Combined structural, microvascular and functional mapping of brain tumors for improved diagnosis and treatment planning

Kyrre Eeg Emblem

Doctoral Thesis

Faculty of Medicine, University of Oslo

2009

Department of Medical Physics

and

The Interventional Centre

Rikshospitalet

Oslo University Hospital

Norway

Series of dissertations submitted to the
Faculty of Medicine, University of Oslo
No. 793

ISBN 978-82-8072-795-4

Cover: Inger Sandved Anfinsen.
Printed in Norway: AiT e-dit AS, Oslo, 2009.

Produced in co-operation with Unipub AS.
The thesis is produced by Unipub AS merely in connection with the
thesis defence. Kindly direct all inquiries regarding the thesis to the copyright
holder or the unit which grants the doctorate.

Unipub AS is owned by
The University Foundation for Student Life (SiO)

Contents:

1. Acknowledgements

The thesis was carried out from 2007 to 2009 at the department of Medical Physics and the Interventional Centre, Rikshospitalet, Oslo University Hospital, Oslo, Norway and financially supported by the Research Council of Norway.

This work is the result of the skills and contributions of many persons. I express my most sincere gratitude to;

My main supervisor Atle Bjørnerud. This thesis could not have been written without you. I am privileged to work with someone of your stature. Your professional skills and your never-ending source of ideas never stop to surprise me. I am thankful for your patience concerning the countless number of questions and requests you have to put up with. Your social intelligence and open-minded approach have made this work fun.

Co-supervisor John K. Hald for your major role in this thesis and the way you have introduced me to the world of neuroradiology. Our discussions, both professional and social, are a great asset to me. Your sense of humor and the way you embrace our many ideas is indeed inspiring.

Other co-authors at the Clinic for Imaging- and Intervention, first and foremost Paulina Due-Tønnessen, Bård Nedregaard, Terje Nome and Bjørn Tennøe, The amount of work you have put into this project has been considerable. Hopefully, I will not ask you to draw so many regions-of-interests in the future.

Other co-authors and collaborating partners, David Scheie, Olivera Casar Borota, Milada Cvancarova, Torstein R. Meling, Inge R. Rasmussen and last but not least Frank G. Zoellner. It has been a pleasure working with so many skilled scientists and I am looking forward to our future projects.

Friends and colleagues at the department of Medical Physics, Clinic for Imaging- and Intervention and the Interventional Centre. In particular, Arne Skretting, Hans-Jørgen Smith, Jarl Jakobsen, Erik Fosse and Marianne Berg for all your valued support and practical help.

My parents, Turid and Knut, for their never-failing support and love. Friends and family for your presence and interest in my work.

My daughter Alma, you make me see the most important aspects of life.

Finally, I would like to express my outmost gratitude and love to my companion in life, Pia, for your patience and support. I could not have completed this work without your enthusiasm and encouragement.

2. List of papers

This thesis is based on the following papers, which will be referred to by their Roman numerals:

I. Emblem KE, Nedregaard B, Nome T, Due-Tonnessen P, Hald JK, Scheie D, Borota OC, Cvancarova M, Bjornerud A. Glioma grading by using histogram analysis of blood volume heterogeneity from MR-derived cerebral blood volume maps. Radiology 2008 Jun;247(3):808-817.

II. Emblem KE, Nedregaard B, Hald JK, Nome T, Due-Tonnessen P, Bjornerud A. Automated Glioma Characterization from Dynamic Susceptibility Contrast Imaging. Brain Tumor Segmentation using Knowledge-based Fuzzy Clustering. (Submitted)

III. Emblem KE, Due-Tonnessen P, Hald JK, Bjornerud A. Automatic Vessel Removal in Gliomas from Dynamic Susceptibility Contrast Imaging. In press, Magn Reson Med 2009

IV. Emblem KE, Zoellner FG, Tennoe B, Nedregaard B, Nome T, Due-Tonnessen P, Hald JK, Scheie D, Bjornerud A. Predictive modeling in glioma grading from MR perfusion images using support vector machines. Magn Reson Med 2008 Oct;60(4):945-952.

3. Abbreviations

γ	Gyromagnetic ratio unique for each isotope possessing a spin
κ	Inter-observer agreement value used in Kappa statistics
υ-SVM	Support vector machines with a υ parameter related to the number of support vectors used and the ratio of the training error
χ	Magnetic susceptibility constant of the tissue (unitless)
ω_0	Angular rotation frequency in units of s^{-1}
2D	2-dimentional
3D	3-dimentional
AIF	Arterial input function
ASL	Arterial spin labeling
AUC	Area under the curve
A_z	Area under ROC curve
B	Magnetic flux density in units of Tesla
C(t)	Concentration-versus-time curve
CAD	Computer aided diagnostics
CBF	Cerebral blood flow measured in ml of blood per 100g of tissue per min
CBV	Cerebral blood volume measured in ml of blood per 100g of brain tissue
CNS	Central nervous system
CSF	Cerebrospinal fluid
DCE	Dynamic contrast enhanced MR imaging
DSC	Dynamic susceptibility contrast MR imaging
EPI	Echo-planar imaging
F	Flow, refer CBF
FCM	Fuzzy c-means clustering
FLAIR	Fluid attenuation inversion recovery
fmAUC	First moment of the area under the curve
FSE	Fast spin-echo
GRE	Gradient-echo
HGG	High-grade gliomas
k	Slope of linear correlation between contrast agent concentration in plasma and change in R_2^* in tissue in units of s^{-1} $mmol^{-1}$
LGG	Low-grade gliomas

LOH	Loss of heterozygosity
M	Magnetization of a medium in units of A/m
MTT	Mean transit time measured in units of s
NMR	Nuclear magnetic resonance
NPV	Negative predictive value
PACS	Picture archiving and communication system
PPV	Positive predictive value
q_0	Quantity of bolus tracer in mmol
R(t)	Residue function
R_1	Longitudinal relaxation rate in units of s^{-1}
R_2	Transverse relaxation rate in units of s^{-1}
R_2^*	Effective transverse relaxation rate in units of s^{-1}
RBF	Radial basis function
RF	Radio frequency
ROC	Receiver operator characteristic
ROI	Regions of interest
S(t)	Signal intensity-versus-time curve
SE	Spin-echo
SI	Signal intensity
SVD	Singular value deconvolution
T_0	Bolus arrival time to tissue of interest measured in s
T_1	Longitudinal relaxation time (spin-lattice) in units of ms
T_2	Transverse relaxation time (spin-spin) in units of ms
T_2*-	Effective transverse relaxation time in units of ms
TE	Echo time in units of ms
TI	Inversion time
TNR	True negative rates
TPR	True positive rates
TR	Repetition time in units of ms
VS	Voxel size
WHO	World health organization

4. Introduction

Epidemiological data from the Cancer Registry of Norway suggests an annual incidence of tumors in the central nervous system (CNS) of 21 per 100,000 persons in Norway. As of 2007, CNS tumors are now the most frequent tumor diagnosis in children and young adults. More than 70% of malignant primary CNS tumors are gliomas, a somewhat heterogeneous group of neoplasms[1]. In spite of important advances in surgery, radiotherapy and chemotherapy, treatment of gliomas constitutes a considerable challenge. Traditionally, the diagnosis and classification of gliomas is based on histopathological characterization of tissue samples from tumor resection or biopsy using the World Health Organization (WHO) classification system[2]. Here, gliomas are graded according to the degree of malignancy (I-IV) of which glioma grades I-II typically are referred to as *low-grade* gliomas while grades III-IV represent *high-grade* gliomas. Although different survival estimates may show large variations, there is a strong correlation between survival rate and glioma grade[1]. Less than 2% of patients diagnosed with a grade IV glioma (Glioblastoma) survive five years after initial diagnosis whereas approximately 60% of patients diagnosed with a grade II glioma (Astrocytoma) survive five years. Because survival rates are strongly correlated to glioma grade, correct tumor grading at an early stage is critical in order to ensure an optimal treatment plan. Furthermore, accurate glioma monitoring is important for the life-long follow-up of tumor recurrence after partial or full tumor resection, and may also become increasingly important for evaluating the response to new chemotherapeutic treatment regimes. Well recognized limitations to invasive grading procedures include sampling error associated with stereotactic biopsy techniques and suboptimal resection due to tumor inaccessibility. Compared to open tumor resection, it was reported in a study that biopsy accuracy of frameless- and frame-based stereotaxis was 89% and 69%, respectively[3]. Also, compared to a small tumor volume ($<50cm^3$), results of a biopsy sample from a large tumor volume ($>50cm^3$) were 8-fold less likely to coincide with the results from open surgery. In addition, repeated invasive procedures to follow tumor growth, tumor recurrence and treatment response are not commonly used.

Based on its high level of soft tissue contrast, magnetic resonance (MR) is the imaging modality of choice to non-invasively characterize brain tumors prior to, during and after treatment. The mainstay of clinical MR brain tumor imaging usually include T2-weighted images with fluid attenuation (FLAIR images) conveying hyper-intense image regions

secondary to tumor infiltration, vasogenic edema and gliosis and T1-weighted images prior to and following administration of a gadolinium-based MR contrast agent. In the latter, high malignancy is commonly (but not always) associated with disruption of the blood-brain barrier which is seen as hyper-intense regions on the contrast enhanced T1-weighted images. Structural tumor appearance and degree of contrast enhancement, however, have a relatively poor correlation with tumor grade and convey little information on tumor patophysiology[4,5]. Furthermore, contrast enhancement as seen on post-contrast T1-weighted images is an unreliable marker for separating tumor recurrence from necrosis induced by radiotherapy, because the two conditions may show similar contrast agent uptake characteristics[6]. As anatomical MR images convey little information about the functional status of the tumor or adjacent tissue, rapid so-called *functional* MR techniques measuring tumor parameters such as tissue perfusion, permeability and water diffusion have been introduced, thereby allowing further insights into tumor pathophysiology. Among these, perfusion MR imaging shows promise through direct assessment of increased tissue vascularity and tumor malignancy. While contrast agent uptake is restricted to brain tumors with disrupted or absent blood-brain-barriers, MR perfusion imaging provides functional information in all vascularized brain tissues. From this, measures of blood flow, blood volume and tissue permeability can be derived, usually by injection of an MR contrast agent. Several studies have shown a stronger correlation between MR perfusion metrics and glioma grade compared to conventional contrast enhanced MR imaging findings[5,7,8]. Also, it has been suggested that MR perfusion may aid in predicting time to progression of glioma grade or survival as an adjunct to histopathology[9,10]. Furthermore, while contrast enhancement on post-contrast T1-weighted images is a relatively unreliable method to guide stereotactic biopsy, it has been suggested that MR perfusion images may aid neurosurgeons in identifying the most malignant tumor area[7,11,12].

4.1 Basics of MR imaging

As a detailed deduction of the principles behind MR imaging is outside the scope of this thesis, only a brief review will be given. The topic, however, is covered in great depth in many textbooks[13].

