

MR perfusion-weighted imaging in the evaluation of high-grade gliomas after treatment: a systematic review and meta-analysis

Praneil Patel, Hadiyah Baradaran, Diana Delgado, Gulce Askin, Paul Christos, Apostolos John Tsiouris, and Ajay Gupta

Department of Radiology, Weill Cornell Medical College/New York-Presbyterian Hospital, New York, New York (P.P., H.B., A.J.T., A.G.); Samuel J. Wood Library & C. V. Starr Biomedical Information Center, Weill Cornell Medical College, New York, New York (D.D.); Department of Healthcare Policy and Research, Weill Cornell Medical College, New York, New York (G.A., P.C.)

Corresponding Author: Praneil Patel, 525 E. 68th St, Box 141 Starr 8A, New York, NY 10065, USA (pmp9007@med.cornell.edu).

Abstract

Background. Distinction between tumor and treatment related changes is crucial for clinical management of patients with high-grade gliomas. Our purpose was to evaluate whether dynamic susceptibility contrast-enhanced (DSC) and dynamic contrast enhanced (DCE) perfusion-weighted imaging (PWI) metrics can effectively differentiate between recurrent tumor and posttreatment changes within the enhancing signal abnormality on conventional MRI.

Methods. A comprehensive literature search was performed for studies evaluating PWI-based differentiation of recurrent tumor and posttreatment changes in patients with high-grade gliomas (World Health Organization grades III and IV). Only studies published in the “temozolomide era” beginning in 2005 were included. Summary estimates of diagnostic accuracy were obtained by using a random-effects model.

Results. Of 1581 abstracts screened, 28 articles were included. The pooled sensitivities and specificities of each study's best performing parameter were 90% and 88% (95% CI: 0.85–0.94; 0.83–0.92) and 89% and 85% (95% CI: 0.78–0.96; 0.77–0.91) for DSC and DCE, respectively. The pooled sensitivities and specificities for detecting tumor recurrence using the 2 most commonly evaluated parameters, mean relative cerebral blood volume (rCBV) (threshold range, 0.9–2.15) and maximum rCBV (threshold range, 1.49–3.1), were 88% and 88% (95% CI: 0.81–0.94; 0.78–0.95) and 93% and 76% (95% CI: 0.86–0.98; 0.66–0.85), respectively.

Conclusions. PWI-derived thresholds separating viable tumor from treatment changes demonstrate relatively good accuracy in individual studies. However, because of significant variability in optimal reported thresholds and other limitations in the existing body of literature, further investigation and standardization is needed before implementing any particular quantitative PWI strategy across institutions.

Key words:

gliomas | meta-analysis | MR perfusion | pseudoprogression | radiation necrosis

Treatment of glioblastoma requires a multifaceted approach, with surgery, radiation, and temozolomide currently considered first-line therapies based on evidence from randomized trial data published in 2005.¹ A similar treatment protocol is typically applied to World Health Organization (WHO) grade III

gliomas with ongoing trials to determine the standard of care.² MRI is essential for evaluating treatment response, though in recent years mounting evidence has shown that tumor and treatment related changes, including pseudoprogression (typically within 6 mo of radiation and chemotherapy) and

the more delayed onset of late radiation necrosis, can appear indistinguishable at a single timepoint on conventional contrast-enhanced structural brain imaging.³⁻⁵ This distinction is crucial for determining the success of a particular treatment and subsequent clinical management, including whether second-line therapeutic agents should be administered and/or repeat surgical resection is necessary.

The use of a more physiological imaging approach in making this distinction by exploiting tumor angiogenesis and neovascularity has gained interest over the past decade.^{4,6,7} Dynamic susceptibility contrast-enhanced (DSC) perfusion is used to generate hemodynamic parameters such as relative cerebral blood volume (rCBV), peak height, and percentage signal intensity recovery.⁶ Dynamic contrast enhanced (DCE) perfusion alternatively uses a T1-relaxivity approach, generating quantitative parameters such as the volumetric transfer constant (K-trans), fractional plasma volume (Vp), and fractional volume of the extracellular extravascular space (Ve), as well as semiquantitative parameters based on the area under the curve (AUC) of signal intensity–time.^{8,9} Both perfusion-weighted imaging (PWI) approaches hold promise in separating radiation and chemotherapy effects from tumor (see [Supplementary Fig. E1](#) for an example). However, individual studies have generally been small and have used varied imaging techniques on heterogeneous patient groups, making it difficult to draw definitive conclusions capable of confidently guiding routine clinical practice. For this reason, we performed a systematic review and meta-analysis to evaluate whether PWI metrics can effectively differentiate between recurrent tumor and posttreatment changes within the enhancing signal abnormality present on conventional MRI.

Materials and Methods

The methodology for this study was based on the Preferred Reporting Items for Systematic Reviews and Meta-Analyses statement.¹⁰

Literature Search and Eligibility Criteria

A systematic search was performed on October 9, 2015 by a medical librarian searching Ovid MEDLINE, Ovid Embase, and the Cochrane Library. Details of the search methodology are provided in the [online Supplementary material](#). After preliminary articles were shortlisted as potentially eligible, additional records were identified by using the “Cited by” and “View references” features in Scopus on November 3, 2015. The titles and abstracts of all of the references were reviewed by 2 independent readers to exclude studies not meeting inclusion criteria. Full manuscripts were reviewed by the same readers to determine final inclusion. Studies specifically focused on PWI-based differentiation of recurrent tumor and post-treatment changes in patients with high-grade gliomas

(WHO grades III and IV) of any age were eligible. Studies with glioblastoma patients who did not undergo standard treatment (“Stupp protocol”)¹ were ineligible. To maximize the comparability of the type of tumors studied, articles that included brain metastases or other primary brain tumors, including low-grade gliomas, in the analysis were ineligible. Specific inclusion criteria were: (1) English language articles or foreign language articles with a provided English translation; (2) articles published in the temozolomide era beginning in 2005; (3) studies with at least 10 subjects on whom gadolinium DSC or DCE PWI of the brain was performed on new posttreatment enhancing lesions.

