Impact of Software Modeling on the Accuracy of Perfusion MRI in Glioma


AJNR Am J Neuroradiol 2015, 36 (12) 2242-2249
doi: https://doi.org/10.3174/ajnr.A4451
http://www.ajnr.org/content/36/12/2242
Impact of Software Modeling on the Accuracy of Perfusion MRI in Glioma


ABSTRACT

BACKGROUND AND PURPOSE: Relative cerebral blood volume, as measured by $T_2^*$-weighted dynamic susceptibility-weighted contrast-enhanced MRI, represents the most robust and widely used perfusion MR imaging metric in neuro-oncology. Our aim was to determine whether differences in modeling implementation will impact the correction of leakage effects (from blood-brain barrier disruption) and the accuracy of relative CBV calculations as measured on $T_2^*$-weighted dynamic susceptibility-weighted contrast-enhanced MR imaging at 3T field strength.

MATERIALS AND METHODS: This study included 52 patients with glioma undergoing DSC MR imaging. Thirty-six patients underwent both non-preload dose– and preload dose–corrected DSC acquisitions, with 16 patients undergoing preload dose–corrected acquisitions only. For each acquisition, we generated 2 sets of relative CBV metrics by using 2 separate, widely published, FDA-approved commercial software packages: IB Neuro and nordiICE. We calculated 4 relative CBV metrics within tumor volumes: mean relative CBV, mode relative CBV, percentage of voxels with relative CBV > 1.75, and percentage of voxels with relative CBV > 1.0 (fractional tumor burden). We determined Pearson ($r$) and Spearman ($\rho$) correlations between non-preload dose– and preload dose–corrected metrics. In a subset of patients with recurrent glioblastoma ($n = 25$), we determined receiver operating characteristic area under the curve for fractional tumor burden accuracy to predict the tissue diagnosis of tumor recurrence versus posttreatment effect. We also determined correlations between rCBV and microvessel area from stereotactic biopsies ($n = 29$) in 12 patients.

RESULTS: With IB Neuro, relative CBV metrics correlated highly between non-preload dose– and preload dose–corrected conditions for fractional tumor burden ($r = 0.96, P = 0.94$), percentage > 1.75 ($r = 0.93, P = 0.91$), mean ($r = 0.87, P = 0.86$), and mode ($r = 0.78, P = 0.76$). These correlations dropped substantially with nordiICE. With fractional tumor burden, IB Neuro was more accurate than nordiICE in diagnosing tumor versus posttreatment effect (area under the curve = 0.85 versus 0.67) ($P < .01$). The highest relative CBV–microvessel area correlations required preload dose and IB Neuro ($r = 0.64, P = 0.58, P = .001$).

CONCLUSIONS: Different implementations of perfusion MR imaging software modeling can impact the accuracy of leakage correction, relative CBV calculation, and correlations with histologic benchmarks.

ABBREVIATIONS: FTB = fractional tumor burden; GBCA = gadolinium-based contrast agents; IBN = IB Neuro; MVA = microvessel area; NICE = nordiICE; PLD = preload dose; pMRI = perfusion MR imaging; rCBV = relative cerebral blood volume