4.1.1 Image generation

The essence of MR imaging is based on the work by Bloch and Purcell[14,15] in the first half of the 19th century showing that a nucleus with a spin angular momentum (spin) can interact

with a magnetic field. This interaction, known as nuclear magnetic resonance (NMR) is described by the linear relationship:

$$\omega_0 = \gamma B_0 \qquad [1]$$

where $\boldsymbol{\omega_0}$ is the angular rotation frequency (Larmor frequency), $\boldsymbol{\gamma}$ is the gyromagnetic ratio unique for each isotope possessing a spin and $\mathbf{B_0}$ is the static magnetic field. As described by classical physics, each nuclear spin will induce a dipolar magnetic moment and the total sum of the individual magnetic moments of all protons in a given macroscopic sample can be described as the net macroscopic magnetization, **M**. In its general form, as a result of magnetic interaction, the behavior of **M** can be described by the Bloch equation:

$$\frac{dM}{dt} = \gamma (M \times B) \qquad [2]$$

Here, the rate of change of **M** is perpendicular to **B** and **M**. In its equilibrium state, **M** points in the direction of $\mathbf{B_0}$, generally referred to as the z-direction and hence the net magnetization is denoted $\mathbf{M_z}$. Information on $\mathbf{M_z}$ is obtained by introducing a second magnetic field $\mathbf{B_1}$, perpendicular to $\mathbf{B_0}$. In the presence of both $\mathbf{B_0}$ and $\mathbf{B_1}$, the Bloch equation [2] can be written as:

$$\frac{dM}{dt} = \gamma M \times (B_0 + B_1) \qquad [3]$$

Since $\mathbf{M_z}$ is rotating at the Larmor frequency, $\mathbf{B_1}$ is also made to oscillate at this frequency which is in the radiofrequency (RF) range (1-100MHz) and the application of the $\mathbf{B_1}$-field is typically referred to as an RF-pulse. During, and immediately after applying the RF-pulse, a component of **M** will be in the xy-plane. This component is known as $\mathbf{M_{xy}}$ and oscillates about the z-axis. The $\mathbf{M_{xy}}$ component can be detected through the induction of a current in a coil placed in the oscillating field. Following the RF-pulse, because of proton interactions and subsequent loss of energy, the observed signal in the $\mathbf{M_{xy}}$ plane will decay rapidly towards zero and the $\mathbf{M_z}$ will return back to its equilibrium state.

The use of NMR in clinical imaging, first described in the 1970's[16,17], is based on the decomposition of an object-of-interest into small sub-volumes in which $\mathbf{M_{xy}}$ of each sub-volume must be obtained from the measured MR signal. This is done by introducing additional so-called magnetic *gradient fields* in all three orthogonal directions where the strength of the fields varies as a function of position. By applying multiple RF-pulses in combination with different gradient fields, it is possible to select limited volumes within the

object. The delay time between each RF-pulse is called the *repetition time*, TR, and the time between the RF-pulse and the actual image sampling is called the *echo time*, TE. A large variety of methods exist in which these parameters are sampled in different ways to generate various MR images with unique image contrasts. Such an imaging scheme is typically referred to as a *pulse sequence*.

4.1.2 Proton relaxation

The process of magnetization decay in the $\mathbf{M_{xy}}$ plane following a RF-pulse and subsequent magnetization recovery in the $\mathbf{M_z}$ plane is known as *proton relaxation*. In the $\mathbf{M_z}$ direction, the magnetization will recover with a time constant referred to as the T_1 relaxation time with a relaxation rate R_1. The T_1 relaxation time is influenced by fluctuations in the magnetic fields due to thermal motion and chemical exchange, and is therefore also known as *spin-lattice* relaxation. At physiological temperature, the T_1 relaxation times vary in different tissue types from a few hundred milliseconds in fat to several seconds in fluids. Using MR pulse sequences sensitive to variations in T_1 relaxation times, so-called T_1-weighted sequences, the tissue-specific T_1 relaxation times result in a range of image contrasts. A sequence is made T_1-sensitive by using a short TR compared to the longest T_1 relaxation times.

In the $\mathbf{M_{xy}}$ direction, the magnetization will decay with a second time constant referred to as the T_2 relaxation time with a relaxation rate R_2. The T_2 relaxation time is influenced by local field inhomogeneities because of proton interactions and is therefore also known as *spin-spin* relaxation. For all mediums except pure water, the T_2 relaxation time in a given tissue type is considerably shorter than the T_1 relaxation time. The reason for this is that local field inhomogeneities introduce variations in the Larmor frequency with consequent loss of phase coherence between proton spins. As for T_1 relaxation, T_2 relaxation times are longer in fluids than in solid tissue types. A T2-weighted sequence, is made optimally T2-sensitive and minimally T1-sensitive by using a TR of several seconds and a short TE of a few milliseconds.

The relaxation effects described above are commonly referred to as *dipolar* relaxation because they result from direct interactions between spin dipoles. In addition to the dipolar relaxation, proton spins always experience a dephasing effect from bulk inhomogeneities in the $\mathbf{B_0}$ field:

$$\Delta B_0 = B_0(1+\Delta\chi) \qquad [4]$$

where $\Delta\mathbf{B_0}$ is the bulk inhomogeneity in a voxel in the final MR image and χ is a so-called magnetic *susceptibility* constant of the tissue. The χ parameter is a physical constant describing the ability of a medium to become magnetized when exposed to a magnetic field. Almost all biological tissues are *diamagnetic* ($\chi < 0$), whereas air, certain iron containing blood products and gadolinium based contrast agents are paramagnetic ($\chi > 0$). Susceptibility variations therefore results in inhomogeneities between different tissue types and interfaces between tissue and air. The susceptibility effect can have a dramatic impact on the final MR image through acceleration of T2 relaxation times and consequent signal loss. Thus, the actual T_2 decay is referred to as T_2^* (with corresponding relaxation rate R_2^*) and can be approximated by:

$$R_2^* = \frac{1}{T_2^*} = \frac{1}{T_2} + \gamma \Delta B_0 \qquad [5]$$

In MR perfusion imaging, the susceptibility effect may also be used advantageously in an imaging scheme known as dynamic susceptibility contrast (DSC) imaging, which will be discussed in this thesis.

4.2 MR perfusion imaging

Perfusion, or blood flow, is a biological process which assures a sufficient supply of cell nourishment, removal of metabolic waste and upkeep of an adequate body temperature. Lack of perfusion in any vital organ can lead to temperature alteration, loss of tissue viability and ultimately cell death. MR perfusion imaging is a collective term describing methods for deriving tissue blood flow or perfusion related parameters from MR images[13,18]. These methods are based on the principle that the temporal effect of a blood *tracer* (such as a contrast agent) on the MR image signal intensity in a given tissue type can be related to tissue perfusion, blood volume or both. Although alternative methods based on completely non-invasive tracers such as arterial spin labeling (ASL) exist[19], the focus of this thesis is MR perfusion imaging by intravenous administration of an MR compatible contrast agent. The contrast agent alters the biophysical properties of tissue, thereby, increasing the proton relaxation rates which result in a local MR signal change in the tissues where the contrast agent is distributed. By injecting the contrast agent as a rapid, intravenous bolus injection and monitoring the change in MR image intensity with sufficient temporal resolution (typically 1-second intervals), the dynamic image intensity curves as a function of time can

be related to blood flow. In this thesis, only cerebral MR perfusion imaging will be addressed.

4.2.1 Tracer kinetic modeling

The concept of deriving functional information from a blood tracer is referred to as *tracer kinetic modeling* of which the methodology was developed in the 1950's[20]. Perfusion imaging is based on the *central volume principle,* stating that the blood volume of a given tissue (i.e brain tissue) is equal to the blood flow into the tissue multiplied by the mean transit time (MTT) of the blood tracer passing through the capillary structure of the tissue:

$$CBV = CBF \cdot MTT \qquad [6]$$

where CBV is cerebral blood volume and CBF is cerebral blood flow. The CBV parameter is defined as the total volume of blood per volume brain tissue and is measured in milliliters of blood per 100 grams of brain tissue [ml/100g]. The CBF parameter is defined as the net blood flow per volume brain tissue and is measured in milliliters of blood per 100 grams of tissue per minute [mL/100 g/min]. The MTT parameter is measured in seconds [s].

In MR perfusion imaging, it is assumed that there is a well-defined linear relationship between the observed signal change in the MR images due to the presence of the contrast agent and the tissue concentration of the contrast agent. In general, this argument holds true if the contrast agent is distributed in the intravascular space of the CNS only, and if the blood-brain barrier is intact. There are two main techniques for contrast enhanced MR perfusion imaging; T_1- and T_2- or T_2*-weighted perfusion imaging. All gadolinium-based MR contrast agents induce enhancements in both T_1- and T_2/T_2* proton relaxation and the effect of the agent on the MR signal intensity depends on whether the acquisition is made T_1- or T_2/T_2* sensitive. In a T_1-weighted acquisition, the MR signal intensity increases dynamically as a function of local contrast agent concentration, whereas in a T_2/T_2*-weighted acquisition the signal intensity decreases as a function of contrast agent concentration. When applied to dynamic MR perfusion analysis, T_1-weighted imaging is commonly referred to as dynamic contrast enhanced (DCE) imaging whereas T_2/T_2* weighted imaging is referred to as DSC imaging.

The main advantage of DCE imaging over DSC imaging is that T1-weighting can be obtained with sufficient temporal resolution using conventional MR imaging techniques that are inherently insensitive to unwanted susceptibility effects. Image distortion and artificial

signal loss can therefore to a large extent be avoided. The main advantage of DSC imaging compared to DCE imaging is the large MR signal change (high contrast agent sensitivity) combined with a high temporal resolution. The T2*-effect is long reaching and affect a large proportion of the extravascuclar water protons in the brain, resulting in a large dynamic MR signal change. The T1-effect however, is primarily a local effect which requires the water protons to come in direct contact with the paramagnetic centre of the contrast agent in order to be affected (i.e. induce T1-relaxation). The result is that less than 10% of local tissue protons are affected. Because all MR contrast agents are confined to the intravascular space in the brain, T1-enhancement is restricted to the intravascular volume in the absence of contrast agent extravasation. The sensitivity of the DCE methodology in terms of the achievable dynamic MR signal change is, therefore, limited by the low blood volume and the rate of intra- extra-vascular water exchange in normal brain tissue. T1-effects do, however, become much larger in the event of extravascular leakage secondary to tumor pathology, and T1-weighted imaging is therefore the preferred technique for estimation of tumor induced changes in capillary permeability[7,18,21]. Permeability estimation typically requires images to be acquired over a period of several minutes in order for appreciable leakage effects to be detected. In contrast, perfusion metrics from DSC imaging can be measured from just the first passage of the contrast agent through the tissue (of the order of 1 minute acquisition time). In this thesis, with focus on the CBV parameter, DSC imaging was used in all studies (Papers I-IV) and discussed from here on out.

Prior to deriving perfusion metrics from T2*-sensitive DSC images, the signal intensity-versus-time curve S(t) in a given image voxel or region is converted into a concentration-versus-time curve C(t) by assuming a mono-exponential signal decay as a function of increase in relaxation rate ΔR2*:

$$C(t) = -\frac{k}{T_E}\ln\left(\frac{S(t)}{S(0)}\right) \propto \Delta R_2^{*}(t) \qquad [7]$$

where k is an unknown proportionality constant and T_1-effects are assumed negligible[22]. In practice, as shown in Figure 1, a $\Delta R_2^{*}(t)$ curve proportional to C(t) is derived from S(t) because of the assumed linear relationship between C(t) and $\Delta R_2^{*}(t)$ in tissue[23].