Data Extraction

Data were extracted by 2 independent readers using a pre-determined data collection template. The data collection template was developed, piloted on 3 randomly selected included studies, and refined as needed. Disagreements on article inclusion and data extraction were resolved by an independent reader as a tiebreaker. Study characteristics extracted included the first author, study design, number and average age of subjects, inclusion/exclusion criteria, treatment protocol, specific treatment effect being studied (pseudoprogression or radiation necrosis) and time between radiation and PWI, and reference standard for determining disease progression. The PWI parameters evaluated in each study were extracted along with any proposed threshold and sensitivity/specificity values. Data about the PWI imaging technique and analysis were also extracted, including MRI vendor and field strength, contrast agent and dose, the use of contrast preloading and/or postprocessing, software for PWI analysis, and methods of tumor segmentation and region of interest (ROI) analysis. In cases of potential overlap between patient cohorts in studies evaluating the same PWI parameter, an attempt was made to contact the corresponding author for clarification. When 2 potentially overlapping studies differed by the treatment effect under evaluation (for example, one study focusing on pseudoprogression and another on late radiation necrosis), data from each study were collected separately for each distinct posttreatment effect. If overlapping cohorts did not provide distinct information, the study with the larger cohort was selected.

Study Quality Assessment

Two independent readers assessed each study using a 14-item list across 4 domains (see [Supplementary material](#) for details). The Quality Assessment of Diagnostic Accuracy Studies (QUADAS)¹¹ and QUADAS-2¹² tools formed the basis of our assessment. The items were piloted and refined as needed. Methodological quality and to a lesser extent reporting quality were addressed for each study. A third reader served as a tiebreaker when necessary.

Statistical Analysis

In our primary meta-analysis, we pooled the sensitivity and specificity data where available of the best performing parameter from each study. We then performed secondary subgroup analyses pooling the sensitivities and specificities of (i) the most commonly evaluated parameters (eligible if evaluated in >3 studies) and (ii) the best performing parameter in studies specifically evaluating pseudoprogression rather than late radiation necrosis or a combination. In all cases, analysis was performed separately for DSC and DCE. Two-by-two contingency tables were reconstructed for each study where possible. Meta-analyses for the sensitivity and specificity proportions were conducted with the use of StatsDirect statistical software v3.0.161. For the sensitivity/specificity proportions of interest, the results of each study were expressed as binary proportions with 95% confidence intervals. Statistical heterogeneity was tested through a chi-square test (ie, Cochran's Q test), and $P < .20$ was used to indicate the presence of heterogeneity. In all cases, the more conservative random-effects models (DerSimonian–Laird) were used to calculate the pooled summary proportions. The presence of publication bias was evaluated by the use of funnel plots and the Begg–Mazumdar rank-correlation test ($P < .05$).

Results

Selection of Studies and Study Characteristics

A total of 1643 abstracts were screened, of which 45 potentially eligible articles were selected for further review (see [Supplementary Fig. E2](#)). Of these articles, 17 did not meet inclusion criteria when read in their entirety (see [Supplementary Fig. E2](#) for specific reasons). The remaining 28 studies met eligibility criteria for the systematic review.^{13–40} Seventeen studies evaluated DSC,^{13,15,17,18,20,23–25,27–31,34,36,39} 8 evaluated DCE,^{16,19,21,22,32,37,38,40} and 3 evaluated both DSC and DCE.^{26,33,35} There were a total of 937 cases (54%) of tumor progression and 806 cases (46%) of treatment change. Investigators reported that pseudoprogression was the specific treatment effect under evaluation in 13 studies,^{14,17,18,29–31,33,34,36–40} radiation necrosis in 4 studies,^{15,16,19,21} and both in 2 studies.^{25,34} Patients meeting inclusion criteria for pseudoprogression developed a new lesion during concurrent radiation and chemotherapy²⁹ out to 6 months after treatment completion.³⁹ Seven different types of PWI processing software were utilized, not including in-house software in several studies. Among DSC studies, 7 utilized contrast preloading, often in combination with postprocessing algorithms,^{20,23–26,33,35} while 10 utilized only postprocessing techniques.^{13,14,17,18,27,28,30,31,34,36} Six of the DCE studies used a model-free method,^{19,22,26,32,33,37} 4 used a 2-compartment model,^{16,35,38,40} and 1 used both.²¹ See [Table 1](#) for an overview of study characteristics and [Supplementary Table E1](#) for an overview of PWI protocols.

Diagnostic Accuracy

A complete depiction of the proposed threshold values can be found in [Supplementary Tables E2 and E3](#). For DSC ([Fig. 1](#)), the pooled sensitivity and specificity using the best performing parameter from each study were 90% (95% CI: 0.85–0.94) and 88% (95% CI: 0.83–0.92), respectively. For DCE ([Fig. 2](#)), the pooled sensitivity was 89% (95% CI: 0.78–0.96) and the pooled specificity was 85% (95% CI: 0.77–0.91). Two DSC PWI parameters, mean rCBV (rCBVmean, $n = 8$) and maximum rCBV (rCBVmax, $n = 5$), were eligible for the subgroup meta-analysis. Using rCBVmean ([Fig. 3](#)) with proposed thresholds ranging from 0.9 to 2.15, the pooled sensitivity and specificity for detecting tumor recurrence were both 88% (95% CI: 0.81–0.94 and 0.78–0.95, respectively). For rCBVmax ([Fig. 4](#)), the pooled sensitivity was 93% (95% CI: 0.86–0.98) and the pooled specificity was 76% (95% CI: 0.66–0.85) with proposed thresholds from 1.49 to 3.1. Subset analysis of DSC studies specifically evaluating pseudoprogression (see [Supplementary Fig. E3](#)) demonstrated a pooled sensitivity of 89% (95% CI: 0.83–0.94) and a pooled specificity of 80% (95% CI: 0.72–0.86). No DCE parameter was eligible for the subgroup meta-analysis because the minimum required number of unique studies ($n = 3$) was not found.