Perfusion MR imaging (pMRI) has emerged as a powerful diagnostic tool in neuro-oncology. Multiple independent studies have shown how measures of microvessel volume, which are linked closely to histologic identity and malignant potential, can facilitate diagnoses that have historically eluded conventional MR imaging.\textsuperscript{1,2} For instance, the metric relative cerebral blood volume (rCBV), as measured by dynamic susceptibility-weighted contrast-enhanced pMRI, can identify high-grade components within nonenhancing glioma,\textsuperscript{5,7} distinguish tumor recurrence from posttreatment effects (ie, pseudoprogression, radiation ne-
PLD as low as 0.05–0.1 mmol/kg. Additionally, modeling correction with most studies showing effective T1 leakage correction with a theoretic dose-dependent risks of nephrogenic systemic fibrosis, the GBCA dose is generally minimized, and removing baseline points collected during the first 5 seconds, after the start of the DSC sequence. All GBCA injections were via a second GBCA injection (0.05-mmol/kg, gadodiamide or gadobenate dimeglumine) by using previously described methods.8,19 In all patients, the PLD amount totaled 0.1 mmol/kg, administered via either a single bolus injection or 2 separate (0.05- mmol/kg) bolus injections, depending on the departmental protocol at the time of imaging. In a subset of patients, we acquired non-PLD-corrected DSC-pMRI data during the initial PLD bolus injection, by using either 0.05- or 0.1-mmol/kg GBCA injections, depending on the clinical perfusion MR imaging protocol used at the time of acquisition. We performed a separate subanalysis to determine the impact of different injection doses as shown in On-line Table 1. All DSC data (gradient-echo echo-planar imaging with TR/TE/flip angle = 1500–2000/20 ms/60°, FOV = 24 × 24 cm, matrix = 128 × 128, 5-mm section, no gap) were acquired during 3 minutes with the bolus injection occurring at the 1-minute mark after the start of the DSC sequence. All GBCA injections were via power injector at 3–5 mL/s, followed by a 20-mL normal saline flush. The final GBCA dose for all patients (irrespective of the method of PLD administration) was 0.15 mmol/kg of body weight.

Perfusion MR Imaging Data Analysis
After transferring all MR imaging data to an off-line workstation and removing baseline points collected during the first 5 seconds, we generated whole-brain rCBV maps by using 2 commonly published commercial software packages: nordicICe (Version 2.3.13) and IB Neuro (Version 1.1), both approved by the US Food and Drug Administration. For NICE, we used all available default options and included leakage correction in all cases. Default options consisted of automatic prebolus baseline selection to define the help work toward consensus and standardization of pMRI methodology.

**MATERIALS AND METHODS**

**Subjects**
We searched our data base (2007–2013) for patients with histopathologically confirmed glioma who had conventional 3T MR imaging with pMRI at 2 different institutions. We included patients in whom the same examination contained 2 separate DSC-pMRI acquisitions (and separate bolus contrast injections) and/or the MR imaging was performed preoperatively for stereotactic resection and/or biopsy within 1 day after imaging. Subjects were pooled from 2 separate institutions: Barrow Neurological Institute at St. Joseph’s Hospital and Medical Center and Mayo Clinic, Arizona. All patient data were anonymized for Health Insurance Portability and Accountability Act compliance. The institutional review board approved our study. All patients undergoing pMRI had estimated glomerular filtration rates of >60 mg/min/1.72 m².

**Perfusion MR Imaging Data Acquisition**
Each 3T examination was performed on 1 of 2 MR imaging magnets (Signa HDx; GE Healthcare, Milwaukee, Wisconsin; or Magnetom Skyra; Siemens, Erlangen, Germany). All patients underwent initial preload dose administration that allowed the acquisition of PLD-corrected DSC-pMRI data, which were all acquired via a second GBCA injection (0.05-mmol/kg, gadodiamide or gadobenate dimeglumine) by using previously described methods.8,19

In this study, we compared 2 commonly published, commercially available implementations of the Boxerman algorithm,17 as integrated within the IB Neuro (IBN, Version 1.1; Imaging Biometrics, Elm Grove, Wisconsin) and nordicICe (NICE, Version 2.3.13; NordicNeuroLab, Bergen, Norway) software packages.8,9,14-18,20,25-28 We present data from a cohort of 52 patients with glioma who underwent DSC-pMRI acquisition at the time of clinical MR imaging. The goals of this study are to determine the equivalency of modeling implementation and rCBV calculation across platforms and to assess whether rCBV variations, if present, will significantly impact correlations with histologic benchmarks. Our overarching goal is to provide information that will help work toward consensus and standardization of pMRI methodology.
We chose the thresholds of 1.0 and 1.75 because of previous studies reporting the biologic significance of these values.6,8,30 On the basis of the rCBV maps generated from NICE and IBN packages, we calculated volume fraction metrics by using the IB Suite and histogram metrics by using custom code written in Matlab (Version R2012a; MathWorks, Natick, Massachusetts). To reduce variability, we used identical normal-appearing white matter ROIs for both software package analyses to generate all rCBV metrics. We calculated multiple previously published rCBV metrics including the following: 1) volume fraction of tumor voxels above the rCBV threshold of 1.75 (percentage >1.75); 2) volume fraction of tumor voxels above the rCBV threshold of 1.0, also known as perfusion MR imaging fractional tumor burden (FTB); 3) histogram mean rCBV; and 4) histogram mode rCBV for all tumor voxels. We chose these thresholds of 1.0 and 1.75 because of previous studies reporting the biologic significance of these values.6,8,30