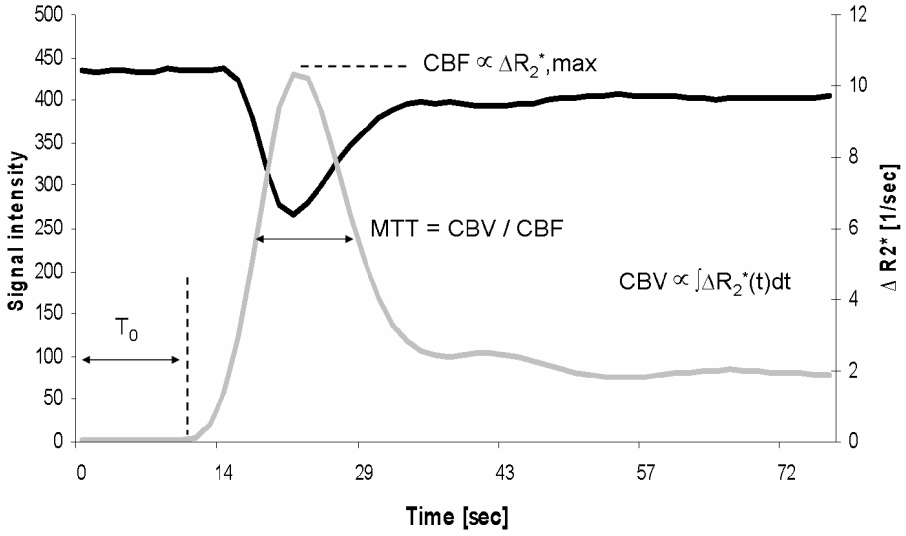


Figure 1: Measured effect of the contrast agent bolus on the MR signal (black line) in a DSC sequence and the corresponding estimated change in the relaxation rate ΔR_2^* (gray line). Relaxation rates are estimated relative to the baseline rate, which is measured from the bolus arrival time T_0. During the first-pass bolus passage, the relaxation rate increases and decreases with a full-width half-maximum equal to MTT. Because of recirculation, a second, smaller bolus peak is observed (at approximately 40-45 seconds) during the post-bolus phase.

It can be shown that CBV is proportional to the integral of the concentration-versus-time curve in a given tissue, $C_{Tissue}(t)$, but this integral needs to be normalized to the corresponding concentration-versus-time curve of a feeding artery, known as the arterial input function (AIF) and denoted $C_{AIF}(t)$, in order to derive (semi-) quantitative CBV estimates[22,24]:

$$CBV = \frac{\int_0^\infty C_{Tissue}(\tau)d\tau}{\int_0^\infty C_{AIF}(\tau)d\tau} \qquad [8]$$

In MR tumor perfusion imaging, CBV values are commonly not normalized to the AIF since AIF determination is time-consuming and lacks the required robustness in a clinical setting. A relative CBV estimate is instead obtained by assuming the AIF to be constant[22,25]. This simplification of equation [8] is based on the principle that the bolus inflow originate

from a common site (i.e. the injection site). In order to deduct the mathematics behind this, a general flow model is introduced as shown in Figure 2.

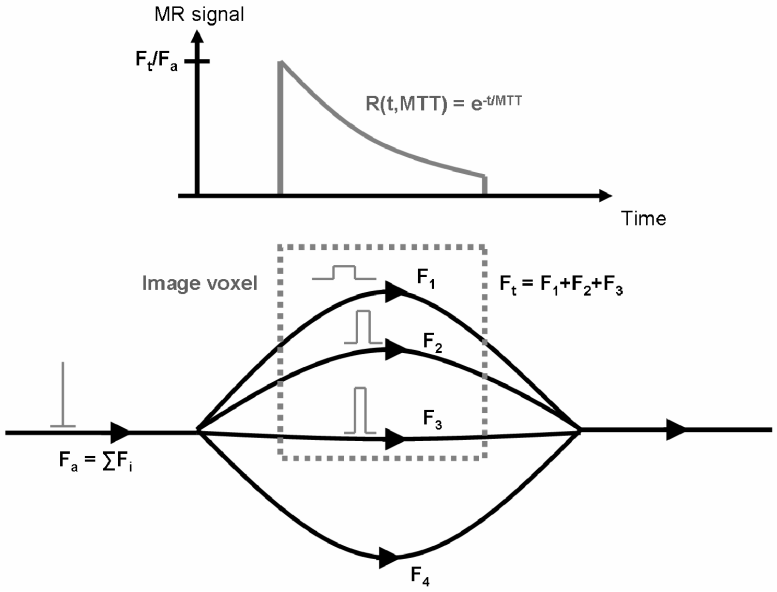


Figure 2: A model of a closed, one-way flow system, from the arterial inflow phase (F_a), through the capillary microvasculature (F_i) and the venous outflow phase. For simplicity, F_a is assumed administrated as a tight bolus with zero duration (i.e. a delta input function) at t=0. In the image voxel volume, the microvascular tubes are assumed to have equal cross-section areas, but various lengths and input functions (in gray). The resulting MR signal is the sum of all F_i in the voxel volume.

In this model, it is assumed that blood flow (or volume) can be assessed from the input tracer and that the total flow in a given region is the sum of all the different vessel branches the tracer use when passing through the system (i.e. image voxel). From figure 2, for a tight bolus in a closed system with known quantity, q_0, the regional C(t) is equal to:

$$C(t) = \frac{kq_0}{F_a} F_t R(t) \qquad [9]$$

where F_t is tissue blood flow in the brain region, F_a is the arterial blood flow and R(t) is a *residue function* describing the total amount of the bolus that remains in the vasculature at time t. The shape of R(t) is generally unknown (and tissue specific) but for simplicity, as shown in Figure 2, R(t) at time t can be expressed as a mono-exponential function of MTT[24]:

$$R(t, MTT) = e^{-\frac{t}{MTT}} \qquad [10]$$

From equations [6], [9] and [10], the area under C(t) in a given brain tissue region can be related to CBV:

$$\int_0^\infty C_{Tissue}(t)dt = \frac{kq_0}{F_a}F_t\int_0^\infty R(t)dt = \frac{kq_0}{F_a}F_t \cdot MTT = \frac{kq_0}{F_a}CBV \qquad [11]$$

Equation [11] is a restatement of equation [8], and estimations of qualitative CBV can be made by assuming kq_0/F_a to be constant, making CBV proportional to the area under C(t)[22,25]. Furthermore, measurement of qualitative CBF values can be made from equation [6] by first estimating an MTT index[25]:

$$MTT_{index} = \frac{fmAUC}{AUC} = \frac{\int_0^\infty C_{Tissue}(t)tdt}{\int_0^\infty C_{Tissue}(t)dt} \qquad [12]$$

where fmAUC is the first moment of the area under the curve. Originally, according to the area/height theorem, MTT_{index} was estimated as the AUC divided by the peak height of the $\Delta R_2^*(t)$ curve[26], and both these approaches are currently used for MTT_{index} estimations. Examples of qualitative CBV, CBF and MTT maps are shown in Figure 3.

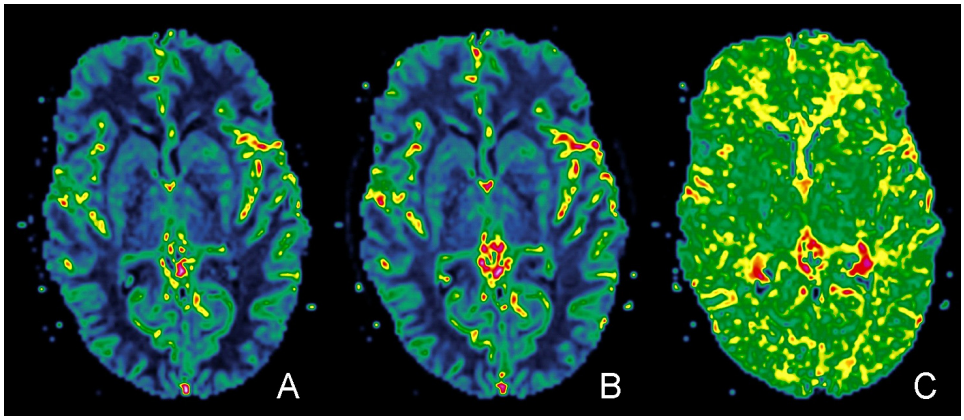


Figure 3: Qualitative CBV **(A)**, CBF **(B)** and MTT maps **(C)** in a healthy brain region of a 29-year old male. The perfusion maps were computed from DSC data during administration of an intravenous contrast agent. According to the central volume principle, CBV is the product of CBF and MTT. As shown in **(A)** and **(B)** respectively, values of CBV and CBF are higher in gray matter (green areas) compared to white matter regions (blue areas). The highest CBV and CBF values are observed in vessels (red/yellow areas). In **(C)**, high MTT values (red/yellow) are observed in tissue areas with long transit times.

Estimation of *quantitative* perfusion metrics from DSC is challenging because the exact relationship between contrast agent concentration and corresponding MR signal change is generally not known. It is, however, possible to estimate perfusion indices which can be compared on an inter-patient basis by assuming the dose-response to be the same in all individuals (and also for all MR scanners used)[22]. To account for the physiological effects of the AIF on the dynamic MR signal in brain tissue (tissue response), a mathematical technique known as *deconvolution* is used. When the tissue response is deconvolved with the AIF one can estimate what the response would be like for an 'ideal' AIF in the form of a spike input with negligible duration (i.e. the delta function shown in Figure 2). This assumption illustrates an ideal situation with no contrast agent dispersion taking place between site of injection and the tissue of interest. The corresponding tissue response to such a spike input is described by the residue function:

$$C_{Tissue}(t) = CBF_{Tissue} C_{AIF}(t) \otimes R(t) = CBF_{Tissue} \int_0^t C_{AIF}(t) R(t-\tau) d\tau \qquad [13]$$

By measuring the AIF in each subject individually, quantitative values of CBV can be obtained from equations [8] or [11], whereas CBF and MTT can be estimated quantitatively from the residue function. As shown in equation [13] and Figure 2, quantitative values of CBF can be derived from the initial height of the residue function, at R(0)=1, obtained by deconvolution of the $C_{Tissue}(t)$ curve from the $C_{AIF}(t)$ curve.

Although a detailed deduction of the mathematics behind equation [13] is beyond the scope of this thesis and the focus of much research and debate[27], the most common deconvolution routine of the tissue response curves (in all image pixels) is based on the mathematical method known as singular value deconvolution (SVD)[24]. It should be emphasized that the final output is semi-quantitative perfusion values only, as absolute quantification requires exact knowledge about tissue density and large/small vessel hematocrit values as well as the exact relaxation properties of the contrast agent used.

4.3 Dynamic susceptibility contrast (DSC) imaging regimes

The susceptibility effects induced by a gadolinium based contrast agent are especially dominant in the brain due to the intravascular distribution of the agent, causing large local variations in the magnetic field in the vicinity of the capillaries and consequent signal change in DSC images. In T_2*-sensitive DSC imaging, so-called echo-planar imaging (EPI)

techniques must be used in order to obtain good T_2*-sensitivity with sufficient temporal resolution. While conventional MR imaging sequences acquire only a small sample of the total MR image per TR, the complete MR image from an EPI-based sequence is formed in just one TR[13]. The EPI technique is inherently sensitive to T2*-effects of the contrast agent, but is also very sensitive to unwanted susceptibility effects present in the brain which cause signal loss and image distortion, especially in brain regions close to air-filled cavities.

