souza MM, Sharma R, Jaimini A, et al. 11C-MET PET/CT and advanced MRI in the evaluation of tumor recurrence in high-grade gliomas. *Clin Nucl Med.* 2014;39(9):791–798.
24. Dandois V, Rommel D, Renard L, Jamart J, Cosnard G. Substitution of 11C-methionine PET by perfusion MRI during the follow-up of treated high-grade gliomas: preliminary results in clinical practice. *J Neuroradiol.* 2010;37(2):89–97.
25. Di Costanzo A, Scarabino T, Trojsi F, et al. Recurrent glioblastoma multiforme versus radiation injury: a multiparametric 3-T MR approach. *Radiol Med.* 2014;119(8):616–624.
26. Gasparetto EL, Pawlak MA, Patel SH, et al. Posttreatment recurrence of malignant brain neoplasm: accuracy of relative cerebral blood volume fraction in discriminating low from high malignant histologic volume fraction. *Radiology* 2009;250(3):887–896.
27. Hu LS, Baxter LC, Pinnaduwage DS, et al. Optimized preload leakage-correction methods to improve the diagnostic accuracy of dynamic susceptibility-weighted contrast-enhanced perfusion MR imaging in posttreatment gliomas. *AJNR Am J Neuroradiol.* 2010;31(1):40–48.
28. Hu X, Wong KK, Young GS, Guo L, Wong ST. Support vector machine multiparametric MRI identification of pseudoprogression from tumor recurrence in patients with resected glioblastoma. *J Magn Reson Imaging.* 2011;33(2):296–305.
29. Kim HS, Kim JH, Kim SH, Cho KG, Kim SY. Posttreatment high-grade glioma: usefulness of peak height position with semiquantitative MR perfusion histogram analysis in an entire contrast-enhanced lesion for predicting volume fraction of recurrence. *Radiology* 2010;256(3):906–915.
30. Kim YH, Oh SW, Lim YJ, et al. Differentiating radiation necrosis from tumor recurrence in high-grade gliomas: assessing the efficacy of 18F-FDG PET, 11C-methionine PET and perfusion MRI. *Clin Neurol Neurosurg.* 2010;112(9):758–765.
31. Kim TH, Yun TJ, Park CK, et al. Combined use of susceptibility weighted magnetic resonance imaging sequences and dynamic susceptibility contrast perfusion weighted imaging to improve the accuracy of the differential diagnosis of recurrence and radionecrosis in high-grade glioma patients. *Oncotarget* 2017;8(12):20340–20353.
32. Kim JY, Park JE, Jo Y, et al. Incorporating diffusion- and perfusion-weighted MRI into a radiomics model improves diagnostic performance for pseudoprogression in glioblastoma patients. *Neuro Oncol.* 2019;21(3):404–414.
33. Kong DS, Kim ST, Kim EH, et al. Diagnostic dilemma of pseudoprogression in the treatment of newly diagnosed glioblastomas: the role of assessing relative cerebral blood flow volume and oxygen-6-methylguanine-DNA methyltransferase promoter methylation status. *AJNR Am J Neuroradiol.* 2011;32(2):382–387.
34. Matsusue E, Fink JR, Rockhill JK, Ogawa T, Maravilla KR. Distinction between glioma progression and post-radiation change by combined physiologic MR imaging. *Neuroradiology* 2010;52(4):297–306.
35. Ozsunar Y, Mullins ME, Kwong K, et al. Glioma recurrence versus radiation necrosis? A pilot comparison of arterial spin-labeled, dynamic susceptibility contrast enhanced MRI, and FDG-PET imaging. *Acad Radiol.* 2010;17(3):282–290.
36. Park JE, Kim HS, Goh MJ, Kim SJ, Kim JH. Pseudoprogression in patients with Glioblastoma: assessment by using volume-weighted voxel-based multiparametric clustering of MR imaging data in an independent test set. *Radiology* 2015;275(3):792–802.
37. Prager AJ, Martinez N, Beal K, et al. Diffusion and perfusion MRI to differentiate treatment-related changes including pseudoprogression from recurrent tumors in high-grade gliomas with histopathologic evidence. *Am J Neuroradiol.* 2015;36(5):877–885.
38. Seeger A, Braun C, Skardelly M, et al. Comparison of three different MR perfusion techniques and MR spectroscopy for multiparametric assessment in distinguishing recurrent high-grade gliomas from stable disease. *Acad Radiol.* 2013;20(12):1557–1565.