**Conventional MR Imaging Acquisition and Analysis**

For each examination, we acquired routine conventional contrast-enhanced MR imaging that included pre- and postcontrast T1-weighted spoiled gradient-echo (inversion recovery prep) stereotactic (ie, volumetric) MR imaging datasets (TI/TR/TE = 300/6.8/2.8 ms, matrix = 320 × 224, FOV = 26 cm, section thickness = 2 mm). Tumor volumes were defined as abnormal enhancing tissue by an experienced neuroradiologist (L.S.H.). In nonenhancing glioma, we defined tumor volumes by using T2-weighted stereotactic MR imaging (TR/TE = 4500/82 ms, matrix = 256 × 256, FOV = 26 cm, section thickness = 2 mm).

**Stereotactic Biopsy, Image Coregistration, and Histologic Microvessel Analysis**

Our cohort included a subset of patients in whom neurosurgeons collected an average of 2–3 tissue specimens from each tumor by using stereotactic surgical localization, following the smallest possible diameter craniotomies to minimize brain shift. Biopsies were performed without knowledge of rCBV analyses. Similar to those in previous studies, biopsy locations and neuronavigational coordinates were recorded and coregistered with MR imaging to enable localized rCBV measurement (3 × 3 voxel-sized ROIs) at corresponding biopsy sites.13,31 Multiple biopsy targets in the same patient were separated by a minimum of 2 cm. The neurosurgeon visually validated stereotactic imaging locations with corresponding intracranial anatomic landmarks, such as vascular structures. Stereotactic biopsy samples were sectioned (10-μm thickness), CD-34 stained, and submitted for quantification of total microvessel area (MVA) by using previously published methods.31-34 Corresponding sections were also stained with hematoxylin-eosin per standard protocol. For each CD-34-stained slide, we measured total microvessel area as previously described.31,32,33 Raw data from 7 of these patients were studied previously.31 The current study differs in the following ways: 1) We used commercial software packages and modeling correction to measure rCBV, 2) we determined test performance differences between packages, and 3) we compared PLD against non-PLD conditions.

**Quantification of Histologic Tumor Fraction in Recurrent Glioblastoma Multiforme**

Our cohort included a subset of 25 patients with recurrent glioblastoma multiforme, previously treated with the protocol of Stupp et al.36 We enrolled each of these patients at the time of recurrence, at which time they underwent preoperative MR imaging (including pMRi) for surgical debulking of newly developed or enlarging lesions suspicious for recurrence identified on surveillance contrast-enhanced MR imaging.

Following debulking, we fixed all surgical tissue specimens in 10% formalin, embedded them in paraffin, sectioned them (10 μm), and stained them with hematoxylin-eosin per standard diagnostic protocol at our institution. Two neuropathologists quantified glioblastoma multiforme and/or posttreatment effect elements for all specimens without knowledge of DSC-MR imaging, by simultaneously estimating histologic fractional volume of tumor relative to nonneoplastic treatment-related features, as previously described.8,30,37,38 Features of tumor recurrence38 and posttreatment effect37,39 were quantified and used to determine the histologic tumor fraction from surgical resection material to diagnose either tumor progression (histologic tumor fraction of ≥50%) or posttreatment effect (histologic tumor fraction of <50%) on the basis of group median values. Raw data from these 25 patients have been studied previously.8 Like the prior study, the current study measures FTB but with several important differences in experimental design: 1) We used and compared 2 separate modeling algorithm implementations to calculate FTB, 2) we assessed performance differences between methods by comparing test accuracies (with receiver operating characteristic analysis), and 3) we use a simplified classification system to establish the clinical presence/absence of tumor progression.