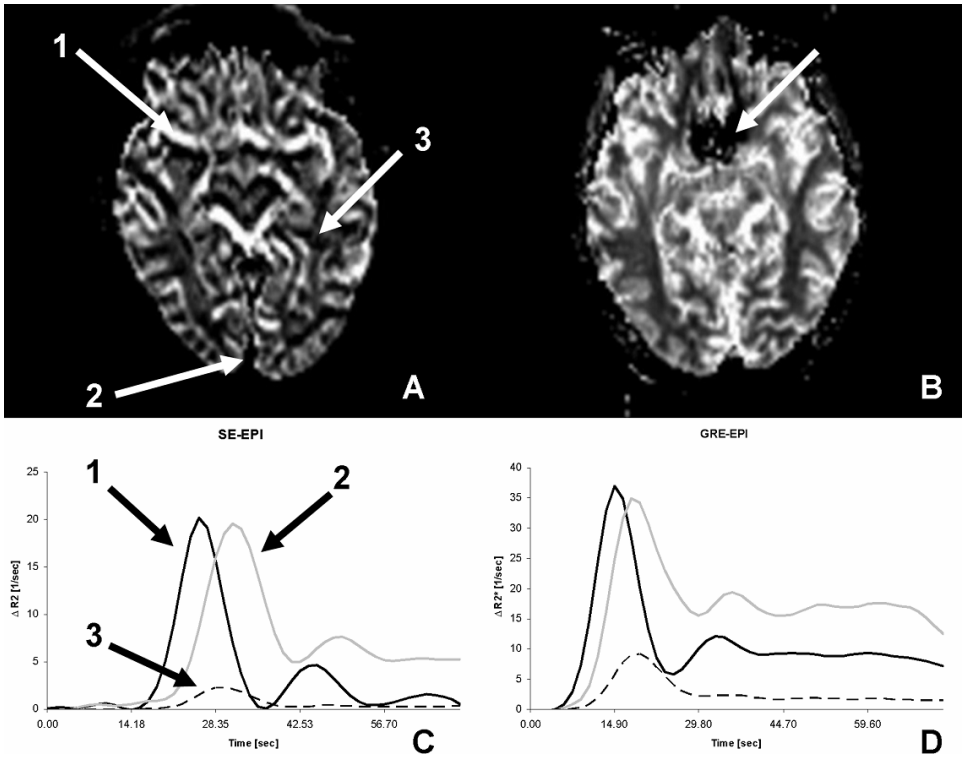


Figure 4: (A) R2-image from SE-EPI and **(B)** R2*-image from GRE-EPI. The corresponding bolus curves from a ROI in **(1)** the middle cerebral artery, **(2)** superior sagittal sinus vein and **(3)** normal appearing white matter tissue are shown as black lines, gray lines and dotted lines in **(C)** and **(D)**, respectively. For both techniques, note the reduced signal change (i.e. reduced CBV) in brain tissue compared to arteries and veins. Compared to **(C)**, the signal change in **(D)** is approximately twice as high because of the susceptibility effect. However, the high sensitivity of the GRE-EPI sequence to susceptibility effects may introduce signal dropouts in brain areas close to air filled cavities as indicated by the white arrow in **(B)**.

Two different types of EPI sequences have been applied in DSC imaging; either based on a spin echo (SE-EPI) readout or a gradient echo (GRE-EPI) readout. Here, SE and GRE refer to two different methods used for encoding the MR signal, and a SE based sequence is mainly T2-sensitve while a GRE based sequence is mainly T2*-sensitive. A potential

advantage of the SE-EPI sequence over GRE-EPI is a more selective sensitivity to relaxation effects occurring at a capillary level (1-2μm), with reduced sensitivity to large vessel relaxation effects (at 3-4μm)[28,29]. The main disadvantage of SE-GRE compared to GRE-EPI, however, is reduced overall sensitivity to the susceptibility induced T2*-relaxation effect, thereby giving a smaller dynamic MR signal response to the contrast bolus as shown in Figure 4. One could, therefore, speculate that the SE-EPI approach better reflects *true* tumor pathophysiology compared to GRE-EPI due to increased sensitivity to capillary relaxation effects. However, it has been reported that GRE-EPI based DSC imaging provides a stronger correlation between glioma grade and tumor CBV compared to SE-EPI sequences[30,31]. Also, GRE-EPI sequences may be a better aid in the differentiation between infiltrating vessels and true tumor CBV elevation[32]. In this thesis, the GRE-EPI sequence was used in all studies (Papers I-IV).

A further issue with DSC imaging is that the kinetic models used to derive CBV values assume the tracer to be confined to the intravascular space. Hence, in areas of extravascular leakage this assumption is no longer valid. The situation is further complicated by the unpredictable relaxation effects which dominate once the contrast agent extravasates. Depending on imaging parameters and probably also local extravascular concentration, the resulting relaxation effect may be predominant T1- or T2/T2*- dependent[33]. Several methods exist for correction of contrast agent extravasation in DSC imaging[7]. One approach is to administer a small contrast bolus prior to the standard bolus injection in order to 'saturate' possible T1-induced relaxation effects caused by extravascular leakage[33]. Alternative methods exist which are based on pure post-processing, either by fitting the dynamic curve to a predefined function which is 'forced' to return to its expected baseline level[34] or to account for the leakage by using a kinetic model similar to the one used in DCE imaging[35].

4.4 DSC-based glioma imaging

The use of MR perfusion imaging in brain tumor patients was first reported in the early 1990's. In one study, areas of strongly enhanced tumor regions during the bolus inflow and a rapid, but incomplete washout effect was observed in five of six patients suggesting a method that might provide useful information when assessing tumor vascularity[36]. Today, MR perfusion sequences are incorporated in the clinical MR protocols at many hospitals

with a comprehensive MR imaging program. Although still relatively new, and the subject of much research and debate, the added value of glioma grading from DSC imaging to conventional MR imaging is promising. Especially qualitative CBV measurements from relative (r)CBV maps (CBV values normalized to a reference CBV value measured in unaffected brain tissue) show a higher correlation to glioma malignancy than conventional MR image metrics[34,37,38] and may also aid in the selection of the optimal target area for biopsy[11]. In addition, as new anti-angiogenic cancer drugs are developed, DSC-based perfusion imaging may aid in assessing anti-angiogenetic treatment effects[39]. Current DSC-based methods for differentiating high-grade gliomas from low-grade are based on measuring the ratio ($rCBV_{max}$) between the most elevated CBV area within the glioma and a CBV value in an unaffected, reference tissue. The regions-of-interests (ROIs), identified by a neuroradiologist, or an operator with good anatomical knowledge, are typically a few voxels wide. This method is often referred to as the *hot-spot method* and the $rCBV_{max}$ parameter shows a strong correlation to glioma grade, as high-grade gliomas tend to have 3-4 times higher mean $rCBV_{max}$ values than low-grade gliomas[34,37,40]. In a study including 160 patients, the hot-spot method resulted in sensitivity and specificity values of 95.0% and 57.5%, respectively, when identifying high-grade gliomas from low-grade[41]. Recently, it has also been reported that the hot-spot method may predict median time to tumor progression in glioma patients independent of histopathologic findings[10]. In that study, tumor progression was defined as a decline in neurologic status or an increase in tumor size of more than 25% as seen on MR images. The results showed that patients with a $rCBV_{max}$ value of < 1.75 had a median time to tumor progression of 3585 days whereas patients with a $rCBV_{max}$ value of > 1.75 had a median time to tumor progression of 265 days. Despite their widespread use in stroke imaging, there has been less focus on assessment of CBF and MTT as predictors for glioma grade. It should be noted, however, that studies have shown that CBF correlates well with both CBV and glioma grade[40,42] whereas MTT does not appear to correlate with glioma grade[43].

Although quantitative assessment of permeability parameters in glioma patients are reported in DCE imaging[44], similar quantitative perfusion assessments in DSC-based glioma grading is rare. This is probably because high diagnostic accuracy[45] values are obtained even using qualitative methods, and also because qualitative analysis is much easier to implement in a clinical setting. However, it has been reported that quantitative values from DSC imaging may also correspond to glioma grade[43].

4.4.1 Current challenges in DSC-based glioma grading

Although current methods for DSC-based glioma grading show promise, some challenges remain.

First, current hot-spot methods suffer from inherent user-dependency. The main reason for this is related to the manual selection of tumor ROIs used in the glioma grading analysis. Even though it has been shown that the metabolic active tumor area is mainly restricted to the tumor border as seen on T2-weighted images, it is well known that malignant tumor infiltration extends beyond the MR visible tumor margins[46,47]. Also, glioma area identification on MR images is inherently difficult because glioma tissue might mimic both unaffected brain tissue and cerebrospinal fluid (CSF) depending on the MR image parameters used. Furthermore, for the hot-spot method, the few voxels that constitute the glioma ROI are inherently prone to image noise and other sources of spurious pixel values (e.g. spikes introduced by the algorithms used to generate the rCBV maps). Consequently, partial or whole tumor area identification is difficult, even for an experienced neuroradiologist making current manual glioma grading methods inherently operator-dependent and time consuming.

Second, separating glioma regions with high rCBV values from similar rCBV values in macroscopic vessels is difficult. Including high rCBV values from vessels in the glioma ROI may result in a false-positive misclassification of a low-grade glioma as a high-grade glioma[48]. For DSC-based glioma imaging, because of the large T2* shortening effect outside the vessel lumen[49,50], anatomical MR images do not reflect the outmost macroscopic vessel margins. To correct for this, vessels can be excluded manually by using coregistered rCBV maps as overlays on the anatomical MR images. This method, however, is very user-dependent.

Third, oligodendroglial tumors (oligodendrogliomas and oligoastrocytomas) tend to show high $rCBV_{max}$ values irrespective of glioma grade[11,51]. A reason for this might be that most oligodendroglial tumors are located in cortical areas and have direct involvement with grey matter with a higher capillary density compared to white matter[51]. Also, the higher oligodendroglial tumor vascularity may be associated with loss of heterozygosity (LOH) on the short arm of chromosome 1 (1p) and the long arm of chromosome 19 (19q), seen in 40-90% of oligodendroglial tumors[52,53]. Consequently, cut-off $rCBV_{max}$ values between glioma

grades may be harder to establish if oligodendroglial tumors are included. Because oligodendroglial tumors constitute approximately 10% of all gliomas[1], it would represent a major limitation to any DSC-based grading method if these tumors had to be characterized by other diagnostic means.

Fourth, reported $rCBV_{max}$ threshold values for optimal differentiation between glioma grades show large variations with values ranging from 1.75 to 5.58[40,41,54]. Although these studies showed high diagnostic accuracy values, this discrepancy suggests that the optimal $rCBV_{max}$ threshold depends on several method specific parameters including contrast agent properties and dose, imaging technique and post-processing routines. This method dependency on the $rCBV_{max}$ threshold may suggest that the applied threshold value must be determined specifically at each site, which complicates comparison of data between sites and also places restrictions on modification of any of the model sensitive parameters in a given institution. Prospective, predictive grading of gliomas from DSC imaging has received little attention in the literature. One possible reason for this is that current $rCBV_{max}$ threshold values are difficult to generalize into a useful model. A predictive model based on one single value per subject will inherently lack sufficient robustness.