Heterogeneity and Publication Bias

See [Supplementary Figs E4–E6](#) for the funnel plots and publication bias and heterogeneity measures from each meta-analysis. There was evidence of borderline to significant heterogeneity in all but 2 of the 10 meta-analyses (specificity proportions for the rCBVmax and DSC pseudoprogression subgroups) performed as determined by the Q test. There was no evidence of significant publication bias in the primary meta-analysis for DSC studies and the secondary meta-analyses for the rCBVmean and rCBVmax subgroups. Only the sensitivity proportions in the primary meta-analysis for DCE studies and secondary meta-analysis of the DSC pseudoprogression subgroup showed borderline significant publication bias as determined by rank correlation tests ($P = .03$ and $.04$, respectively).

Quality of Studies

The results of the study quality assessment are summarized in [Fig. 5](#). There was potential introduction of bias in patient selection, the reference standard, and patient flow.¹² Only a minority of studies, 7 of 28, was prospective.^{13,16,23–25,29,40} While imaging protocols were typically well described, the reporting of treatment protocols was less clear. Patients on potentially confounding anti-angiogenic drugs or steroids were not excluded in at least 10 studies.^{19,22,25,26,28,31,33,34,37,40} Neuroradiologists were often blinded to the outcome at the time of PWI analysis, but most studies did not report whether the pathologist was blinded to the imaging results. Five studies relied solely on clinical and/or radiological follow-up for confirmation of PWI findings without any histopathological confirmation.^{30,35,36,38,40} Among the other 23 studies with at least one subject with histological follow-up, 14 provided some histological criteria for tumor

Table 1 Overview of study and patient characteristics

Study First Author and Year	Study Design	PWI Technique	# Patients (# GBM if mixed)	Mean Age and/or Range	Specific Treatment Effect*	Time Since Radiation (range if provided)	Reference Standard (# path)
Alexiou 2014 ¹³	P	DSC	30 (27)	61	UR	12 mo (3–24 mo)	Both (2)
Baek 2012 ¹⁴	R	DSC	79	50.6 (19–83)	PSP	Within 4 wk of end of CCRT	Both (5)
Barajas 2009 ¹⁵	R	DSC	66*	54.2	RN	1.7–50.2 mo	Both (55)
Bisdas 2011 ¹⁶	P	DCE	18*	UR	RN	1.8–7 mo	Both (5)
Cha 2014 ¹⁷	R	DSC	35	49 (24–70)	PSP	Within 180 days of CCRT	Both (3)
Choi 2013 ¹⁸	R	DSC	62	49.3 (22–79)	PSP	Within 4 wk end of CCRT	Both (5)
Chung 2013 ¹⁹	R	DCE	57	52.1 (25–69)	RN	39.6 wk	Path
Gasparetto 2009 ²⁰	R	DSC	12	33–71	UR	UR	Path
Hamilton 2015 ²¹	R	DCE	24 (15)	51	RN	2.6 ± 2.2 y	Path
Heo 2015 ²²	R	DCE	48	53.9 (27–73)	UR	14 wk	Both (9)
Hu 2009 ²³	P	DSC	11+ (9)	46.9 (31–62)	UR	3–28.5 mo	Path
Hu 2010 ²⁴	P	DSC	11+ (9)	46.9 (31–62)	UR	3–28.5 mo	Path
Hu 2012 ²⁵	P	DSC	9*	50 (25–73)	Both	64% <6 mo (1–5.5) 36% >6 mo (8–53)	Path
Kim 2010 ²⁶	R	DSC	39	48.2 (18–78)	UR	UR	Path
Kim 2014 ²⁷	R	DSC, DCE	169	52.2 (25–69)	UR	46.5 wk	Both (87)
Kim 2014 ²⁸	R	DSC	51	51.5 (25–72)	UR	43.2 wk	Path
Kong 2011 ²⁹	P	DSC	59	25–74	PSP	During CCRT or within 1–2 cycles of adjuvant chemo	Both (4)
Mangla 2010 ³⁰	R	DSC	19	65 (41–90)	PSP	1 mo after CCRT	Radiological
Martinez 2012 ³¹	R	DSC	34	54.6	PSP	UR	Both
Narang 2011 ³²	R	DCE	22 (17)	51.8 (18–70)	UR	UR for HGG	Both
Park 2015 ³³	R	DSC, DCE	54	45.5	PSP	Within 12 wk of end of CCRT	Both (11)
Prager 2015 ³⁴	R	DSC	68 (55)	54.9 (22.6–79.4)	Both with PSP subgroup	6.1 mo (0.4–40.4)	Path
Seeger 2013 ³⁵	R	DSC, DCE	40	53.6	UR	UR	
Song 2013 ³⁶	R	DSC	20	50.8 (24–68)	PSP	Within 2 mo end of CCRT	CR
Suh 2013 ³⁷	R	DCE	79	51.2 (25–69)	PSP	Within 5 wk of end of CCRT	Both (24)
Thomas 2015 ³⁸	R	DCE	37	37–87	PSP	UR	CR
Young 2013 ³⁹	R	DSC	20	9–84	PSP	0.5–6.3 mo	Both (9)
Yun 2015 ⁴⁰	P	DCE	33	54.6 (28–82)	PSP	UR	Radiological

