Department of Radiology & Diagnostic Medical Imaging

Clinical Validation of Automatable Gaussian Normalized CBV in Brain Tumor Analysis: Superior Reproducibility and Slightly Better Association with Survival than Current Standard Manual Normal Appearing White Matter Normalization.

Lei Qin

Xiang Li

Angie Li

Suchun Cheng

Jinrong Qu

See next page for additional authors

Authors

Lei Qin, Xiang Li, Angie Li, Suchun Cheng, Jinrong Qu, Katherine Reinshagen, Jiani Hu, Nathan Himes MD, Gao Lu, Xiaoyin Xu, and Geoffrey S Young

Clinical Validation of Automatable Gaussian Normalized CBV in Brain Tumor Analysis: Superior Reproducibility and Slightly Better Association with Survival than Current Standard Manual Normal Appearing White Matter Normalization[1,2]

Lei Qin[*,†,3], Xiang Li[‡,§,3], Angie Li[‡,¶], Suchun Cheng[#], Jinrong Qu[‡,§], Katherine Reinshagen[†,‡,**], Jiani Hu[#], Nathan Himes[‡,††], Gao Lu[‡,‡‡], Xiaoyin Xu[‡,‡‡] and Geoffrey S. Young[*,†,‡]

[*]Dana-Farber Cancer Institute, Department of Imaging, Boston, MA, USA; [†]Harvard Medical School, Department of Radiology, Boston, MA, USA; [‡]Brigham and Women's Hospital, Department of Radiology, Boston, MA, USA; [§]Affiliated Cancer Hospital of Zhengzhou University, Department of Radiology, Zhengzhou, Henan, China; [¶]The Robert Larner, M.D. College of Medicine at the University of Vermont, Burlington, VT, USA; [#]Dana-Farber Cancer Institute, Department of Biostatistics and Computational Biology, Boston, MA, USA; [**]Massachusetts Eye and Ear Infirmary, Department of Radiology, Boston, MA, USA; [††]Medical Imaging of Lehigh Valley, Lehigh Valley Hospital, Allentown, PA, USA; [‡‡]Peking Union Medical College Hospital, Department of Neurosurgery, Beijing, China

Abstract

PURPOSE: To validate Gaussian normalized cerebral blood volume (GN-nCBV) by association with overall survival (OS) in newly diagnosed glioblastoma patients and compare this association with current standard white matter normalized cerebral blood volume (WN-nCBV). *METHODS:* We retrieved spin-echo echo-planar dynamic susceptibility contrast MRI acquired after maximal resection and prior to radiation therapy between 2006 and 2011 in 51 adult patients (28 male, 23 female; age 23-87 years) with newly diagnosed glioblastoma. Software code was developed in house to perform Gaussian normalization of CBV to the standard deviation of the whole brain CBV. Three expert readers manually selected regions of interest in tumor and normal-appearing white matter on CBV maps. Receiver operating characteristics (ROC) curves associating nCBV with 15-month OS were calculated for both GN-nCBV and WN-nCBV. Reproducibility and interoperator variability were compared using within-subject coefficient of variation (wCV) and intraclass correlation coefficients (ICCs). *RESULTS:* GN-nCBV ICC (≥ 0.82) and wCV ($\leq 21\%$) were superior to WN-nCBV ICC (0.54-0.55) and wCV ($\geq 46\%$). The area under the ROC curve analysis demonstrated both GN-nCBV and WN-nCBV to be good predictors of OS, but GN-nCBV was consistently superior, although the difference was not statistically significant. *CONCLUSION:* GN-nCBV has a slightly better association with clinical gold standard OS than conventional WM-nCBV in our glioblastoma patient cohort. This equivalent or superior validity, combined with the advantages of higher reproducibility, lower interoperator variability, and easier automation, makes GN-nCBV superior to WM-nCBV for clinical and research use in glioma patients. We recommend widespread adoption and incorporation of GN-nCBV into commercial dynamic susceptibility contrast processing software.

Translational Oncology (2018) 11, 1398–1405

Introduction

Dynamic susceptibility contrast (DSC) MRI estimates of brain tumor cerebral blood volume (CBV) reflect tumor vascularity and neoangiogenesis [1], are predictive of glioma grade and survival [2–9], and aid in assessment of treatment response [10,11] and differentiation of pseudoprogression from true tumor progression [12,13]. DSC detects the transient decrease in signal intensity (ΔSI (t)) on continuously acquired echo-planar T2 or T2*-weighted images caused by passage of bolus gadolinium contrast through the brain capillaries. Integrating the area under the transverse relaxation rate (ΔR2 or ΔR2*) curve derived from the ΔSI curve yields the CBV for each voxel [14].

In addition to the number, size, and distribution of vessels within each voxel, CBV estimates vary with intravascular concentration, dispersion, delay, flow rate, choice of acquisition parameters [gradient vs spin-echo, repetition time (TR), echo time (TE), flip angle, contrast agent, contrast dose, leakage-reduction preload etc.] and choice of postprocessing algorithm that may or may not include gamma-variate fitting [15], baseline subtraction [16], and leakage correction [10,17]. To compensate, white matter normalization (WN) is typically performed whereby tumor CBV is divided by the mean CBV in a region of interest (ROI) selected manually in the contralateral normal-appearing white matter (NAWM), yielding a unitless "normalized CBV" (nCBV) ratio [14,17–20].