4.5 K-means clustering

Classification of data into sub-classes is desirable in medical imaging for separation of image regions or image metrics with similar properties, such as CBV values above or below a threshold value. In statistical analysis, cluster analysis is a common term for discrimination of data based on iterative algorithms that computes the optimal separation between a given set of classes. One simple, *partitional* (non-overlapping) cluster algorithm is the k-means algorithm[55,56]. Here, n objects are divided into a user-specified number of cluster classes, k, of which the objects in a specific class share a common set of attributes. For image segmentation, the algorithm initially selects a set of random cluster centroids positions, and the objective of the clustering algorithm is then, through an iterative process, to minimize the within-class deviation from the class centroid and at the same time maximize the between-class centroid distance.

K-means clustering is commonly implemented as a two-step iterative process. First, based on the complete image or a smaller sub-sample, all points (i.e. image intensity values) are

reassigned at once to their nearest cluster centroid, followed by a recalculation of the cluster centroids. This iteration results in an approximate solution, which reduce the computation time of the second step. Second, each data point is individually reassigned in order to assess whether the new distance reduces the sum of all distances. Although several methods exist, the distance *d* between the *m*-dimensional data point $x_1, x_{2,...}, x_m$ and a centroid position $y_1, y_{2,...}, y_m$ is typically assessed by deriving the squared Euclidean distance:

$$d = \sum_{i=1}^{m} (x_i - y_i)^2 \qquad [14]$$

Finally, the cluster centroids are recomputed until a global minimum is reached which is the optimal separation of the *k* cluster classes. Although a *true* global minimum can only be achieved by performing an exhaustive search over all starting points, a procedure with several varying staring points will generally converge towards a global minimum.

The advantage of using k-means clustering in medical imaging is that the algorithm is fast, even for a large datasets, and produce tight clusters equivalent to homogenous, compact brain tissue areas such as white or gray matter. Because a data point can belong to one cluster class only, a disadvantage of the k-means cluster algorithm is that it is relatively intolerant of imprecise data from heterogeneous image regions. Also, the number of cluster classes is an input parameter which must be selected carefully in order to avoid poor discrimination.

4.6 Summary of introduction

The added value of MR imaging in the diagnosis of glioma patients is unquestioned, providing excellent anatomical and functional tumor information. Since the early 1990's, considerable progress has been made within the field of MR perfusion imaging with respect to MR sequence development and post-processing routines. With focus on cost-benefit and cost-effectiveness in the medical workflow, fast user-friendly diagnostic methods are becoming increasingly important with potential for improved diagnostic efficacy. Current DSC-based glioma grading methods however, suffer from inherent user-dependence and inter-institutional variations suggesting that these methods may not provide the required robustness and reproducibility for a non-invasive alternative to biopsy. Hence, introduction of objective, user-independent methods for glioma characterization is warranted and the focus of this thesis.

5. Aims of the study

The aim of this study was to introduce and evaluate alternative methods for DSC-based glioma characterization with focus on minimal user-dependency. The diagnostic efficacy[45] of the proposed methods was assessed in patients diagnosed with a glioma after an MR exam and subsequent surgery and compared to current methods in the literature.

5.1 Specific aims

- Paper I: To retrospectively compare the diagnostic accuracy of an alternative method to grade gliomas, based on histogram analysis of rCBV values from the entire tumor volume (*histogram method*), with the hot-spot method, using histology as the reference standard.

- Paper II: To compare the diagnostic efficacy of DSC MR imaging, in terms of pre-surgical diagnostic accuracy and expected patient survival, using automatically segmented and manually selected glioma volumes.

- Paper III: To assess whether macroscopic vessels could reliably be removed automatically from DSC-based rCBV maps and whether the proposed vessel removal procedure improved the diagnostic efficacy of glioma grading using histology as reference.

- Paper IV: To implement a predictive model for glioma grading using tumor blood volume histogram signatures derived from DSC images and to assess whether the diagnostic accuracy of the predictive model improved with the size of the reference database.

6. Summary of papers

Paper I. Glioma grading by using histogram analysis of blood volume heterogeneity from MR-derived cerebral blood volume maps. Radiology 2008 Jun;247(3):808-817.

Purpose: To introduce an alternative method to DSC-based glioma grading based on histogram analysis of rCBV values from the entire tumor volume with the overall aim of reducing user-bias.

Materials and Methods: Summaries of patient data and MR imaging parameters are shown in Tables 1 and 2, respectively. Relative (r)CBV maps were created using established tracer kinetic models[22,24] and coregistered to the anatomical MR images. Four neuroradiologists performed the image analysis independently as follows; ROIs of the complete tumor were drawn in each slice according to the combined rCBV map overlay / anatomical MR image underlay information taking care to avoid areas of necrosis, edema or nontumor macrovessels evident on the post-contrast T1-weighted images[30,35]. Signal hyper-intensities thought to represent tumor tissue as seen on the T2-weighted images were used to define the outmost tumor margin, and areas of contrast enhancement seen on the post-contrast T1-weighted images were always included. Examples of glioma ROI identifications in a low-grade, grade II diffuse astrocytoma and a high-grade, grade IV glioblastoma are shown in Figure 5. Each observer recorded the time used to perform the analysis, and evaluated how difficult the two methods were to perform (easy, intermediate or difficult).

Based on the glioma ROIs defined by the four observers, the distribution of whole tumor rCBV values was normalized to a mean white matter rCBV value and analyzed by classifying the complete distribution of $rCBV_{max}$ values into an area-normalized histogram with a predefined number of histogram bins. The reason for the area-normalization is to account for inter-patient variations in tumor size, ensuring that the area under the normalized histogram is always equal to one. Another advantage of normalizing the rCBV histogram distribution is that the heterogeneity is then inversely proportional to the peak height of the normalized distribution. Glioma grade was determined by assessing the peak height of the normalized histogram distribution, under the hypothesis that $rCBV_{max}$ heterogeneity is related to tumor malignancy, and inversely proportional to the peak height

of the histogram distribution. Using logistic regression, McNemar's test, receiver operator characteristic (ROC) curves and inter-observer agreement (Fleiss-Kappa analysis[57]), the proposed method was compared to a reference "hot-spot" method[58].

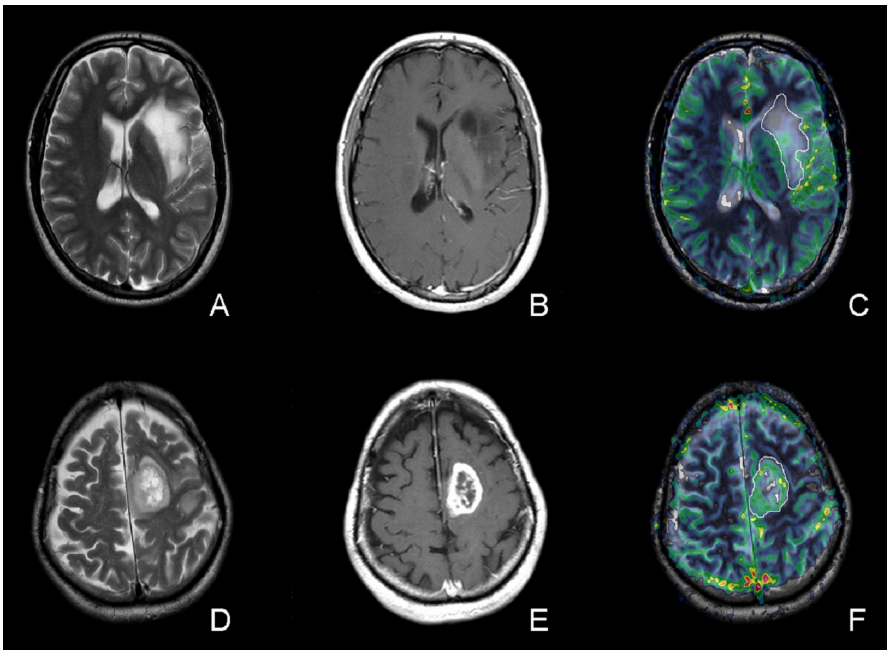


Figure 5: (A) Axial T2-weighted FSE image, **(B)** axial post-contrast T1-weighted SE image and **(C)** coregistered $rCBV_{max}$ map overlaid on **(A)**, respectively, of a patient with a low-grade, grade II diffuse astrocytoma. **(D)** Axial T2-weighted FSE image, **(E)** axial post-contrast T1-weighted SE image and **(F)** coregistered $rCBV_{max}$ map overlaid on **(D)**, respectively, of a patient with a high-grade, grade IV glioblastoma. Tumor areas as identified by a neuroradiologist are shown as white ROIs in **(C)** and **(F)**. Note the higher $rCBV_{max}$ values in the high-grade glioma area in **(F)** compared to the low-grade glioma **(C)**.

Results:

Examples of histogram distributions for glioma grades I-IV are shown in Figure 6. On average, the four observers reported using 7 minutes and 11 minutes per patient when using the reference hot-spot method and the proposed histogram method, respectively. The two methods were reported equally difficult to perform by all observers (intermediate).

For the hot-spot method, optimal $rCBV_{max}$ cut-off values between high-grade and low-grade gliomas ranged from 3.75 to 5.58. For the histogram method, optimal histogram peak height cut-off values between high-grade and low-grade gliomas ranged from .10 to .12. For the histogram method, the diagnostic accuracy was equal for all observers. There was an

increase in sensitivity (90%) and negative predictive value (NPV) (87%) when using the histogram method compared to the sensitivity (55-76%) and NPV (54-74%) of the hot-spot method. For one observer, the sensitivity values of the histogram method was significantly higher than those of the hot-spot method (McNemar's test; P=.002). The specificity and positive predictive value (PPV) was 83% and 87% when using the histogram method and 63-88% and 64-88% when using the hot-spot method, respectively. The mean areas (A_z values ±standard errors) under the ROC curves were larger for all observers when using the histogram method (A_z=.905±.041-.914±.039) compared to the hot-spot method (A_z=.698±.072-867±.055). For one observer, the A_z value when using the histogram method was significantly higher than the A_z value when using the hot-spot method (P<.001).

For the Fleiss-Kappa (κ) analysis, there was a moderate inter-observer agreement between the four observers when using the hot-spot method (κ=.559) and an almost perfect inter-observer agreement (κ=.923) when using the histogram method.

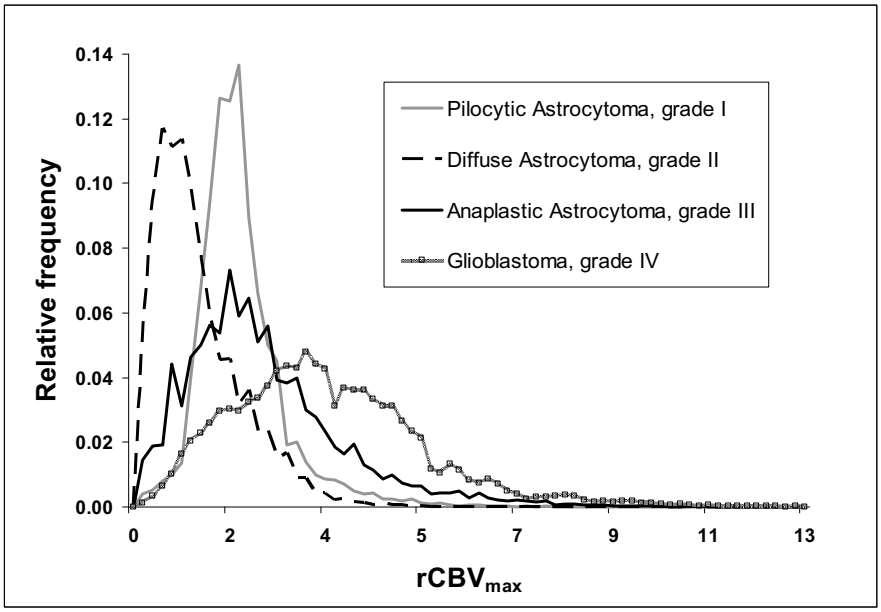


Figure 6: Normalized histogram distributions of $rCBV_{max}$ values from total glioma areas in four patients diagnosed with glioma grades I-IV. Note the lower histogram peak values and wider distributions of the high-grade gliomas (III-IV) compared to the low-grade gliomas (I-II), attributed to increased vascular heterogeneity.