**Statistical Analysis**

A biostatistician performed all analyses. We first determined Pearson and Spearman correlations between non-PLD- and PLD-corrected conditions for all rCBV metrics as calculated by IBN and NICE. Second, we used receiver operating characteristic anal-
ysis to determine the accuracy of FTB (as measured by IBN and NICE) to diagnose tumor versus posttreatment effect. Finally, we determined Pearson and Spearman correlations between local-ized rCBV and MVA from corresponding stereotactic biopsies. $P < .05$ was statistically significant.

RESULTS

Subjects and Tumor Types
We enrolled 52 patients (17 women, 35 men; mean age, 53 years), of whom 87% (45/52) had high-grade gliomas with 78% (35/45) presenting at recurrence after standard multimodal therapy. On-line Table 2 summarizes the tumor types for primary and recurrent cases.

Comparing rCBV Measurements in the Presence and Absence of Preload Dose
Comparing rCBV between PLD and non-PLD conditions gives an indication of how well modeling implementation corrects T1 leakage errors. We acquired both PLD- and non-PLD-corrected rCBV values in a subset of patients ($n = 36$) for whom we calculated 4 separate rCBV metrics (mean, mode, percentage >1.75, and FTB) by using both IB Neuro and nordicICE software packages. When we used IBN (Fig 1), rCBV thresholding metrics correlated very highly between non-PLD- and PLD-corrected conditions (FTB: $r = 0.96$, $\rho = 0.94$; percentage >1.75: $r = 0.93$, $\rho = 0.91$); correlations were also high for mean rCBV ($r = 0.87$, $\rho = 0.86$) and mode rCBV ($r = 0.78$, $\rho = 0.76$). With NICE modeling, these correlations dropped substantially (Fig 1) for thresholding metrics (FTB: $r = 0.70$, $\rho = 0.71$; percentage >1.75: $r = 0.59$, $\rho = 0.60$), mean rCBV ($r = 0.43$, $\rho = 0.62$), and mode rCBV ($r = 0.51$, $\rho = 0.65$). When we added $\gamma$ variate fitting, correlations for mean rCBV by using NICE decreased though the other metrics remained largely unchanged (Table 1). On visual inspection of thresholding maps, non-PLD and PLD-corrected voxels showed greater spatial correspondence when using IBN compared with NICE (Fig 2). Table 1 summarizes correlations for all conditions.

![Figure 1](https://via.placeholder.com/150)

**FIG 1.** A–D. Scatterplots correlating rCBV metrics with and without preload dosing (PLD), as measured by 2 separate modeling algorithms (IBN, NICE without $\gamma$ variate fitting). PLD- and non-PLD corrected values are shown in the x- and y-axes, respectively. Overall, IBN measurements demonstrate consistently higher Pearson ($r$) and Spearman ($\rho$) correlations for mean rCBV, mode rCBV, fractional tumor burden (FTB), and percentage of voxels >1.75. The thresholding metrics (FTB, percentage >1.75) correlate most strongly between PLD- and non-PLD-corrected conditions.
Statistical significance is between PLD and non-PLD metrics (compared with NICE), suggesting higher rCBV accuracy in the absence of PLD correction. The study results here underscore the importance of how soft-

### Table 1: Pearson (r) and Spearman (ρ) correlations between rCBV metrics under PLD-corrected and non-PLD-corrected conditions, as measured by IBM and NICE perfusion software algorithms

<table>
<thead>
<tr>
<th>rCBV Metric</th>
<th>PLD (IBM)</th>
<th>Non-PLD vs PLD (NICE) with gvf</th>
<th>Non-PLD vs PLD (NICE) without gvf</th>
<th>PLD (IBM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>r = 0.87</td>
<td>&lt;.001</td>
<td>0.54</td>
<td>ρ = 0.43</td>
</tr>
<tr>
<td></td>
<td>ρ = 0.86</td>
<td>&lt;.001</td>
<td>0.52</td>
<td>ρ = 0.62</td>
</tr>
<tr>
<td></td>
<td>r = 0.78</td>
<td>&lt;.001</td>
<td>0.01</td>
<td>ρ = 0.51</td>
</tr>
<tr>
<td></td>
<td>ρ = 0.76</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
<td>ρ = 0.65</td>
</tr>
<tr>
<td>% &lt; 175</td>
<td>r = 0.93</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
<td>r = 0.59</td>
</tr>
<tr>
<td></td>
<td>ρ = 0.91</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
<td>ρ = 0.60</td>
</tr>
<tr>
<td>FTB</td>
<td>r = 0.96</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
<td>r = 0.70</td>
</tr>
<tr>
<td></td>
<td>ρ = 0.94</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
<td>ρ = 0.71</td>
</tr>
</tbody>
</table>