The few published studies suggest substantial coefficient of variation (CV) in NAWM ROI CBV measurements including test-retest CV of 12%-14% in healthy volunteers [21] and an interscanner CV of 25%-30% in glioma patients [22]. The roughly 20% variation in NAWM measurements is likely responsible for a substantial part of the intra- and interobserver CV of 30%-41% reported in white matter normalized brain lesion nCBV [18]. Methods proposed to reduce this variation in CBV estimates include standardization [23], Z-score normalization [22], and Gaussian normalization (GN) [22]. In GN, the tumor ROI CBV is normalized to the standard deviation (SD) of CBV throughout the whole brain rather than ROI measurements of NAWM CBV. GN eliminates completely the subjectivity of NAWM ROI selection, reduces operator time, makes automation simpler, and has been reported to provide the lowest CV in NAWM and the highest tumor contrast for glioblastoma (GBM) [21].

Although GN decreases CV in NAWM, its effect on tumor nCBV estimates has not been studied in detail. Whether glioma GN-nCBV is as valid as current standard WM-nCBV remains to be established [24]. To address this, we compared GN-nCBV association with OS to the known strong OS association of WN-nCBV in a well-characterized cohort of newly diagnosed GBM.

Methods

Human Subjects

The study was in compliance with the Health Insurance Portability and Accountability Act and approved by our institutional review board. Informed consent was waived for this retrospective study.

For this study, 51 patients were retrieved from a database of adult glioma patients newly diagnosed between 2006 and 2011 (28 men, 23 women; age mean: 56.6 years and range: 23-87 years). All patients had WHO grade IV glioblastoma, known survival, and SE-EPI DSC acquired after maximal surgical resection and before radiation therapy.

Data Acquisition

Thirty-five patients underwent 1.5-T MRI and 16 patients underwent 3-T MRI on whole body MRI scanners (GE Medical Systems, Milwaukee, WI). Axial DSC was performed utilizing a series of SE-EPI images (scan parameters: 1900-2000 milliseconds TR/80 milliseconds TE, 128*128 matrix size; 10-mm slice thickness; 40 time points) acquired 10 seconds prior to, during, and after intravenous administration of gadopentetate dimeglumine (Magnevist, Bayer Healthcare) with a power injector at a rate of 4 ml/s, followed by 20-ml saline flush. Double-dose (0.2 mmol/kg) contrast was used if acquisition was performed with 1.5 T, while one and half dose (0.15 mmol/kg) was used for 3 T, with a maximum 30 ml.

Image Analysis

The dynamic source images were visually inspected for to exclude datasets with substantial patient motion. No substantially motion-degraded scans were detected. DSC analysis was performed using the Functool software package on Advantage Window workstation (GE Medical Systems). A lower threshold was manually adjusted to remove the background noise. All pixels with lower intensity than the threshold were removed. The remaining pixels define the brain volume used to calculate whole brain SD for GN. The beginning and end of the bolus passage were defined on the time-intensity curve to set the integration range for calculation of CBV maps.

Three readers (clinical radiologists with 15, 5, and 5 years of experience, respectively) independently selected tumor and NAWM ROIs directly on the resulting CBV maps. Each reader picked three tumor ROIs on the high CBV spots and another three ROIs on the contralateral NAWM. The radius of these ROIs was 1-2 image pixels (2-4 mm). Because it has been previously shown that the maximum or mean CBV of several tumor ROIs generates better intra- and interobserver reproducibility than a single ROI and that the maximum is slightly better than the mean [18], we calculated both the maximum and the mean of the three tumor ROIs. For NAWM, the mean of the three ROIs was calculated and used for normalization of the measured tumor CBV values.

To assess the influence of variation in ROI size on NAWM estimates, one of the three readers (reader 2) picked one large ROI (radius = 10 mm) on NAWM for each patient, and the mean CBV inside the ROI was recorded for normalization. This second evaluation was performed more than 2 months later than the first evaluation to avoid recall bias. A fourth reader (with clinical experience of 14 years) picked a single large ROI in the NAWM using the same radius independently.

GN was performed using Matlab (Mathworks, Natick, MA) code produced in-house [22]. The SD of the whole brain CBV map, $\sigma_{CBV_{Whole\ Brain}}$, was calculated and used for normalization:

$$nCBV_G = \frac{CBV}{\sigma_{CBV_{Whole\ Brain}}} \tag{1}$$

Conventional NAWM ROIs are selected to exclude tumor, but whole brain SD used in GN includes tumor. To examine whether the presence of variable amounts of tumor introduces significant variation between patients in whole brain SD, we selected 10 patients with the largest volume of enhancing tumor in this cohort. One investigator manually contoured the ROIs slice by slice on both contrast-enhanced pGd-T1WI and fluid-attenuated inversion recovery T2WI

Table 1. Summary Statistics of Tumor CBV^{mean}, Tumor CBV^{max}, and NAWM CBV^{mean}

Reader	ROI Size	Tumor CBV^{mean}		Tumor CBV^{max}		NAWM CBV^{mean}	
1 [*]	Small (2-4 mm)	125.19 ± 85.04	$P = .002$ [¶]	146.61 ± 102.22	$P < .001$ [¶]	31.41 ± 11.38	$P = .04$ [¶]
2 [*]		135.95 ± 83.92		160.39 ± 100.39		30.93 ± 12.62	
3 [*]		140.87 ± 86.48		172.00 ± 100.85		27.95 ± 14.95	
(1-2) [†]		−10.76 ± 3.59 [§]	$P = .004$	−13.78 ± 4.28 [§]	$P = .002$	0.48 ± 1.12	$P = .67$
(1-3) [†]		−15.68 ± 4.88 [§]	$P = .002$	−25.39 ± 6.44 [§]	$P < .001$	3.45 ± 1.67 [§]	$P = .044$
(2-3) [†]		−4.91 ± 4.84	$P = .31$	−11.61 ± 6.86	$P = .10$	2.98 ± 1.64	$P = .075$
2 [*,‡]	Large (10 mm)	N/A		N/A		31.26 ± 12.14	
4 [*]		N/A		N/A		36.37 ± 13.70	
(2 [‡]-4) [†]		N/A		N/A		−5.11 ± 1.23 [§]	$P < .001$

[*] Data are represented as mean ± SD.
[†] Data are represented as mean ± SE.
[‡] Data from reader 2 in big ROIs.
[§] The measurements from the two readers differed statistically (paired t tests).
[¶] Differences among three readers were assessed by mixed-effects models.