Paper II. Automated Glioma Characterization from Dynamic Susceptibility Contrast Imaging. Brain Tumor Segmentation using Knowledge-based Fuzzy Clustering. (Submitted)

Purpose: To assess whether glioma volumes from knowledge-based fuzzy c-means (FCM) clustering of multiple MR image classes can provide similar diagnostic efficacy values as manually defined tumor volumes when characterizing gliomas from DSC imaging.

Materials and Methods: Summaries of patient data and MR imaging parameters are shown in Tables 1 and 2, respectively. Relative (r)CBV maps were created using established tracer kinetic models[22,24] and coregistered to the anatomical MR images. Two methods for identifying the glioma areas were performed; (1) four neuroradiologists manually defined the glioma areas as described in Paper I, (2) glioma areas were automatically segmented from all anatomical image types using a previously published knowledge-based FCM clustering routine shown to correspond well with manually defined tumor volumes[59-61].

A fuzzy clustering approach differ from a k-means clustering approach in that each data element (single pixel) can belong to more than one cluster thereby providing a more flexible approach than the k-mean clustering. The strength of the association between a data element and a class is indicated using a value between zero and one, in which a value closer to 1 indicate a sharper partitioning. One common fuzzy clustering approach is the FCM algorithm[59]. This approach is based on the selection of an initial guess for the n cluster centroids, representing the mean location for each cluster. From this starting point, the cluster centroids are iteratively updated to its optimal location by minimizing an object function:

$$J_m = \sum_{i=1}^{N}\sum_{j=1}^{C} u_{ij}^{m} \| x_i - c_j \|^2, \qquad 1 \le m \le \infty \qquad [15]$$

where u_{ij} is the strength of the association between a given data point x_i and a cluster class j, and c_j is the cluster centroid. The object function in equation [15] represents the distance from x_i to c_j weighted by u_{ij}. The complementary knowledge-based operation consists of a linear sequence of low-level image processing operations based on known MR image properties secondary to brain structures or pathology[59,60]. Here, for all anatomical MR image types, the glioma class was identified as the FCM cluster class with the highest

image intensity (Figure 7). Finally, a set of binary morphological operations[60] was performed to *clean* the cluster image and a 3D seed growing algorithm was applied on the complete image stack to identify and connect tumor regions in neighboring image slices.

To compare how well (on a pixel-by-pixel basis) the automatically segmented glioma volumes corresponded with the manual glioma volumes, sensitivity values and PPV were derived for high- and low-grade gliomas separately. Using the histogram method (Paper I), sensitivity and specificity values when using manually defined and automatically segmented glioma volumes to pre-surgically grade gliomas were assessed by pair-wise comparisons of the areas (A_z) under the ROC curves. Also, regardless of glioma grade, Kaplan-Meier survival curves with log-rank tests (Mantel-Cox) were used to compare the manual and automatically segmented glioma volumes with respect to separating a "high-risk" patient group from a "low-risk" group. Here, the high-risk group was defined as those patients with an expected survival from MR examination date to death of less than 2 years, whereas the low-risk group was defined as those patients with an expected survival of more than 2 years.

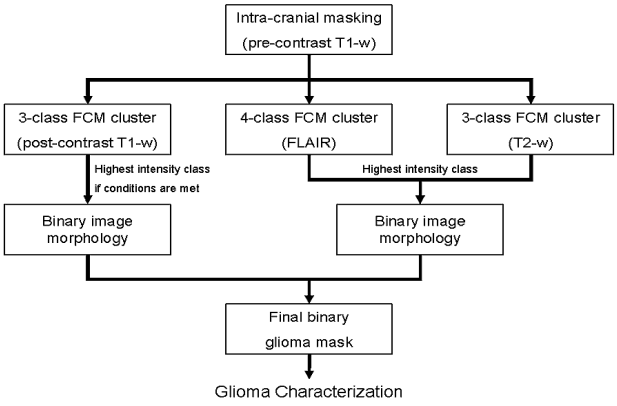


Figure 7: Schematic flow-diagram of the entire automatic segmentation procedure performed in our study. Prior to glioma segmentation, the images were standardized using adaptive histogram equalization[62] and brain tissue pixels were identified using an intra-cranial brain mask procedure in Statistical Parametric Mapping (SPM5)[63]. After FCM cluster analysis, a set of 2D and 3D binary morphological image operations was performed on the segmented glioma images to remove non-tumor pixel areas mimicking tumor tissue[60].

Results:

Examples of the knowledge-based FCM cluster analysis are shown in Figure 8. Across the four observers, the pixel-by-pixel sensitivity and PPV when identifying glioma pixels

manually was 59% (±2%) and 89% (±1%) for the low-grade gliomas and 57% (±2%) and 87% (±1%) for the high-grade gliomas, respectively. For the automatically segmented glioma volumes, the pixel-by-pixel sensitivity values for the low-grade gliomas (83%±2%) and high-grade gliomas (69%±4%) were significantly higher than the manual method (Mann-Whitney; P<.001 and P=.005, respectively). The corresponding pixel-by-pixel PPVs for the low-grade gliomas (66%±3%) and high-grade gliomas (73%±4%) were significantly lower than the manual method (P<.001 and P=.004, respectively). For all observers, there was no significant difference (P=.576-.970) between the areas under the ROC curves when using manually defined glioma volumes (A_z=.875±.049-.908±.040) and automatically segmented glioma volumes (A_z=.890±.046) to grade gliomas. For the survival analysis, a higher log-rank value between the low-risk group and high-risk group (Chi-Square = 14.984, P<.001) was observed when using the automatically segmented glioma volumes compared to the manually defined glioma volumes (9.441-12.022, P=.001-.002).

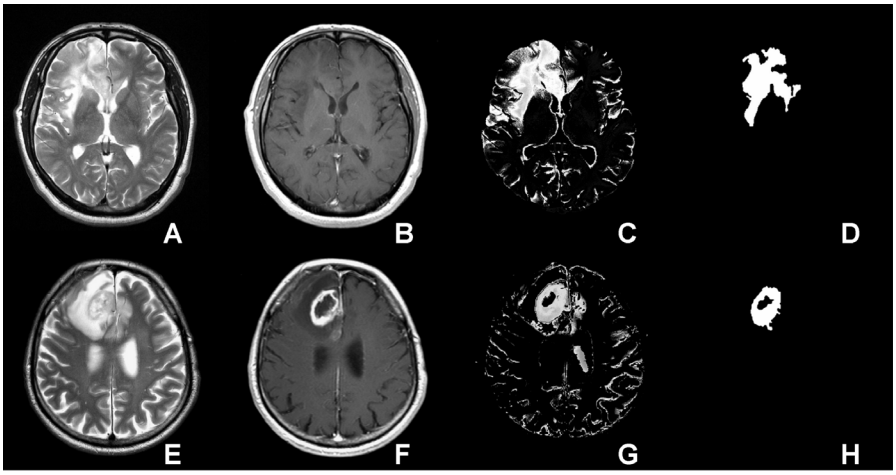


Figure 8: Results of the FCM clustering procedure in a low-grade, grade II oligodendroglioma patient **(A-D)** and high-grade, grade IV glioblastoma patient **(E-H)**. Using anatomical MR images such as T2-weighted **(A, E)** and post-contrast T1-weighted **(B, F)** images, the resulting cluster class representative of the highest mean pixel values was thought to represent the tumor class **(C, G)**. The final glioma volumes **(D, H)**, used in the histogram analysis, were assessed by performing a set of knowledge-based operations, thereby removing non-tumor pixels with similar properties as tumor tissue.

Paper III. Automatic Vessel Removal in Gliomas from Dynamic Susceptibility Contrast Imaging. In press, Magn Reson Med 2009

Purpose: To propose an automatic vessel segmentation technique based on clustering of multiple parameters derived from the DSC first-pass curve with the aim of reducing user-dependency in DSC-based glioma grading.

Materials and Methods: Summaries of patient data and MR imaging parameters are shown in Tables 1 and 2, respectively. Relative (r)CBV maps were created using established tracer kinetic models[22,24] and coregistered to the anatomical MR images. Glioma ROIs were identified manually by a neuroradiologist as follows; (1) as described in Paper I and (2) as described in Paper I but with vessel included in the glioma ROIs.

Using an iterative 5-class k-means cluster analysis approach[64], pixels in the DSC images thought to represent arteries and veins were separated from tissue based on the temporal characteristics of the first-pass curve, including the parameters T_0, AUC and wash-out characteristics of the contrast agent which can be assessed by deriving fmAUC as shown in equation [12]. From this, it was expected that both arteries and veins would exhibit a high AUC (high rCBV). Additionally, arteries (veins) would exhibit a short (long) T_0 and a low (high) fmAUC[65,66]. Combining these features, it was hypothesized that arteries and veins could be distinguished from highly perfused tumor tissue. The remaining three cluster classes were thought to represent tumor and/or gray- and white matter and cerebrospinal fluid.

Since malignant tumor tissue can mimic arterial signal with similarly short T_0 and large rCBV, two additional pre-cluster steps were also tested; (1) a MTT corrected vessel masking procedure and (2) a signal intensity (SI) corrected vessel masking procedure. In both cases, an initial iterative 5-class k-means cluster analysis routine was performed to identify and remove the cluster class with the (1) highest mean MTT value and (2) highest mean SI values as seen on T2*-w baseline DSC images during T_0. For the uncorrected- , the MTT corrected- and the SI corrected vessel masks, pixels thought to represent vessels by the automatic method were removed from the glioma ROIs with vessels included. Examples of binary vessel maps from the three vessel segmentation procedures are shown in Figure 9.

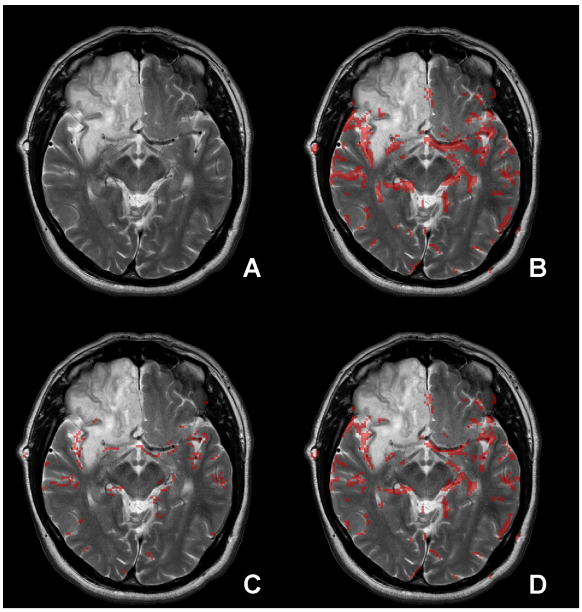


Figure 9: (A) Axial T2-weighted images of a patient with a low-grade, grade II oligodendroglioma. Binary vessel masks overlaid on **(A)** using; **(B)** uncorrected vessel masks, **(C)** MTT corrected vessel masks and **(D)** SI corrected masks. Note the identification of the main arteries and veins in **(B-D)**, also recognized on **(A)**. Compared to the uncorrected **(B)** and SI corrected **(D)** vessel masks, the MTT corrected vessel masks **(C)** are generally more conservative in masking out pixels representing vessels in normal appearing tissue regions.