Note: gvf indicates γ-variate fitting. NICE calculations were performed with and without gvf. IBM modeling shows substantially higher correlation between PLD and non-PLD metrics (compared with NICE), suggesting higher rCBV accuracy in the absence of PLD correction. Statistical significance is P value < .05.

**FIG 2.** Image of a representative case in a 39-year-old patient with recurrent high-grade ganglioglioma shows an enhancing mass (A). Color overlay percentage ≥ 175 thresholding maps (B–E) depict orange voxels with high rCBV > 1.75, compared with intermediate yellow voxels (rCBV 1.0–1.75) and low green voxels (rCBV < 1.0). With NICE, both spatial distribution and percentage of orange voxels show high discrepancy between non-PLD: (70%, B) and PLD-corrected (35%, C) maps. With IBM, the percentage of orange voxels on the non-PLD map (54%, D) approximates that on the PLD-corrected map (51%, E) with high spatial congruence.

### DISCUSSION

Relative CBV represents the most robust and widely used perfusion MR imaging metric in neuro-oncology. Of the techniques that measure rCBV, DSC is the most commonly used method because of wide availability, straightforward postprocessing, and easy-to-use software programs. DSC uses the indicator dilution theory based on susceptibility (T2*-T2*+weighted signal drop) from first-pass transit of a single GBCA bolus injection. DSC assumes an intact BBB with no extravascular GBCA leakage or recirculation and thus requires correction methods when these factors occur (discussed below). Dynamic contrast-enhancement MRI and arterial spin-labeling offer alternative approaches to DSC for calculation of rCBV. The theory and limitations of these techniques have been described previously.

Correctly performing DSC requires several technical considerations based on comparison data from prior studies validating optimal conditions for best practice. First, DSC-pMRI generally necessitates both PLD and mathematical modeling to achieve the highest degree of T1 leakage correction and rCBV accuracy. Results from our study support this requirement (Table 2). Regarding PLD amount, most groups use a single dose (0.1 mmol/kg) of GBCA, particularly at 1.5T, though adequate PLD correction could be achieved with a GBCA dose as low as 0.05 mmol/kg at 3T. Second, gradient-echo T2*-weighted DSC represents the most preferred and widely published method for DSC. While spin-echo T2-weighted DSC offers a higher signal-to-noise ratio and fewer susceptibility artifacts, double or triple GBCA injection doses (0.2–0.3 mmol/kg) are typically needed during the acquisition of spin-echo DSC to overcome the lower contrast-to-noise ratio (ie, signal drop in response to the GBCA first-pass bolus). Compared with spin-echo, gradient-echo DSC offers advantages such as the following: 1) superior contrast-to-noise ratio (ie, greater signal drop during GBCA first-pass), which allows lower contrast dosage during DSC acquisition (0.05–0.1 mmol/kg) and improves the quality of rCBV data, minimizing the need for signal denoising; 2) greater sensitivity to microvessels of all sizes (including larger tortuous glomeruloid-type vessels commonly observed in high-grade gliomas); and 3) the ability to use flip angles of <90° to minimize T1 leakage effects. Finally, in regard to mathematical modeling, the algorithm published by Boxerman et al remains the most highly cited and validated method to date and has been implemented commercially for widespread use.