(FLAIR-T2). The volume of enhancing and nonenhancing tumor in these ROIs was calculated as a percentage of whole brain volume. After the two sets of ROIs were copied to the aligned CBV maps, the whole brain SD was calculated after first excluding contrast-enhancing tumor ($\sigma_{CBV_{Whole\ Brain\ -\ enhancing}}$) and then after excluding both enhancing and nonenhancing tumor ($\sigma_{CBV_{Whole\ Brain\ -\ FLAIR\ nonenhancing}}$). The percentage change in SD resulting from exclusion of the tumor ROI was calculated as:

$$SD\ change = \frac{\left(\sigma_{CBV_{Whole\ Brain-enhancing}} - \sigma_{CBV_{Whole\ Brain}}\right)}{\sigma_{CBV_{Whole\ Brain}}} \times 100\% \quad (2A)$$

Or

$$SD\ change = \frac{\left(\sigma_{CBV_{Whole\ Brain-FLAIR\ nonenhancing}} - \sigma_{CBV_{Whole\ Brain}}\right)}{\sigma_{CBV_{Whole\ Brain}}} \times 100\% \quad (2B)$$

Because the ROI size subanalysis demonstrated that the small NAWM ROI mean CBV was similar to the large NAWM ROI CBV (see results section), only small ROI NAWM estimates were used for nCBV normalization. Twelve nCBVs in total were generated: $nCBV_i^{mean} = \frac{tumor\ CBV_i^{mean}}{NAWM\ CBV_i^{mean}}$; $nCBV_{Gi}^{mean} = \frac{tumor\ CBV_i^{mean}}{\sigma_{CBV\ Whole\ Brain}}$; and $nCBV_i^{max} = \frac{tumor\ CBV_i^{max}}{NAWM\ CBV_i^{mean}}$, $nCBV_{Gi}^{max} = \frac{tumor\ CBV_i^{max}}{\sigma_{CBV\ Whole\ Brain}}$, where i = reader 1,2,3.

Statistical Analysis

Tumor CBV$_i^{mean}$, Tumor CBV$_i^{max}$, NAWM CBV$_i^{mean}$ inside the selected ROIs, nCBV$_i^{mean}$ and nCBV$_i^{max}$, nCBV$_{Gi}^{mean}$ and nCBV$_{Gi}^{max}$ measurements were summarized by means and SDs for i = reader 1,2,3. Pairwise comparisons among readers in these outcome measurements were summarized by mean differences and corresponding standard errors (SEs) and assessed by paired t tests, separately. For reader 2 who selected both NAWM large ROIs and small ROIs, a paired t test was also used to evaluate the effect of ROI size. Mixed-effects models were built for each outcome variable with readers as fixed effects to evaluate differences among the three readers and subjects as random effects to account for correlation among measurements from the three readers on the same subject. Only the small ROI measurements were used in building mixed-effects models.

The between-subject standard deviation (bSD), within-subject standard deviation (wSD), repeatability coefficient (RC), within-subject coefficient of variation (wCV), intraclass correlation coefficient (ICC), and concordance correlation coefficient (CCC) were calculated based on variance components of the aforementioned mixed-effects models [25] to compare repeatability across readers..

Bland-Altman plots provided an intuitive methodology using the concept of limits of agreement for assessing agreement between two readers. Lastly, we calculated the receiver operating characteristics (ROC) curves to associate nCBV measurements with 15-month overall survival (OS) based on the reported 12-15 month median post-operative OS in GBM patients [26,27]. We evaluated the strength of association by the estimated area under the ROC curve (AUC). A $P \leq .05$ was considered to be statistically significant. Statistics were computed using Stata (Stata v14, StataCorp., College Station, TX), with the exception that CCC was calculated in R-package [28].

Results

Summary statistics of reader-specific Tumor CBV^{mean}, Tumor CBV^{max}, and NAWM CBV^{mean} measurements are listed in Table 1 and plotted in Figure 1. All three measured parameters using small ROIs showed
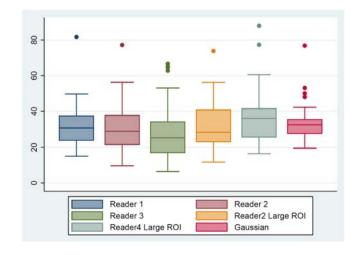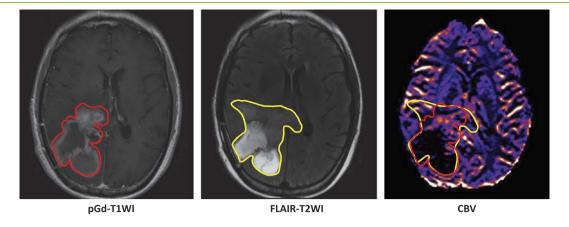

Figure 1. Box plots of NAWM mean CBV measurements among different readers, varying ROI radius size, and the SDs of the brains used for normalization in GN.