GBM, glioblastoma multiforme; R = retrospective, P = prospective, * treatment effect was as described by author, + indicates study where some patients had more than one PWI study or biopsy contribute to analysis, UR = unreported, PSP = pseudoprogression, RN = radiation necrosis, CCRT = concurrent chemotherapy and radiation therapy, HGG = high-grade glioma, path = pathology, CR = clinical and radiological.

recurrence, with widely divergent definitions of tumor ranging from qualitative description to a variety of quantitative thresholds.^{15,16,19–28,34,39} All patients had PWI findings confirmed by a histological, clinical, and/or radiological reference standard, but in 14 studies the standard varied among patients.^{13–18,22,26,29,31–33,37,39}

Discussion

Discriminating between tumor recurrence and treatment change has been a topic of extensive research interest spanning multiple imaging modalities. In a prior

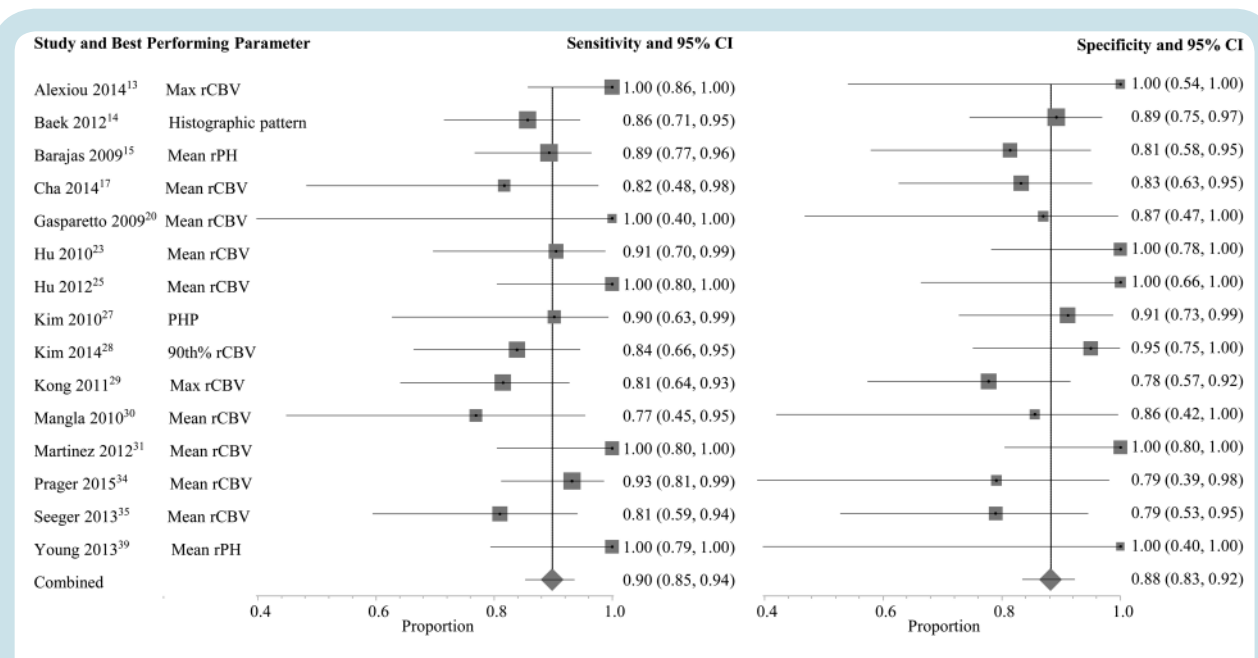


Fig. 1 Individual forest plots of the sensitivity and specificity of the best performing PWI parameter from each DSC study. Squares represent the sensitivity or specificity proportion with the size proportional to the weighting of the study. Diamond represents pooled estimate. Lines represent 95% CIs. rPH = relative peak height; PHP = peak height position; xth% = xth histogram percentile.

systematic review and meta-analysis, Shah et al compared MR, CT, and nuclear medicine techniques in differentiating tumor primarily from delayed radiation necrosis.⁴¹ Deng et al compared ¹¹C-Met PET and DSC for the detection of glioma recurrence and included studies predating the introduction of temozolomide.⁴² In contrast, we included

both early and delayed treatment changes while focusing specifically on DSC and DCE in high-grade gliomas, the techniques and clinical context that are most commonly encountered in practice.

DSC and DCE techniques each have their own established limitations. DSC has poorer spatial resolution and