To compare the automatic vessel segmentation procedure with manual vessel removal, the neuroradiologist recorded the time used to perform the analysis and how difficult it was to perform (easy, intermediate or difficult). Glioma grades were assessed for both methods using the histogram analysis method described in Paper I. Logistic regression was used to derive sensitivity values, specificity values, NPV and PPV when separating high-grade gliomas from low-grade. In addition, areas (A_z) under the ROC curves and Kaplan-Meier survival analysis over a 900 day period was performed as described in Paper II.

Results:

The neuroradiologist used approximately 10 minutes per patient when drawing the glioma ROIs with vessels excluded and 5 minutes with vessels included. When excluding vessels, the degree of difficulty was intermediate. When including vessels, the degree of difficulty was easy to intermediate. Compared to the manual method with vessels excluded by the neuroradiologist, applying the MTT corrected vessel mask resulted in a change of estimated glioma grade in 9 patients. Here, for the 32 patients diagnosed with a histopathological low-

grade glioma, three low-grade gliomas were correctly classified as low-grade whereas and one high-grade glioma was misclassified as low-grade. For the 45 patients diagnosed with a histopathological high-grade glioma, three high-grade gliomas were correctly classified as high-grade whereas two low-grade gliomas were misclassified as high-grade. The three low-grade gliomas with a correct change of glioma grade were all located within the M1 or M2 segments of the middle cerebral artery distribution, whereas none of the misclassified high-grade gliomas were located in close proximity to any of the larger intra-cranial vessels.

For all measures of diagnostic accuracy, the highest value was obtained using the MTT corrected vessel mask. For the manual reference method, the sensitivity, specificity, NPV and PPV was 87%, 78%, 81% and 85%, respectively. For the MTT corrected vessel masking, the corresponding values were 91%, 81%, 87% and 87%, respectively. For the ROC analysis, the A_z value for the reference method was .881 (±.038) and .935 (±.026) for the MTT corrected vessel masking. Similarly, a higher log-rank value was observed between the low-risk and high-risk group when using the MTT corrected vessel masking (Chi-Square = 20.390, P<.001) compared to the manual reference method (11.484, P=.001). Figure 10 show results of the MTT corrected vessels segmentation routine on the CBV map.

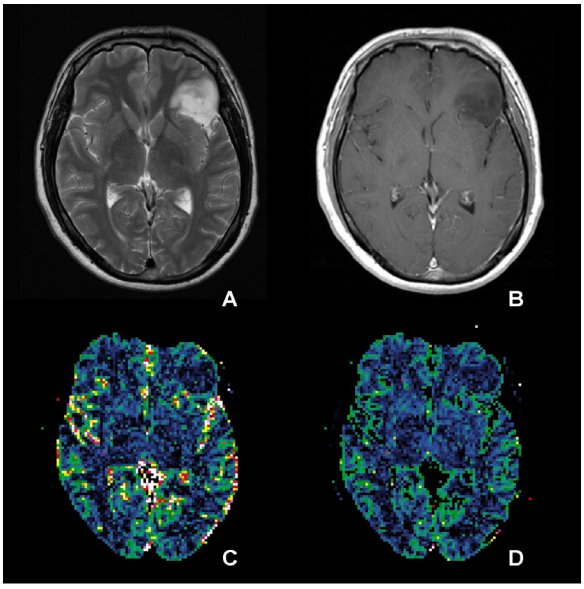


Figure 10: Axial T2-weighted **(A)** and T1-weighted post-contrast **(B)** MR images of a patient with a low-grade, grade II oligoastrocytoma. Coregistered rCBV maps with vessel included and vessel excluded by the MTT corrected vessel masking procedure are shown in **(C)** and **(D)**, respectively. Note the removal of rCBV values from the middle cerebral artery in **(D)** compared to **(C)**, in immediate proximity to the tumor area.

Paper IV. Predictive modeling in glioma grading from MR perfusion images using support vector machines. Magn Reson Med 2008 Oct;60(4):945-952.

Purpose: To implement a predictive model based on support vector machines (SVM) for glioma grading using tumor blood volume histogram signatures derived from DSC images and to assess the diagnostic accuracy of the model and the sensitivity to sample size.

Materials and Methods: Summaries of patient data and MR imaging parameters are shown in Tables 1 and 2, respectively. Relative (r)CBV maps were created using established tracer kinetic models[22,24] and coregistered to the anatomical MR images. Glioma ROIs were identified manually as described in Paper I. Four neuroradiologists drew tumor ROIs in 53 of the 86 patients, whereas a fifth neuroradiologist drew tumor ROIs in the remaining 33 patients. The four sets of 53 histogram signatures from the four observers were used to create a training dataset whereas the remaining 33 histogram signatures from the fifth observer were used to create a test dataset.

To predict glioma grades in the test dataset based on the information in the training dataset, a SVM model was applied. A predictive SVM model is based on generalized linear classification algorithms, which can be used to create a model for classification of new objects based on a set of training data[67,68]. In this, the SVM model creates a separating *hyperplane*, a higher-dimensional generalization, so that it optimally discriminates two or more classes. During a minimization procedure, the hyperplane is tuned so that the SVM model generalization error is minimized, thus achieving an optimal solution to the classification problem:

$$f(\vec{x}) = \langle \vec{w}, \vec{x} \rangle + \vec{b} \qquad [16]$$

Here, $f(\vec{x})$ defines a hyperplane with parameters $\vec{w}$ (the direction perpendicular to the hyperplane), input data $\vec{x} = (x_1, \ldots, x_n) \in X$ and $\vec{b}$ (position vector). However, this classification may be difficult since real-world data are noisy, which limits the accuracy with which the resulting hyperplane can be determined in feature space[67]. One approach to address this problem is to transform the original input space into a higher dimensional space by using so–called *kernel* functions that optimally discriminates the two classes. The choice of kernel function has to be selected carefully in order to avoid poor discrimination power[69]

and a non-linear radial basis function (RBF) kernel was found to conform to the low-dimension histogram data used in our study[70]:

$$K(x_i, x_y) = e^{-\gamma \|x_i - x_j\|^2}, \qquad \gamma > 0 \quad [17]$$

Here, a Gaussian based RBF kernel is used. The parameter γ is related to the standard deviation, i.e. the width of the Gaussian distribution. Maximal reduction of the generalization error is reached by optimizing the distance between the margin of the function separating the classes (functional margin) and the input examples during a training process. Figure 11 depicts an illustration of the optimal margin (dotted lines) and hyperplane (solid line) for a 2 dimensional binary example.

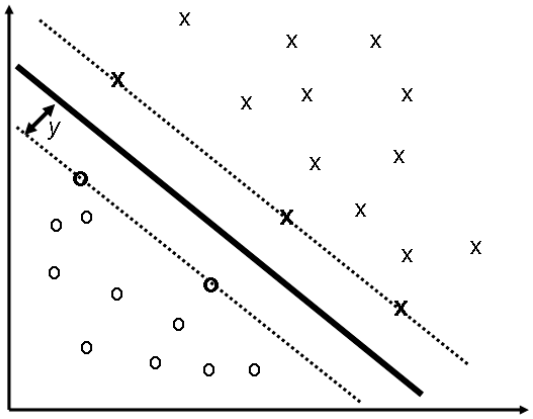


Figure 11: Illustration of the optimal margins (dotted lines) and hyperplane y (solid line) for arbitrary binary data (◦, x). The data points defining the margins are called support vectors and their positions should be such as to maximize the distance between the support vectors and the hyperplane.

Although numerous SVM models exists[69], a SVM model known as υ-SVM was considered practical for our data[71,72]. The υ-SVM has the advantage of using a parameter, υ, which is related to the number of support vectors used and the ratio of the training error. The range of the υ-parameters is between [0,1]. A low υ value result in less softness of the classification margins (i.e. less generalized) but fewer misclassifications, whereas a high υ value result in more softness of the classification margins but with more misclassifications allowed. The optimal υ-SVM model has a υ value giving minimal misclassifications and maximal generalization.

For all values of υ and γ, the diagnostic accuracy of the υ-SVM model on the test dataset was evaluated by assessing true positive rates (TPR) and true negative rates (TNR) of the glioma grading using histology as a reference. Based on the optimal υ and γ parameter values only, the effect of increasing the sample size (including more patients) of the training dataset was evaluated using linear regression and Fleiss-Kappa (κ) statistics[57] as described in Paper I.

Results:

To correct for over-adaptation of the υ-SVM model and to provide high generalization, the optimal υ value (0.56) was found within 99% of the maximal filtered accuracy value. Similarly, the optimal γ value (0.0065) was found within 99% of the maximal filtered accuracy value. Using these values, the TPR and TNR of the model on the test dataset was 0.76 and 0.82, respectively.

A significant increase (R^2=0.46, P<0.001) in the Fleiss-Kappa agreement values was observed when including more patients in the training dataset (5 patients; κ=0.139 up to 53 patients; κ=0.766). As shown in Figure 12, including more patients in the training dataset resulted in a significant increase in TPR (R^2=0.311, P<0.001). The TNR however, remained unchanged (R^2=0.004, P=0.692).

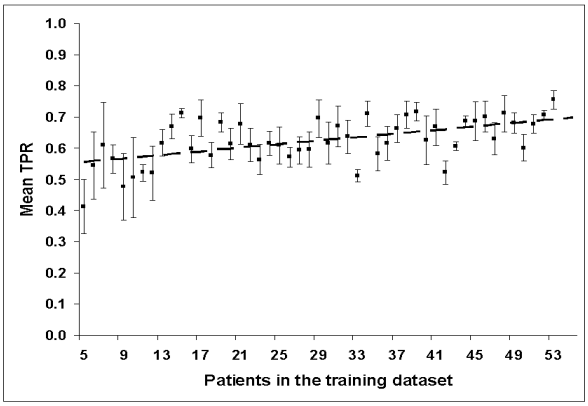


Figure 12: For an increasing number of patients included in the training dataset (5-53), the scatter plot shows mean TPR across the four observers when using the υ-SVM model on the test dataset (using optimal υ and γ values only). The error bars indicate standard errors of the mean values. As illustrated by the dotted trend line, TPR increased significantly with sample size (R^2=0.311, P<0.001).

7. Discussion

7.1 General discussion

The infiltrative and aggressive natures of many intra-axial brain tumors make these neoplasms a formidable challenge to treatment. Both accurate preoperative assessment of tumor location and function as well as lifelong monitoring of postoperative tumor recurrence and growth are important. Non-invasive methods for glioma characterization are desirable as sampling error and inaccessible tumors may limit stereotactic biopsy sampling. It has been suggested that DSC imaging can provide an alternative to histopathology for early identification of high- and low-risk patient groups with respect to time to progression of glioma grade or patient survival[10]. Hence, including DSC imaging on a routine basis will provide clinicians with important information concerning treatment planning[9].