pGd-T1WI   **FLAIR-T2WI**   **CBV**

Figure 2. Conventional images and perfusion imaging CBV map demonstrating ROIs manually contoured on contrast-enhanced T1WI and FLAIR-T2WI, respectively, and copied to CBV map.
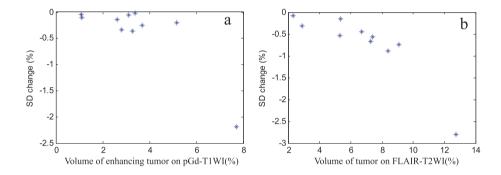

Figure 3. Change (%) in SD of the whole brain CBV as a function of the excluded volume of contrast-enhancing tumor (A) and enhancing and nonenhancing tumor (B).

statistically significant differences among readers. The differences came largely from reader 1 in Tumor CBV$^{\mathrm{mean}}$ and Tumor CBV$^{\mathrm{max}}$, and between readers 1 and 3 in NAWM CBV$^{\mathrm{mean}}$. The NAWM measurements using large ROI also showed a significant difference between readers 2 and 4. The variations in NAWM mean measurements were similar regardless of ROI size (Table 1). This is illustrated by the data from reader 2 who performed measurements using both small and large ROIs while achieving comparable means and SDs ($P = .80$). This suggests

that a larger ROI radius does not offer better precision. Therefore, in order to reduce the overall number of comparisons, when calculating nCBV for testing the effects of normalization, we only used the NAWM CBV$^{\mathrm{mean}}$ measured from the small ROI measurements.

The substantially smaller variation among patients resulting from $\sigma_{CBV_{Whole\ Brain}}$ used in GN compared to manually selected NAWM CBV$^{\mathrm{mean}}$ used in WN is illustrated in Figure 1.

Figure 2 shows an example of ROIs selected on T1 and FLAIR images, and the same ROIs copied onto the CBV map. These two ROIs were excluded to calculate SD change (%) using Eq. (2A or B). Figure 3 shows the SD change when enhancing and nonenhancing lesions were excluded in the 10 selected patients. The x-axis of Figure 3 is the tumor volume percentage in the whole brain. The SD variation was less than 1% for all

Table 2. Comparison of nCBV Measurements Derived from NAWM Normalization (WN) Versus GN

Reader	nCBV$^{\mathrm{mean}}$		nCBV$^{\mathrm{max}}$	
WN 1[*]	4.04 ± 2.26	$P < .001$[§]	4.69 ± 2.62	$P < .001$[§]
WN 2[*]	4.70 ± 3.09		5.58 ± 3.55	
WN 3[*]	6.53 ± 4.95		7.85 ± 5.67	
WN (1-2)[†]	−0.66 ± 0.33	$P = .051$	−0.88 ± 0.37[‡]	$P = .02$
WN (1-3)[†]	−2.50 ± 0.54[‡]	$P < .001$	−3.16 ± 0.62[‡]	$P < .001$
WN (2-3)[†]	−1.83 ± 0.56[‡]	$P = .002$	−2.27 ± 0.63[‡]	$P < .001$
GN 1[*]	3.65 ± 1.93	$P = .004$[§]	4.26 ± 2.27	$P < .001$[§]
GN 2[*]	3.99 ± 1.94		4.73 ± 2.31	
GN 3[*]	4.12 ± 1.92		5.06 ± 2.33	
GN (1-2)[†]	−0.34 ± 0.12[‡]	$P = .005$	−0.46 ± 0.15[‡]	$P = .003$
GN (1-3)[†]	−0.47 ± 0.15[‡]	$P = .004$	−0.80 ± 0.21[‡]	$P < .001$
GN (2-3)[†]	−0.13 ± 0.16	$P = .40$	−0.34 ± 0.22	$P = .14$

[*] Data are represented as mean ± SD.
[†] Data are represented as mean ± SE.
[‡] nCBV measurements between two readers were significantly different (paired t tests).
[§] Differences among three readers were assessed by mixed-effects models.

Table 3. ICC, CCC, wCV, bSD, wSD, and RC for Tumor CBV$^{\mathrm{mean}}$, Tumor CBV$^{\mathrm{max}}$, and NAWM CBV$^{\mathrm{mean}}$ and nCBV Measurements Among Three Independent Readers

	ICC	CCC	RC	wCV	bSD	wSD
Tumor CBV$^{\mathrm{mean}}$	0.93	0.92	62.60	16.87%	82.10	22.60
Tumor CBV$^{\mathrm{max}}$	0.91	0.90	83.43	18.86%	96.57	30.12
NAWM CBV$^{\mathrm{mean}}$	0.66	0.66	20.94	25.11%	10.66	7.56
nCBV$^{\mathrm{mean}}$	0.54	0.48	6.82	48.39%	2.64	2.46
nCBV$^{\mathrm{max}}$	0.55	0.48	7.74	46.28%	3.06	2.80
nCBV$_{\mathrm{G}}^{\mathrm{mean}}$	0.86	0.85	1.99	18.49%	1.79	0.72
nCBV$_{\mathrm{G}}^{\mathrm{max}}$	0.82	0.79	2.74	21.10%	2.08	0.99

A value <0.40 was considered poor, values between 0.40 and 0.59 were considered fair, values between 0.60 and 0.74 were considered good, and values >0.75 was considered excellent.
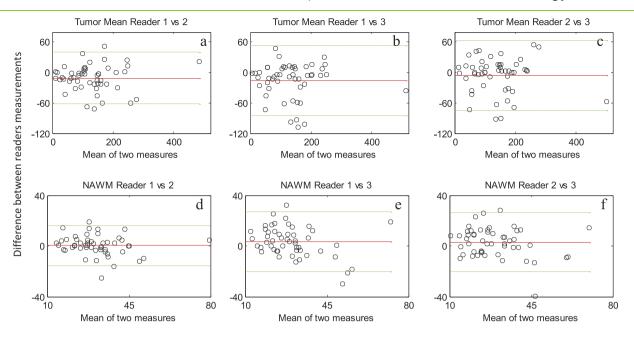
Figure 4. Bland-Altman scatter plots of tumor and NAWM mean CBV. The dotted lines show the 95% confidence interval. *Y*-axis is the difference between the two readers' measurements.

but 1 of the 10 patients analyzed. In the outlying patient with tumor occupying 12% of the brain, the SD changed by only 2.5%-3%. This analysis shows that variation in tumor volume has minimal effect on SD even in the worst case where tumor involves a very large fraction of brain.