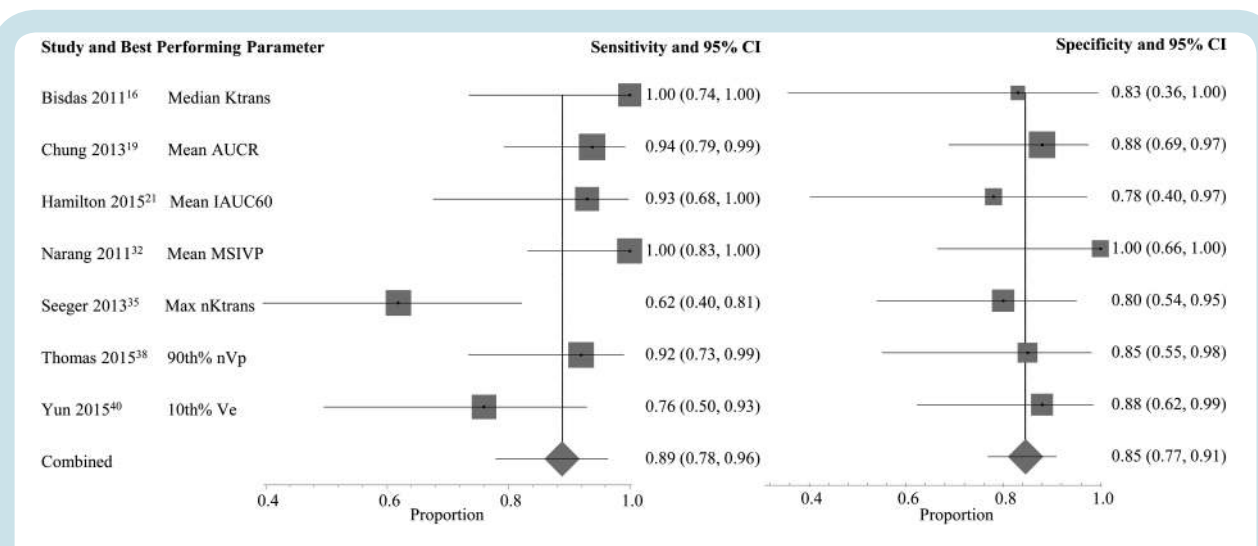


Fig. 2 Individual forest plots of the sensitivity and specificity of the best performing PWI parameter from each DCE study. Squares represent the sensitivity or specificity proportion with the size proportional to the weighting of the study. Diamond represents pooled estimate. Lines represent 95% CIs. K-trans = volumetric transfer coefficient, AUCR = area under curve ratio; IAUC60 = initial area under the curve after 60 s; MSIVP = maximum slope of enhancement in the initial vascular phase; nVp = normalized fractional plasma volume; Ve = fractional volume of the extracellular extravascular space; xth% = xth histogram percentile.

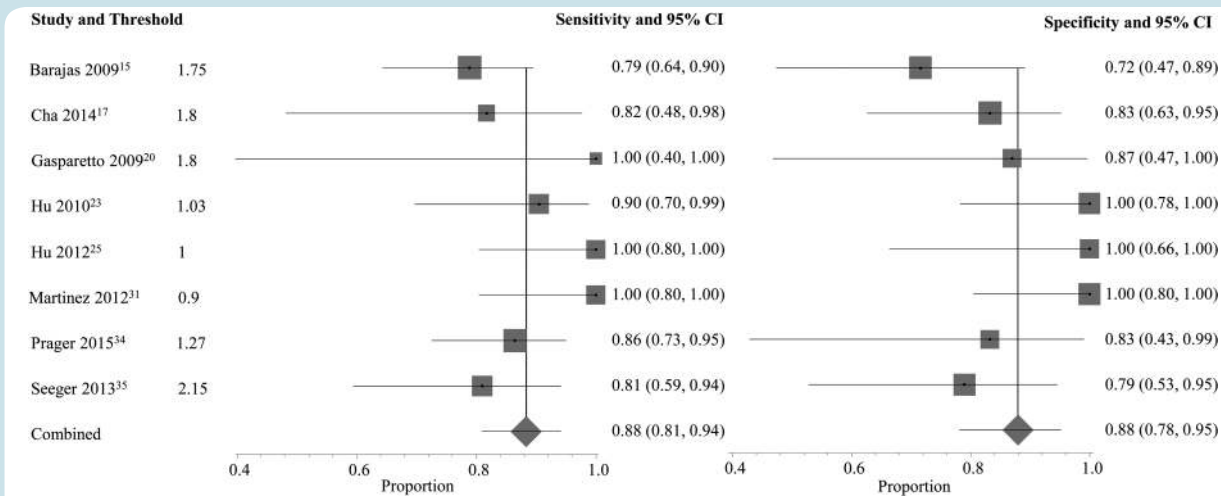


Fig. 3 Individual forest plots of the sensitivity and specificity of mean rCBV. Squares represent the sensitivity or specificity proportion with the size proportional to the weighting of the study. Diamond represents pooled estimate. Lines represent 95% CIs.

is more sensitive to susceptibility effects. Recirculation and leakage of contrast into the extravascular, extracellular space causes T1 relaxation and residual T2/T2* effects that must be taken into consideration.⁶ Although DCE has the potential to overcome many of the shortcomings of DSC, the technique is relatively understudied and therefore is employed less frequently in clinical practice. DCE does suffer from the need for complex pharmacokinetic models to account for nonlinearity between signal intensity and contrast agent concentration. Non-model-based analyses with DCE avoid this problem but have an unclear physiological basis.^{6,8} The result has been the diversity of methods of imaging, processing, and analysis that are evident in the literature. Despite these limitations, our study shows that within individual studies, PWI parameters separate viable

tumor from treatment changes with relatively good sensitivity and specificity using study-specific thresholds. The accuracy is similar among the best performing DSC and DCE parameters from each study in aggregate, the most commonly evaluated parameters, rCBVmean and rCBVmax, and in the pseudoprogression subgroup with sensitivities and specificities predominantly in the 80%–90% range.

However, our study illustrates major limitations of the current literature on quantitative characterization of new posttreatment enhancing lesions by PWI, especially with regard to the reproducibility and generalizability of the techniques utilized given the presence of statistically significant heterogeneity. First, investigators differed in their choice of postprocessing software, which can have