The study results here underscore the importance of how soft-
Table 2: Correlations between rCBV and fractional MVA under different PLD and modeling conditions

<table>
<thead>
<tr>
<th>Conditions for rCBV Measurement</th>
<th>Pearson Correlation (r)</th>
<th>P Value</th>
<th>Spearman Correlation (p)</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fractional MVA</td>
<td>1.00</td>
<td>–</td>
<td>1.00</td>
<td>–</td>
</tr>
<tr>
<td>No PLD (IBN)</td>
<td>0.46</td>
<td>.02</td>
<td>0.33</td>
<td>.12</td>
</tr>
<tr>
<td>No PLD (NICE + gvf)</td>
<td>0.51</td>
<td>.01</td>
<td>0.26</td>
<td>.19</td>
</tr>
<tr>
<td>No PLD (NICE – gvf)</td>
<td>0.35</td>
<td>.10</td>
<td>0.18</td>
<td>.39</td>
</tr>
<tr>
<td>PLD (IBN)</td>
<td>0.64</td>
<td>&lt;.001</td>
<td>0.58</td>
<td>.001</td>
</tr>
<tr>
<td>PLD (NICE + gvf)</td>
<td>0.53</td>
<td>&lt;.01</td>
<td>0.28</td>
<td>.15</td>
</tr>
<tr>
<td>PLD (NICE – gvf)</td>
<td>0.59</td>
<td>.001</td>
<td>0.40</td>
<td>.04</td>
</tr>
</tbody>
</table>

Note: + indicates with, –, without; –, not applicable; gvf, γ variate fitting.
*Both PLD correction and IBN software modeling were needed to achieve maximal correlation.

FIG 3. Receiver operating characteristic curves for fractional tumor burden to predict histopathology (tumor versus posttreatment effect) in patients with recurrent glioblastoma multiforme (n = 25). FTB by IB Neuro (blue) demonstrates a significantly larger area under the curve (AUC) compared with nordicICE (without γ variate fitting, orange) FTB measurements (0.85 versus 0.70, P < .01), suggesting that different modeling algorithms can impact the accuracy in predicting histopathologic diagnosis. Adding γ variate fitting further reduces NICE estimates of FTB (0.67, green).

In this study, we chose to validate rCBV measurements against histopathology rather than outcomes. Imaging measurements such as rCBV are most directly related to histologic correlates such as microvessel volume and histologic identity (eg, tumor grade, tumor versus posttreatment effect). How these histologic features (and their imaging correlates) predict survival may be confounded by a number of different factors such as age, molecular markers (isocitrate dehydrogenase), methylation status (eg, O6-methylguanine-DNA methyltransferase), extent of resection, salvage therapy at the time of recurrence, and so forth.5-16,47-49 While clinical outcomes are desirable as end points, they must be correlated with imaging and histologic features together in a controlled trial with a larger patient cohort, which is beyond the scope of this article. Our purpose in this study was simply to determine which method of rCBV measurement (ie, which software package) came closest to informing histopathologic benchmarks. We think that this context justifies the rationale for validating rCBV against histopathologic benchmarks.

We recognize potential study limitations. First, we limited the scope of the evaluation to 2 specific software packages, though many commercial options exist. We simplified the project to maximize the potential clinical impact because we evaluated the most published and validated modeling algorithm to date (Boxerman method,8,9,14-21,24,26-28 and we eval-

While we observed strong correlations between non-PLD and PLD measures (when using IBN), further studies are likely needed to determine the following: 1) whether PLD can or should be omitted, 2) what the optimal conditions would be to allow PLD omission (ie, modeling implementation, 3T field strength), and 3) whether this omission would significantly impact prognostic and diagnostic accuracy. Under PLD conditions, separate experiments confirmed significantly higher FTB accuracy with IBN (area under the curve = 0.84), compared with NICE (area under the curve = 0.67, p < 0.01), in diagnosing histopathologically confirmed tumor versus posttreatment effect (ie, pseudoprogression, radiation necrosis). IBN also provided the highest degree of correlation between localized rCBV and tissue microvessel area (Table 2).
REFERENCES


15. Schmaida KM, Prah M, Connelly J, et al. Dynamic-susceptibility contrast agent MRI measures of relative cerebral blood volume pre-
dict response to bevacizumab in recurrent high-grade glioma. Neuro Oncol 2014;16:880–88 CrossRef Medline