Table 2 summarizes the statistical analysis of the GN-nCBV and WM-nCBV estimates. Both $nCBV^{mean}$ and $nCBV^{max}$ measurements obtained from WN differed significantly among the three independent readers ($P < .001$). The difference is largely attributable to reader 3 whose measurements were significantly higher than readers 1 and 2. GN of nCBV measurements did not resolve the significant differences among the three readers for either $nCBV^{mean}$ or $nCBV^{max}$

measurements ($P = .004$ and $P < .001$, respectively) but did substantially decrease the SEs of each pairwise comparison.

Reproducibility metrics obtained for Tumor CBV^{mean}, Tumor CBV^{max}, and NAWM CBV^{mean} and nCBV measurements among the three readers are listed in Table 3. Tumor CBV^{mean}, Tumor CBV^{max}, and nCBV measurements derived from GN all had an excellent ICC value (≥ 0.82) and CCC (≥ 0.79). NAWM CBV^{mean} measurements had good ICC and CCC values both of 0.66. On the other hand, nCBV measurements derived from WN had a fair ICC value between 0.54 and 0.55 and CCC of 0.48. The RC of the absolute measured values on CBV maps was relatively high, with Tumor CBV^{max}RC of 83.43 and NAWM
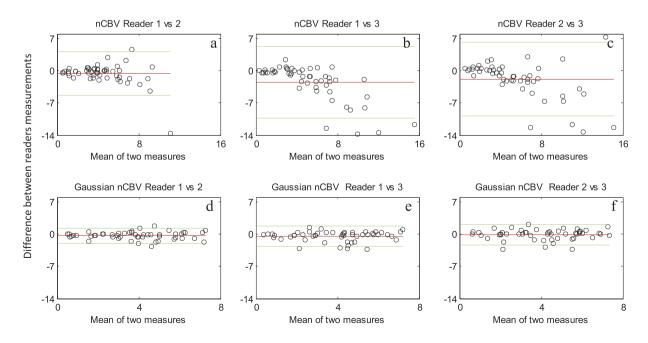

Figure 5. Bland-Altman scatter plots of manual ROI normalized and Gaussian normalized nCBV between readers. The dotted lines show the 95% confidence interval. *Y*-axis is the difference between the two readers' measurements.

Table 4. AUC Correlating nCBV with 15-Month OS

Reader	AUC (SE)			
	$nCBV^{mean}$		$nCBV^{max}$	
	WN	GN	WN	GN
1	0.60 (0.08)	0.67 (0.08)	0.62 (0.08)	0.69 (0.08)
	$P = .25$	$P = .03$	$P = .16$	$P = .01$
2	0.67 (0.08)	0.74 (0.08)	0.67 (0.08)	0.71 (0.07)
	$P = .03$	$P = .002$	$P = .02$	$P = .004$
3	0.70 (0.07)	0.71 (0.08)	0.72 (0.07)	0.75 (0.07)
	$P = .008$	$P = .003$	$P = .003$	$P = .001$
Mean	0.68 (0.08)	0.73 (0.07)	0.70 (0.07)	0.73 (0.07)
	$P = .02$	$P = .002$	$P = .007$	$P = .001$

P values evaluate the differences of the estimated AUC to 0.5.

CBV^{mean} RC of 20.94. But because RC measures the within-subject variance, it is expected to be relatively high for high mean values. The RC of normalized nCBV was smaller using GN compared to WN, demonstrating a better reproducibility of using GN. The wCV, already scaled by the mean, can be used to compare across all measurement types. The Tumor CBV^{mean}, Tumor CBV^{max}, and wCV were less than 20%, but NAWM CBV^{mean} wCV was higher than 25%, suggesting that readers were more consistent in placing hot spot ROI in tumor than in NAWM. Using WN, the wCV of nCBV was >46%. GN reduced the wCV to 18.5% and 21.1% for $nCBV_G^{mean}$ and $nCBV_G^{max}$, respectively.

Figure 4 shows Bland-Altman scatter plots of Tumor CBV^{mean} and NAWM CBV^{mean} measurements. No trend was observed in the difference of Tumor CBV^{mean} (*A-C*) and NAWM CBV^{mean} (*D-F*) between readers. Figure 5 reveals that, for all pairwise comparisons, the nCBV difference between readers increases with increasing WN-nCBV measurements (*A-C*). On the contrary, no systematic trend is observed in GN-nCBV measurements between readers (*D-F*). In addition, the *y*-axis scale is the same across the plots, clearly illustrating that the GN-nCBV has a much smaller variation between readers than the WN-nCBV. Note that the *x*-axis scale representing the mean of two readers nCBV is smaller for the GN-nCBV (0-8) than WN-nCBV (0-16) plot.