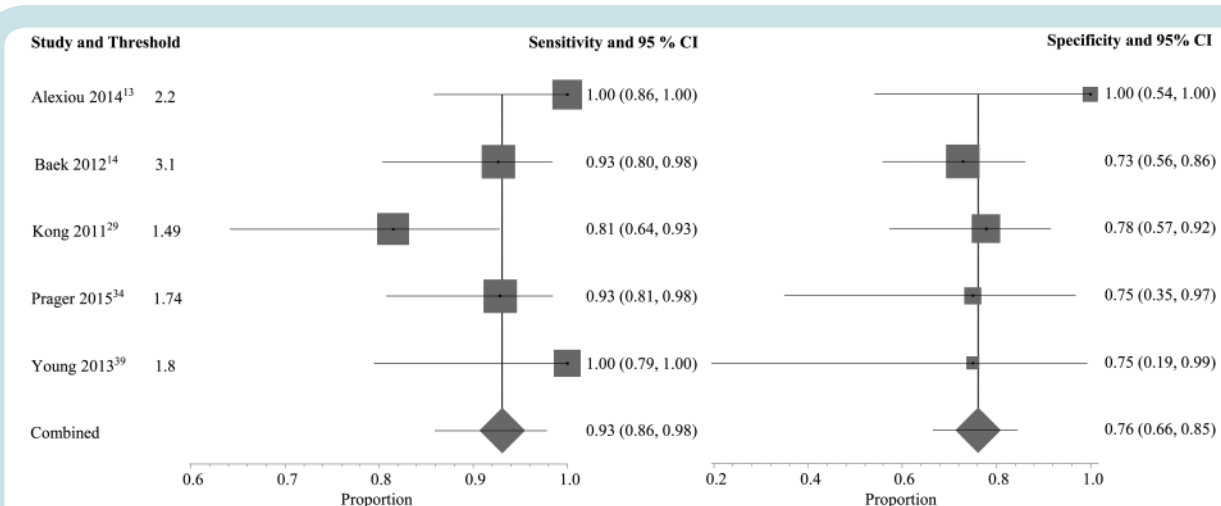


Fig. 4 Individual forest plots of the sensitivity and specificity of maximum rCBV. Squares represent the sensitivity or specificity proportion with the size proportional to the weighting of the study. Diamond represents pooled estimate. Lines represent 95% CIs.

Study	Prospective?	Patient and treatment characteristics compared?	Adequately described the treatment protocol?	Potentially confounding adjuvant treatments?	Did the study avoid inappropriate exclusions?	Interval between treatment completion and index test results interpreted without knowledge of the results of the reference standard?	MRP analysis techniques reproducible as described?	Did more than one investigator process the PWI data?	Evaluation of inter-rater reliability?	Reference standard adequately defined?	Path interpreted without knowledge of results of the index test?	Did patients receive the same reference standard?	All patients included in the analysis?
Alexiou 2014 ¹³	●	●	●	●	●	●	●	●	●	●	●	●	●
Baek 2012 ¹⁴	●	●	●	●	●	●	●	●	●	●	●	●	●
Barajas 2009 ¹⁵	●	●	●	●	●	●	●	●	●	●	●	●	●
Bisdas 2011 ¹⁶	●	●	●	●	●	●	●	●	●	●	●	●	●
Cha 2014 ¹⁷	●	●	●	●	●	●	●	●	●	●	●	●	●
Choi 2013 ¹⁸	●	●	●	●	●	●	●	●	●	●	●	●	●
Chung 2013 ¹⁹	●	●	●	●	●	●	●	●	●	NA	●	●	●
Gasparetto 2009 ²⁰	●	●	●	●	●	●	●	●	●	NA	●	●	●
Hamilton 2015 ²¹	●	●	●	●	●	●	●	●	●	NA	●	●	●
Heo 2015 ²²	●	●	●	●	●	●	●	●	●	●	●	●	●
Hu 2009 ²³	●	●	●	●	●	●	●	●	●	NA	●	●	●
Hu 2010 ²⁴	●	●	●	●	●	●	●	●	●	NA	●	●	●
Hu 2012 ²⁵	●	●	●	●	●	●	●	●	●	NA	●	●	●
Kim 2010 ²⁶	●	●	●	●	●	●	●	●	●	NA	●	●	●
Kim 2014 ²⁷	●	●	●	●	●	●	●	●	●	●	●	●	●
Kim 2014 ²⁸	●	●	●	●	●	●	●	●	●	NA	●	●	●
Kong 2011 ²⁹	●	●	●	●	●	●	●	●	●	●	●	●	●
Mangla 2010 ³⁰	●	●	●	●	●	●	●	●	●	NA	NA	●	●
Martinez 2012 ³¹	●	●	●	●	●	●	●	●	●	●	●	●	●
Narang 2011 ³²	●	●	●	●	●	●	●	●	●	●	●	●	●
Park 2015 ³³	●	●	●	●	●	●	●	●	●	●	●	●	●
Prager 2014 ³⁴	●	●	●	●	●	●	●	●	●	NA	●	●	●
Seeger 2013 ³⁵	●	●	●	●	●	●	●	●	●	NA	NA	●	●
Song 2013 ³⁶	●	●	●	●	●	●	●	●	●	NA	NA	●	●
Suh 2013 ³⁷	●	●	●	●	●	●	●	●	●	●	●	●	●
Thomas 2015 ³⁸	●	●	●	●	●	●	●	●	●	NA	NA	●	●
Young 2013 ³⁹	●	●	●	●	●	●	●	●	●	●	●	●	●
Yun 2015 ⁴⁰	●	●	●	●	●	●	●	●	●	NA	NA	●	●

Fig. 5 Results of the study quality assessment. ● = yes, ● = no, ● = unclear, NA = not applicable; PWI = perfusion weighted imaging.

a significant impact on PWI analysis.^{43,44} Second, they also differed in techniques utilized for contrast leakage and recirculation correction for DSC. Hu et al measured rCBV under different acquisition and postprocessing conditions and found that diagnostic accuracy varied based on the use of contrast preloading and baseline subtraction techniques, illustrating the importance of optimizing these variables for DSC.²³ However, multiple factors affect

optimization, including type of contrast agent, incubation time, pulse-sequence parameters, and correction algorithm, which limits comparison of quantitative metrics across studies. Third, the size, number, location, and method of selection of ROIs similarly limit comparability across studies. For example, normalization to contralateral white matter was variable among DCE studies, and when one DSC study also included a contralateral gray matter

ROI, the result was 2 different proposed threshold values for the same patient cohort.^{23,24} Fourth, perhaps the most apparent source of heterogeneity among studies was the wide array of parameters that were evaluated. Even the most commonly evaluated DSC parameter, normalized rCBV, differed based on whether mean, maximum, or histogram-derived percentile values were obtained. Similar variability was present among DCE parameters, including some semiquantitative metrics (such as maximum slope of enhancement in initial vascular phase³² and initial area under the curve ratio³⁷) that are unlikely to be amenable to routine use unless incorporated into automated commercial software. Our study suggests that the net result of this methodological heterogeneity is the wide range of optimal values reported across studies, making it impossible to find clinically meaningful pooled PWI thresholds to reliably distinguish tumor from treatment effect.