39. Shin KE, Ahn KJ, Choi HS, et al. DCE and DSC MR perfusion imaging in the differentiation of recurrent tumour from treatment-related changes in patients with glioma. *Clin Radiol.* 2014;69(6):e264–e272.
40. Sugahara T, Korogi Y, Tomiguchi S, et al. Posttherapeutic intraaxial brain tumor: the value of perfusion-sensitive contrast-enhanced MR imaging for differentiating tumor recurrence from nonneoplastic contrast-enhancing tissue. *Am J Neuroradiol.* 2000;21(5):901–909.
41. Wang S, Martinez-Lage M, Sakai Y, et al. Differentiating tumor progression from pseudoprogression in patients with glioblastomas using diffusion tensor imaging and dynamic susceptibility contrast MRI. *Am J Neuroradiol.* 2016;37(1):28–36.

42. Young RJ, Gupta A, Shah AD, et al. MRI perfusion in determining pseudoprogression in patients with glioblastoma. *Clin Imaging*. 2013;37(1):41–49.
43. Zakhari N, Taccone MS, Torres CH, et al. Prospective comparative diagnostic accuracy evaluation of dynamic contrast-enhanced (DCE) vs. dynamic susceptibility contrast (DSC) MR perfusion in differentiating tumor recurrence from radiation necrosis in treated high-grade gliomas. *J Magn Reson Imaging*. 2019;50(2):573–582.
44. Hardy RJ, Thompson SG. A likelihood approach to meta-analysis with random effects. *Stat Med*. 1996;15(6):619–629.
45. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*. 1982;143(1):29–36.
46. Hosmer DW, Jr., Lemeshow S, Sturdivant RX. *Applied Logistic Regression*. 3rd ed. Hoboken, New Jersey: Wiley; 2013.
47. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ*. 2003;327(7414):557–560.
48. Zwinderman AH, Bossuyt PM. We should not pool diagnostic likelihood ratios in systematic reviews. *Stat Med*. 2008;27(5):687–697.
49. Trikalinos TA, Balion CM, Coleman CI, et al. Chapter 8: meta-analysis of test performance when there is a "gold standard". *J Gen Intern Med*. 2012;27(Suppl 1):S56–S66.
50. Singh S, Chang SM, Matchar DB, Bass EB. Chapter 7: grading a body of evidence on diagnostic tests. *J Gen Intern Med*. 2012;27(Suppl 1):S47–S55.
51. Li Z, Zhou Q, Li Y, et al. Mean cerebral blood volume is an effective diagnostic index of recurrent and radiation injury in glioma patients: a meta-analysis of diagnostic test. *Oncotarget*. 2017;8(9):15642–15650.
52. Suh CH, Kim HS, Jung SC, Choi CG, Kim SJ. Multiparametric MRI as a potential surrogate endpoint for decision-making in early treatment response following concurrent chemoradiotherapy in patients with newly diagnosed glioblastoma: a systematic review and meta-analysis. *Eur Radiol*. 2018;28(6):2628–2638.
53. Berkman ND, Lohr KN, Ansari M, et al. Grading the strength of a body of evidence when assessing health care interventions for the effective health care program of the agency for healthcare research and quality: an update. 2013 Nov 18. In: *Methods Guide for Effectiveness and Comparative Effectiveness Reviews [Internet]*. Rockville (MD): Agency for Healthcare Research and Quality (US); 2008. https://www.ncbi.nlm.nih.gov/books/NBK174881/.
54. Boxerman JL, Quarles CC, Hu LS, et al. Consensus recommendations for a dynamic susceptibility contrast MRI protocol for use in high-grade gliomas. *Neuro Oncol*. 2020;22(9):1262–1275.
55. Ellingson BM, Bendszus M, Boxerman J, et al. Consensus recommendations for a standardized Brain Tumor Imaging Protocol in clinical trials. *Neuro Oncol*. 2015;17(9):1188–1198.
56. Woodworth GF, Garzon-Muvdi T, Ye X, et al. Histopathological correlates with survival in reoperated glioblastomas. *J Neurooncol*. 2013;113(3):485–493.
57. Holdhoff M, Ye X, Piotrowski AF, et al. The consistency of neuropathological diagnoses in patients undergoing surgery for suspected recurrence of glioblastoma. *J Neurooncol*. 2019;141(2):347–354.