Table 4 shows the AUC estimate to evaluate the association between nCBV and 15-month OS in 51 subjects, where 46 deaths were observed and the other 5 were still alive at the cutoff for this analysis. Twenty of the 46 deaths occurred within 15 months of follow-up and the other 26 after 15 months. We compared GN-nCBV and WN-nCBV measurements among the three readers. As reflected in Table 4, all nCBV measurements, except for reader 1 WN-nCBV, were significantly different from 0.5, suggesting that both WN and GN are good predictors of OS. The mean GN-nCBV among the three readers (i.e.,

$$nCBV_m^{mean} = 1/3 \sum_{i=1}^{3} nCBV_i^{mean} \;\; ; \qquad nCBV_m^{max} = 1/3 \sum_{i=1}^{3} nCBV_i^{max}),$$

was slightly better than the mean WN-nCBV maps. This held true for both tumor mean nCBV (AUC of $nCBV_m^{mean}$ 0.68 vs $nCBV_{Gm}^{mean}$ 0.73, $P = .33$) and tumor max nCBV (AUC of $nCBV_m^{max}$ 0.70 vs $nCBV_{Gm}^{max}$ 0.73, $P = .54$) measurements.

Discussion

Translation of novel quantitative, reproducible, validated imaging biomarkers into clinical practice is an urgent need and the mission of the Quantitative Imaging Network supported by the National Cancer Institute [29]. For early detection of tumor response to treatment, the algorithm that generates the smallest measurement noise will perform best [30]. DSC is widely available, but interoperator variation and time needed for ROI measurement have limited clinical translation of published semiquantitative nCBV analysis techniques.

GN, a promising automatable statistical method for producing nCBV maps, has been demonstrated to be superior to other calculation-based methods [22] but has not been validated in comparison to clinical gold standard OS or compared to current standard manual NAWM ROI normalization.

Validation

In our sample of newly diagnosed GBM patients, GN-nCBV had a slightly stronger association with OS than expert reader WN-nCBV (Table 4). WN and GN $nCBV^{max}$ and $nCBV^{mean}$ were comparable, with $nCBV^{max}$ offering slightly higher AUC for prediction of survival. GN-nCBV, the average of all readers' WN-nCBV, and individual reader 2 and reader 3 WN-nCBV were good predictors of OS (Table 4), but reader 1 WN-nCBV did not correlate with OS, and reader 2 and reader 3 $nCBV^{mean}$ differed significantly (Table 2). These results both validate GN-nCBV and illustrate the importance of such automatable methods to reduce interoperator variation in clinical practice and research.

Reproducibility

Interobserver wCVs for tumor $nCBV^{max}$ (46%) and $nCBV^{mean}$ (48%) in our sample were similar to previous literature when using comparable statistical methods [18]. Applying the formula CV = 100*SD/mean to each lesion in our sample individually before averaging across all patients, per Wetzel et al. [18], yields "average" CV of 32% for both $nCBV^{max}$ and $nCBV^{mean}$, nearly identical to their published "average" CV of 30% for $nCBV^{max}$ and 35% for $nCBV^{mean}$. For NAWM CBV, our "average" CV of 23% accords well with the 20% in that report.

As an operator-independent statistical method, GN is completely reproducible within individual patients. The $\sigma_{CBV_{Whole\ Brain}}$ used for GN is very robust and reliable (Figure 3). It has substantially smaller interpatient variation than NAWM CBV regardless of reader or NAWM ROI size (Figure 1), which may be an important advantage for use in clinical trial grouped analyses.

Sources of Variation

Tumor CBV measured by reader 1was significantly different from reader 2 ($P = .004$) and reader 3 ($P = .002$) (Table 1). While neither WN nor GN eliminated interreader variation completely (Table 2), SD is much lower (1.9 vs 2.3-5.0) in the GN-nCBV data, suggesting that the variation in NAWM ROI selection is responsible for a substantial degree of interoperator variation in nCBV. Analysis of ICC between readers supports this, revealing excellent Tumor CBV^{mean} (0.93) and Tumor CBV^{max} (0.91) ICC (Table 3) but relatively poor NAWM CBV^{mean} ICC (0.66.). WN markedly decreases tumor CBV ICC from 0.93 to 0.54 for $nCBV^{mean}$ and 0.91 to 0.55 for $nCBV^{max}$. Normalization inevitably decreases ICC somewhat by combining measurement error from the denominator (NAWM CBV^{mean} or $\sigma_{CBV_{Whole\ Brain}}$) with that of the numerator (Tumor CBV^{mean} or Tumor CBV^{max}), but this marked decrease produced by WN contrasts with the relatively slight ICC decrease (GN-$nCBV^{mean}$ 0.86 and GN-$nCBV^{max}$ 0.82), providing further evidence that NAWM ROI variability is the primary source of nCBV variation (Table 3).

Possible explanations for the observed lower reproducibility of NAWM ROI measurements compared to tumor ROI measurements include larger degree of freedom in selection of NAWM ROI compared to tumor ROI and greater impact of intravoxel noise and partial volume averaged blood vessels and gray matter on lower NAWM CBV estimates compared with higher tumor CBV estimates. We varied the size of NAWM ROI to test the effect of differences in partial volume averaging and voxel SNR but detected no significant difference in NAWM mean or SE between large and small ROI (Table 1).

Weaknesses

The major weakness of our analysis is that the nCBV ROC AUC in our cohort is lower than previously reported AUC of 0.86 and 1.0 for all gliomas and pure astrocytic tumors respectively [1]. In part, this difference likely reflects the larger sample size and more heterogeneous mix of tumor histologies and treatments in our patient group. As such, 0.7 may be a more realistic estimate of the diagnostic value of nCBV in a typical mixed clinical practice. Also, we chose not to use leakage correction methods for this analysis in order to avoid introducing additional computational model complexity and/or sources of variation. This may have contributed to the lower AUCs observed.