Other limitations of the current literature related to study design include retrospective analysis, small sample sizes, potentially confounding use of anti-angiogenic drugs or steroids, and lack of blinding, raising the possibility of ascertainment bias. Additionally, in many studies there was deficient delineation of the timing of the treatment effect. While the physiology of pseudoprogression and late radiation necrosis could possibly be similar enough for the purposes of PWI, the distinction is clinically important. Pseudoprogression typically spontaneously stabilizes or improves without further treatment and is associated with increased survival, likely reflecting an exuberant positive treatment response.³ Late radiation necrosis, on the other hand, is a severe tissue reaction that often requires treatment to address the associated morbidity.⁶ There was also an absent emphasis on histological confirmation of PWI findings in many studies, a crucial factor given that an adequate reference standard is required to ensure that results have validity. Studies that did provide histological criteria demonstrated the variation inherent in the “gold standard,” which is of paramount importance given that tumor and treatment change often coexist. Mean-derived as well as many histogram-based metrics are expectedly affected by this overlap, possibly reducing their predictive value due to increased variability. Finally, there was evidence of borderline significant publication bias in a minority of the meta-analyses performed, implying that smaller and less precise studies were not published and are thus missing. However, this is less of a concern in meta-analyses of single outcome proportions because no effect estimate is calculated (ie, no comparison between groups is made).

There are several limitations in our study to consider. First, some investigators evaluated multiparametric approaches, which our study does not take into account, noting that these approaches compound the heterogeneity and complexity of techniques in the literature. Second, we did not exclude studies with patients diagnosed with WHO grade III gliomas. While glioblastoma patients constituted the majority even in those studies, differences in tumor biology and treatments may have affected results. Third, we included studies that did not require histological confirmation of PWI findings. This was perhaps less of an issue in studies evaluating pseudoprogression, which is often a clinicoradiological diagnosis, but as mentioned previously, histology provides

greater validity. Fourth, although we included only studies published in the so-called temozolomide era, articles did not always report on specific chemotherapy regimens for patients.

Based on our study, relatively small, individual studies taken in isolation demonstrate promising accuracy in distinguishing treatment effect from tumor using PWI. However, optimal imaging techniques and threshold values remain difficult to identify with highly variable proposed cutoff values, which are potentially useful only as general guides. A particular metric and threshold value that is optimized and validated at a single institution might be sufficient if applied consistently to a patient being followed over time. On the contrary, if a universally applicable approach with threshold values reproducible across institutions is desired, standardization of techniques and postprocessing methods is required alongside a consensus histological definition for recurrent tumor. Further investigation including prospective comparative evaluation of postprocessing methods, additional efforts with the relatively understudied DCE, and the comparative accuracy of a qualitative approach are warranted.

Supplementary Material

Supplementary material is available at *Neuro-Oncology Journal* online (<http://neuro-oncology.oxfordjournals.org/>).

Funding

This work was supported by a Weill Cornell Medical College grant UL1-TR000457-06 to G.A. and P.C.

Conflict of interest statement. None declared.

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 Wick W, Wiestler B, Platten M. Treatment of anaplastic glioma. *Cancer Treat Res*. 2015;163:89–101.
- 3 Hygino da Cruz LC Jr, Rodriguez I, Domingues RC et al. Pseudoprogression and pseudoresponse: imaging challenges in the assessment of post-treatment glioma. *AJNR Am J Neuroradiol*. 2011;32(11):1978–1985.
- 4 Mabray MC, Barajas RF Jr, Cha S. Modern brain tumor imaging. *Brain Tumor Res Treat*. 2015;3(1):8–23.
- 5 Young RJ, Gupta A, Shah AD et al. Potential utility of conventional MRI signs in diagnosing pseudoprogression in glioblastoma. *Neurology*. 2011;76(22):1918–1924.
- 6 Verma N, Cowperthwaite MC, Burnett MG et al. Differentiating tumor recurrence from treatment necrosis: a review of neuro-oncologic imaging strategies. *Neuro Oncol*. 2013;15(5):515–534.