Genetic differences between tumors were not considered in this study. Many known genetics tumor markers reflect differences in tumor biology that affect the patient's OS, including IDH-1 mutation and MGMT methylation, among others. In addition, although all subjects received standard of care consisting of maximal resection followed by Stupp protocol temozolamide chemoradiation, many subsequently underwent different experimental treatments on trial. None of these experimental treatments has been proven to improve OS, but it is possible that variations in therapy may have affected survival in some patients. This genetic and treatment heterogeneity in our dataset may have lowered the overall AUC for nCBV association with OS in our dataset, but the observed association with OS nevertheless remains substantial. Since our analysis was designed to validate the newer more reproducible and automatable GN method of CBV analysis and compare it to the existing standard white matter normalization method, rather than to investigate the absolute strength of nCBV association with OS per se, this should not affect our conclusions. In other words, because the same test population with the identical perfusion and survival data was used for both the GN and conventional WN analyses, the heterogeneity should not introduce any bias into our comparison between these normalization methods. From a clinical translation point of view, this heterogeneous population is advantageous. Since it closely simulates a typical clinical population of patients with different tumor genetics undergoing different treatments, our results suggest that GN-nCBV should perform robustly in the clinic.

Conclusion

Both brain tumor nCBV maps produced by GN and by current standard manual NAWM ROI normalization correlate strongly with 15-month OS in our newly diagnosed GBM patients, but the GN-nCBV had a consistently stronger association and far lower interoperator and intersubject variability. This slightly better validity and superior reproducibility, combined with computational simplicity and potential for full automation, argue for implementation of fully automated GN in DSC processing software and for clinical and research use of the GN-nCBV maps produced. Implementation of GN by MRI and postprocessing software vendors is needed to allow widespread use of this technique and can be expected to improve patient care.

References

[1] Barajas RF, Phillips JJ, Parvataneni R, Molinaro A, Essock-Burns E, Bourne G, Parsa AT, Aghi MK, McDermott MW, and Berger MS, et al (2012). Regional variation in histopathologic features of tumor specimens from treatment-naive glioblastoma correlates with anatomic and physiologic MR Imaging. *Neuro Oncol* **14**(7), 942–954.

[2] Law M, Yang S, Babb JS, Knopp EA, Golfinos JG, Zagzag D, and Johnson G (2004). Comparison of cerebral blood volume and vascular permeability from dynamic susceptibility contrast-enhanced perfusion MR imaging with glioma grade. *Am J Neuroradiol* **25**(5), 746–755.

[3] Boxerman JL, Schmainda KM, and Weisskoff RM (2006). Relative cerebral blood volume maps corrected for contrast agent extravasation significantly correlate with glioma tumor grade, whereas uncorrected maps do not. *Am J Neuroradiol* **27**(4), 859–867.

[4] Hakyemez B, Erdogan C, Ercan I, Ergin N, Uysal S, and Atahan S (2005). High-grade and low-grade gliomas: differentiation by using perfusion MR imaging. *Clin Radiol* **60**(4), 493–502.

[5] Bisdas S, Kirkpatrick M, Giglio P, Welsh C, Spampinato MV, and Rumboldt Z (2009). Cerebral blood volume measurements by perfusion-weighted MR imaging in gliomas: ready for prime time in predicting short-term outcome and recurrent disease? *Am J Neuroradiol* **30**(4), 681–688.

[6] Mangla R, Ginat DT, Kamalian S, Milano MT, Korones DN, Walter KA, and Ekholm S (2013). Correlation between progression free survival and dynamic susceptibility contrast MRI perfusion in WHO grade III glioma subtypes. *J Neurooncol* **116**(2), 325–331.

[7] Mangla R, Singh G, Ziegelitz D, Milano MT, Korones DN, Zhong J, and Ekholm SE (2010). Changes in relative cerebral blood volume 1 month after radiation-temozolomide therapy can help predict overall survival in patients with glioblastoma. *Radiology* **256**(2), 575–584.

[8] Schmainda KM, Zhang Z, Prah M, Snyder BS, Gilbert MR, Sorensen AG, Barboriak DP, and Boxerman JL (2015). *Dynamic susceptibility contrast MRI measures of relative cerebral blood volume as a prognostic marker for overall survival in recurrent glioblastoma: results from the ACRIN 6677/RTOG 0625 multicenter trial.* *Neuro Oncol* **17**(8), 1148–1156.

[9] Spampinato MV, Schiarelli C, Cianfoni A, Giglio P, Welsh CT, Bisdas S, and Rumboldt Z (2013). Correlation between cerebral blood volume measurements by perfusion-weighted magnetic resonance imaging and two-year progression-free survival in gliomas. *Neuroradiol J* **26**(4), 385–395.

[10] Kickingereder P, Wiestler B, Burth S, Wick A, Nowosielski M, Heiland S, Schlemmer H-P, Wick W, Bendszus M, and Radbruch A (2015). Relative cerebral blood volume is a potential predictive imaging biomarker of bevacizumab efficacy in recurrent glioblastoma. *Neuro Oncol* **17**(8), 1139–1147.

[11] Sorensen AG, Emblem KE, Polaskova P, Jennings D, Kim H, Ancukiewicz M, Wang M, Wen PY, Ivy P, and Batchelor TT, et al (2012). Increased survival of glioblastoma patients who respond to antiangiogenic therapy with elevated blood perfusion. *Cancer Res* **72**(2), 402–407.

[12] Hu LS, Baxter LC, Smith KA, Feuerstein BG, Karis JP, Eschbacher JM, Coons SW, Nakaji P, Yeh RF, and Debbins J, et al (2009). 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. *Am J Neuroradiol* **30**(3), 552–558.

[13] Ramon J, Barajas F, Chang JS, Segal MR, Parsa AT, McDermott MW, Berger MS, and Cha S (2009). Differentiation of recurrent glioblastoma multiforme from radiation necrosis after external beam radiation therapy with dynamic susceptibility-weighted contrast-enhanced perfusion MR imaging. *Radiology* **253**(2), 486–496.

[14] Cha S, Knopp EA, Johnson G, Wetzel SG, Litt AW, and Zagzag D (2002). Intracranial mass lesions: dynamic contrast-enhanced susceptibility-weighted echo-planar perfusion MR imaging. *Radiology* **223**(1), 11–29.

[15] Boxerman JL, Rosen BR, and Weisskoff RM (1997). Signal-to-noise analysis of cerebral blood volume maps from dynamic NMR imaging studies. *J Magn Reson Imaging* **7**(3), 528–537.

[16] Hu LS, Baxter LC, Pinnaduwage DS, Paine TL, Karis JP, Feuerstein BG, Schmainda KM, Dueck AC, Debbins J, and Smith KA, et al (2010). 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* **31**(1), 40–48.

[17] Perkio J, Aronen HJ, Kangasmaki A, Liu Y, Karonen J, Savolainen S, and Ostergaard L (2002). Evaluation of four postprocessing methods for determination of cerebral blood volume and mean transit time by dynamic susceptibility contrast imaging. *Magn Reson Med* **47**(5), 973–981.

[18] Wetzel SG, Cha S, Johnson G, Lee P, Law M, Kasow DL, Pierce SD, and Xue X (2002). Relative cerebral blood volume measurements in intracranial mass lesions: interobserver and intraobserver reproducibility study. *Radiology* **224**(3), 797–803.

[19] Østergaard L (2005). Principles of cerebral perfusion imaging by bolus tracking. *J Magn Reson Imaging* **22**(6), 710–717.

[20] Jafari-Khouzani K, Emblem KE, Kalpathy-Cramer J, Bjørnerud A, Vangel MG, Gerstner ER, Schmainda KM, Paynabar K, Wu O, and Wen PY, et al (2015). Repeatability of cerebral perfusion using dynamic susceptibility contrast MRI in glioblastoma patients. *Transl Oncol* **8**(3), 137–146.

[21] Henry ME, Kaufman MJ, Lange N, Schmidt ME, Purcell S, Cote J, Perron-Henry DM, Stoddard E, Cohen BM, and Renshaw PF (2001). Test-retest reliability of DSC MRI CBV mapping in healthy volunteers. *Neuroreport* **12**(8), 1567–1569.

[22] Ellingson BM, Zaw T, Cloughesy TF, Naeini KM, Lalezari S, Mong S, Lai A, Nghiemphu PL, and Pope WB (2012). Comparison between intensity normalization techniques for dynamic susceptibility contrast (DSC)-MRI estimates of cerebral blood volume (CBV) in human gliomas. *J Magn Reson Imaging* **35**(6), 1472–1477.

[23] Bedekar D, Jensen T, and Schmainda KM (2010). Standardization of relative cerebral blood volume (rCBV) image maps for ease of both inter- and intrapatient comparisons. *Magn Reson Med* **64**(3), 907–913.

[24] Akgoz A, Rahman R, You H, Qu J, Hamdan A, Seethamraju R, Wen PY, and Young GS (2014). Spin-echo echo-planar perfusion prior to chemoradiation is a strong independent predictor of progression-free and overall survival in newly diagnosed glioblastoma. *J Neurooncol* **119**(1), 111–119.

[25] Barnhart HX and Barboriak DP (2009). Applications of the repeatability of quantitative imaging biomarkers: a review of statistical analysis of repeat data sets. *Transl Oncol* **2**(4), 231–235.

[26] Stupp R, Mason WP, van den Bent MJ, Weller M, Fisher B, Taphoorn MJ, Belanger K, Brandes AA, Marosi C, and Bogdahn U, et al (2005). Radiotherapy plus concomitant and adjuvant temozolomide for glioblastoma. *N Engl J Med* **352**(10), 987–996.

[27] Oh J, Henry RG, Pirzkall A, Lu Y, Li X, Catalaa I, Chang S, Dillon WP, and Nelson SJ (2004). Survival analysis in patients with glioblastoma multiforme: predictive value of choline-to-n-acetylaspartate index, apparent diffusion coefficient, and relative cerebral blood volume. *J Magn Reson Imaging* **19**(5), 546–554.

[28] Carrasco JL, Phillips BR, Puig-Martinez J, King TS, and Chinchilli VM (2013). Estimation of the concordance correlation coefficient for repeated measures using SAS and R. *Comput Methods Programs Biomed* **109**(3), 293–304.

[29] Clarke LP, Nordstrom RJ, Zhang H, Tandon P, Zhang Y, Redmond G, Farahani K, Kelloff G, Henderson L, and Shankar L, et al (2014). The Quantitative Imaging Network: NCI's historical perspective and planned goals. *Transl Oncol* **7**(1), 1–4.

[30] Meyer CR, Armato SG, Fenimore CP, McLennan G, Bidaut LM, Barboriak DP, Gavrielides MA, Jackson EF, McNitt-Gray MF, and Kinahan PE, et al (2009). Quantitative imaging to assess tumor response to therapy: common themes of measurement, truth data, and error sources. *Transl Oncol* **2**(4), 198–210.