- 7 Jahng GH, Li KL, Ostergaard L et al. Perfusion magnetic resonance imaging: a comprehensive update on principles and techniques. *Korean J Radiol.* 2014;15(5):554–577.
- 8 Paldino MJ, Barboriak DP. Fundamentals of quantitative dynamic contrast-enhanced MR imaging. *Magn Reson Imaging Clin N Am.* 2009;17(2):277–289.
- 9 Yankeelov TE, Gore JC. Dynamic contrast enhanced magnetic resonance imaging in oncology: theory, data acquisition, analysis, and examples. *Curr Med Imaging Rev.* 2009;3(2):91–107.
- 10 Liberati A, Altman DG, Tetzlaff J et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *J Clin Epidemiol.* 2009;62(10):e1–34.
- 11 Whiting P, Rutjes AW, Reitsma JB et al. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol.* 2003;3:25.
- 12 Whiting PF, Rutjes AWS, Westwood ME et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med.* 2011;155(8):529–536.
- 13 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.
- 14 Baek HJ, Kim HS, Kim N et al. 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.
- 15 Barajas RFJr, 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.
- 16 Bisdas S, Naegele T, Ritz R et al. Distinguishing recurrent high-grade gliomas from radiation injury. A pilot study using dynamic contrast-enhanced MR imaging. *Acad Radiol.* 2011;18(5):575–583.
- 17 Cha J, Kim ST, Kim H-J et al. Differentiation of tumor progression from pseudoprogression in patients with posttreatment glioblastoma using multiparametric histogram analysis. *AJNR Am J Neuroradiol.* 2014;35(7):1309–1317.
- 18 Choi YJ, Kim HS, Jahng GH et al. Pseudoprogression in patients with glioblastoma: added value of arterial spin labeling to dynamic susceptibility contrast perfusion MR imaging. *Acta Radiologica.* 2013;54(4):448–454.
- 19 Chung WJ, Kim HS, Kim N et al. Recurrent glioblastoma: optimum area under the curve method derived from dynamic contrast-enhanced T1-weighted perfusion MR imaging. *Radiology.* 2013;269(2):561–568.
- 20 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.
- 21 Hamilton JD, Lin J, Ison C et al. Dynamic contrast-enhanced perfusion processing for neuroradiologists: model-dependent analysis may not be necessary for determining recurrent high-grade glioma versus treatment effect. *AJNR Am J Neuroradiol.* 2015;36(4):686–693.
- 22 Heo YJ, Kim HS, Park JE et al. Uninterpretable dynamic susceptibility contrast-enhanced perfusion mr images in patients with post-treatment glioblastomas: cross-validation of alternative imaging options. *PLoS ONE.* 2015;36(12):2242–2249.
- 23 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.
- 24 Hu LS, Baxter LC, Smith KA et al. Relative cerebral blood volume values to differentiate high-grade glioma recurrence from posttreatment radiation effect: direct correlation between image-guided tissue histopathology and localized dynamic susceptibility-weighted contrast-enhanced perfusion MR imaging measurements. *AJNR Am J Neuroradiol.* 2009;30(3):552–558.
- 25 Hu LS, Eschbacher JM, Heiserman JE et al. Reevaluating the imaging definition of tumor progression: perfusion MRI quantifies recurrent glioblastoma tumor fraction, pseudoprogression, and radiation necrosis to predict survival. *Neuro Oncol.* 2012;14(7):919–930.
- 26 Kim HS, Goh MJ, Kim N et al. Which combination of MR imaging modalities is best for predicting recurrent glioblastoma? Study of diagnostic accuracy and reproducibility. *Radiology.* 2014;273(3):831–843.
- 27 Kim HS, Kim J-H, Kim S-H et al. 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.
- 28 Kim HS, Suh CH, Kim N et al. Histogram analysis of intravoxel incoherent motion for differentiating recurrent tumor from treatment effect in patients with glioblastoma: initial clinical experience. *AJNR Am J Neuroradiol.* 2014;35(3):490–497.
- 29 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.
- 30 Mangla R, Singh G, Ziegelitz D et al. Changes in relative cerebral blood volume 1 month after radiation-temozolomide therapy can help predict overall survival in patients with glioblastoma. *Radiology.* 2010;256(2):575–584.
- 31 Martinez-Martinez A, Martinez-Bosch J. Perfusion magnetic resonance imaging for high grade astrocytomas: can cerebral blood volume, peak height, and percentage of signal intensity recovery distinguish between progression and pseudoprogression? *Radiologia.* 2014;56(1):35–43.
- 32 Narang J, Jain R, Arbab AS et al. Differentiating treatment-induced necrosis from recurrent/progressive brain tumor using nonmodel-based semiquantitative indices derived from dynamic contrast-enhanced T1-weighted MR perfusion. *Neuro Oncol.* 2011;13(9):1037–1046.
- 33 Park JE, Kim HS, Goh MJ et al. 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.
- 34 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. *AJNR Am J Neuroradiol.* 2015;36(5):877–885.
- 35 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.
- 36 Song YS, Choi SH, Park CK et al. True progression versus pseudoprogression in the treatment of glioblastomas: a comparison study of normalized cerebral blood volume and apparent diffusion coefficient by histogram analysis. *Korean J Radiol.* 2013;14(4):662–672.
- 37 Suh CH, Kim HS, Choi YJ et al. Prediction of pseudoprogression in patients with glioblastomas using the initial and final area under the curves ratio derived from dynamic contrast-enhanced T1-weighted perfusion MR imaging. *AJNR Am J Neuroradiol.* 2013;34(12):2278–2286.
- 38 Thomas AA, Arevalo-Perez J, Kaley T et al. Dynamic contrast enhanced T1 MRI perfusion differentiates pseudoprogression from recurrent glioblastoma. *J Neurooncol.* 2015;125(1):183–190.

- 39 Young RJ, Gupta A, Shah AD et al. MRI perfusion in determining pseudoprogression in patients with glioblastoma. *Clin Imaging*. 2013;37(1):41–49.
- 40 Yun TJ, Park CK, Kim TM et al. Glioblastoma treated with concurrent radiation therapy and temozolomide chemotherapy: differentiation of true progression from pseudoprogression with quantitative dynamic contrast-enhanced MR imaging. *Radiology*. 2015;274(3):830–840.
- 41 Shah AH, Snelling B, Bregy A et al. Discriminating radiation necrosis from tumor progression in gliomas: a systematic review what is the best imaging modality? *J Neurooncol*. 2013;112(2):141–152.
- 42 Deng SM, Zhang B, Wu YW et al. Detection of glioma recurrence by 11C-methionine positron emission tomography and dynamic susceptibility contrast-enhanced magnetic resonance imaging: a meta-analysis. *Nucl Med Commun*. 2013;34:758–766.
- 43 Hu LS, Kelm Z, Korfiatis P et al. Impact of software modeling on the accuracy of perfusion MRI in glioma. *AJNR Am J Neuroradiol*. 2015;36(12):2242–2249.
- 44 Kelm ZS, Korfiatis PD, Lingineni RK et al. Variability and accuracy of different software packages for dynamic susceptibility contrast magnetic resonance imaging for distinguishing glioblastoma progression from pseudoprogression. *J Med Imaging (Bellingham)*. 2015;2(2):026001.