Furthermore, as most patients suspected of an intra-axial tumor will undergo surgery and subsequently receive a histopathological diagnosis, a second important application of MR-based glioma imaging methods will be related to the evaluation of therapy response from anti-angiogenetic markers[39,47,73,74]. Invasive procedures to follow tumor growth and treatment response cannot be performed repeatedly, and non-invasive imaging methods may convey similar information with minimal patient discomfort. On conventional anatomical MR images, tumor response to therapy is currently based on changes in tumor volume or contrast enhancement patterns on T1-weighted images[6]. These changes may take time to develop and slow the decision making as to whether the ongoing treatment strategy should be continued, terminated or changed. It has been suggested, however, that DSC-based glioma imaging can demonstrate significant increases in tumor rCBV, secondary to malignant transformation, up to 12 months before contrast enhancement changes are apparent on T1-weighted MR images[75]. With such exciting results, the need for robust, reproducible DSC-based glioma characterization methods with short processing times and minimal user-dependence is evident. This argument is further strengthened by the focus on cost-benefit and cost-effectiveness in the medical workflow.

7.2 The histogram analysis method (Paper I)

Results from histopathology show that low-grade gliomas tend to incorporate preexisting vessels whereas high-grade gliomas no longer can rely on this vascular supply and consequently develop new vessels[74,76]. Tumor angiogenesis with vessel tortuosity, apoptosis and necrotic tumor components will result in reduced intra-tumoral vascular homogeneity. To exploit this phenomenon and to address some of the shortcomings of the hot-spot method discussed in this thesis, the histogram analysis approach has been proposed as an alternative method to grade gliomas from DSC imaging. Instead of measuring a single $rCBV_{max}$ value, it is suggested that assessing the complete distribution of glioma rCBV values from either a single MR image slice or the complete glioma volume provides a more robust and reproducible estimate of tumor malignancy. In concordance with histopathology, and as suggested by others[77], our results from the histogram analysis show that the rCBV distribution in a high-grade glioma conveys a more heterogeneous spectrum of rCBV values than a low-grade glioma.

The most important result from our study was that the interobserver agreement between the four observers for the histogram method was almost perfect. As confirmed by others[78], the histogram method may be less sensitive to user-bias in glioma grading thereby providing a more robust method for potential life-long monitoring of treatment response from angiogenetic inhibitors which require low observer variability. Furthermore, compared to the reference hot-spot method[58], our study and other studies using a histogram analysis approach convey similar or higher diagnostic accuracy values than the reference method[78,79]. As shown in Paper I, at similar specificity (~ 83%), the sensitivity of the histogram method was higher for all observers (90%) compared to the reference hot-spot method (55-76%). In another study excluding oligodendroglial tumors[78], a battery of histogram measures was compared showing that mean values and standard deviations of the histogram distribution also correlated to glioma grade. In that study, however, the histogram peak height was not correlated to tumor grade. With the other histogram metrics showing good correlation, this discrepancy from our results is somewhat surprising and may be related to the choice of histogram bin numbers (40 bins) and the use of an ellipsoid tumor ROI from a single slice. In our study, an iterative procedure was used to derive the histogram bin number giving highest diagnostic accuracy (108 bins) and the histogram signatures included $rCBV_{max}$ values from whole tumor ROIs with vessels excluded.

Furthermore, the shape of the normalized rCBV histogram distribution for a specific glioma is fixed even though the rCBV values are in arbitrary units. Thus, the histogram method can in theory be independent of reference tissue[79]. Because rCBV values in gray matter of a healthy individual is approximately twice as high as rCBV values in white matter, the hot-spot method is however, critically dependent on correct selection of reference tissue because the determination of $rCBV_{max}$ is based on this parameter alone. Nevertheless, because estimation of CBV scales with global flow F_a as shown in equation [11], a normalization procedure may be warranted in order to reduce inter-patient variations due to differences in cardiac output. Also, the normalization procedure may reduce the effect of variations in the dynamic response between image slices.

The histogram analysis method has also been used to differentiate between low-grade oligodendroglial tumors *with* or *without* loss of LOH on 1p/19q[80]. In concordance with the higher $rCBV_{max}$ values observed in low-grade oligodendrogliomas with LOH on 1p/19q[53], it has been reported that low-grade oligodendrogliomas with LOH on 1p/19q convey a more heterogeneous distribution of rCBV values compared to low-grade oligodendrogliomas without LOH on 1p/19q[80]. However, contrary to studies using the hot-spot method[53], this difference is not sufficient to consistently misclassify a low-grade oligodendroglioma as high-grade. This feature of the histogram method is attractive, in that a glioma grading method aiming at optimal differentiation between high- and low-grade gliomas should have minimal variations within each grade.

A controversial issue related to the histogram method is the choice of tumor area included in the histogram analysis[6]. Although identification of *correct* tumor areas as seen on anatomical MR images is notoriously difficult[46,47], approaches using ellipsoid ROIs or peritumoral glioma ROIs on a single MR image[78], freehand ROI drawings in all tumor-slices[54] or semi-automatically segmented ROIs[79] all show high diagnostic accuracy values. As shown in Papers I and II, even though glioma ROI definitions between the four observers show large variations, the high diagnostic accuracy obtained by all observers suggest that potential imperfect glioma delineation is relatively unimportant when using the histogram method. Also, the large number of pixels included in the glioma ROI may limit the influence of rCBV values from non-tumor pixels erroneously included in the tumor ROI.

A further challenge with the histogram method was that the observers reported using longer analysis time per patient for the histogram method compared to the hot-spot method. The reason for this was that the tumor volume had to be identified in every slice. However, the methods were considered equally difficult to perform. With recent advents in advanced tumor segmentation routines from anatomical MR images, an automated method for total volume identification as suggested in Paper II may further improve the utility of the histogram analysis method.

7.3 Automatic tumor segmentation in DSC-based glioma grading (Paper II)

In spite of obvious potential advantages, automatic tumor segmentation procedures in DSC-based glioma grading have received little attention. The additional gain of replacing the manual tumor definition with an automatic procedure includes removal of inter-observer variations, time efficiency and standardized criteria for tumor characterization. A reason for the lack of automated segmentation procedures in DSC-based glioma imaging might be that a traditional brain tumor MR imaging protocol used in the clinical routine typically consist of several 2D anatomical MR images sets which are sub-optimal for tumor segmentation compared to 3D MR images[81,82]. Furthermore, an automated tumor segmentation model must include tumor areas with contrast enhancement without including contrast enhanced blood vessels. While automatic methods for tumor identification in DSC-based glioma grading are rare, several approaches based on supervised (template based) and unsupervised segmentation methods for identification of brain tissue structures from anatomical MR images only, have been proposed. Automatic tissue or tumor segmentation based on multi-spectral data analysis[83], neural networking[84], support vector machines[85] and knowledge-based FCM clustering techniques[59,60] all show great promise. While a supervised segmentation routine require a very large volume of data in order to create a robust training template[81,84], an unsupervised procedure combining identification of various image classes and prior knowledge of image properties is not dependent on sample size.

One unsupervised segmentation approach yielding promising results when segmenting grade IV glioblastoma from anatomical MR images is the FCM clustering approach. In one study[60], a tumor segmentation method based on multi-spectral analysis of multiple MR image types conveyed good correspondence to radiologist-labeled ‘ground truth’ tumor volumes. In Paper II, a similar approach based on knowledge-based FCM clustering was

used. The main advantage of this approach is that the proposed segmentation routine is user-independent, thereby removing user-induced variations in the glioma grading. A second attractive feature of the FCM procedure is that it is tolerant of imprecise image data, often experienced in heterogeneous tumor tissue. Our results suggest that using the proposed segmentation routine on DSC-based glioma grading correspond well with the results using manually defined tumor volumes. The diagnostic accuracy values obtained in our study were similar to those reported in the literature and there was no significant difference between the A_z values of the automatic and manual method[34,41,86,87]. The significantly lower PPV of the FCM clustering routine compared to the manual method suggest that the automatic segmentation routine is less conservative, consequently including more hyper-intense areas as seen on the T2-weighted and FLAIR images. Although the metabolic active tumor area is mainly restricted to the visible T2 tumor border[46,47,88], the infiltrating nature of gliomas should in turn favor a less conservative method for tumor delineation such as the proposed automatic segmentation method.

A potential drawback with knowledge-based FCM is that a perfect delineation between the resulting tissue classes is difficult since each pixel can belong to more than one cluster class. Hence, although not a focus of Paper II, the manual and automatic segmentation routine used in our study may not be an adequate measure of tumor volume for quantitative assessment of tumor growth and for aiding neurosurgeons intra-operatively. However, it was shown in Papers I and II that imperfect glioma delineation is relatively unimportant in pre-surgical glioma grading by histogram analysis of DSC images. Thus, using a histogram analysis method on DSC values from automatically segmented tumor volumes should favor the unsupervised knowledge-based FCM approach used in Paper II.

Also, no special care was taken to exclude rCBV pixels from large vessels within the segmented glioma volume. Although large vessels can be appreciated as dark areas on T2-weighted images and thus should be excluded during the knowledge-based operations[88,89], it is well known that the T2*-effect from large vessels in a GRE-EPI sequence result in an over-estimation of the intravascular susceptibility effect in pixels adjacent to large vessels[49,50]. The consequence of this is overestimated rCBV values in pixels not recognized as vessel tissue by the segmentation procedure. To address this problem, we have proposed an automatic vessel removal approach based on cluster analysis of the DSC dynamic first-pass parameters (Paper III).

7.4 Automatic vessel segmentation (Paper III)

Presence of elevated rCBV values from large vessels within the tumor area may confound the glioma grading[11,48,58]. Low-grade gliomas are typically located in the frontal- or temporal lobes[2,89,90], in close proximity to the anterior- or middle cerebral arteries, which might result in an overestimation of glioma grade[48]. To correct for this, large vessels within the tumor area can be manually excluded from the glioma ROI prior to glioma grading[54,58], which make current glioma grading methods inherently user-dependent. However, manual (Paper I) or automatic (Paper II) glioma segmentation from anatomical MR images do not account for the extravascular susceptibility effect in DSC imaging[49]. A second approach is to introduce a cut-off rCBV value to remove the high-end spectrum of rCBV values thought to represent vessels[78]. This approach however, does not discriminate between arterial, venous and tissue image pixels as the AUC can convey similar characteristics even though the shapes of the first-pass curves are different.

The proposed vessel segmentation technique in Paper III is fast and provides a direct estimate of pixels which are actually affected by the vascular susceptibility effect. Hence, the need for an expert user with good anatomical knowledge to mask out vessel regions is removed. In the case of both manually defined and automatically segmented tumor area identification, tumor ROI drawings with vessels included are easier and faster to perform compared to tumor ROIs with vessel excluded. In clinical routine, this feature is attractive as it reduces the time spent on post-processing and potentially increases the reproducibility of a DSC-based glioma grading method.

The diagnostic accuracy values obtained in our study was similar to other studies using the histogram method on gliomas[54,78,87]. Compared to the manual reference method, the use of a pre-mask to exclude areas of elevated MTT values, secondary to glioma malignancy, was found to improve the diagnostic accuracy of the glioma grading. The uncorrected vessel masking procedure, however, resulted in reduced sensitivity, NPV and PPV whereas the specificity increased. The increased number of false negatives and reduced number of false positives suggest that the uncorrected vessel masking routine is too radical in removing rCBV values in the tumor bed. This result suggests that the main challenge of the proposed vessel segmentation routine in glioma grading is to differentiate vessels from elevated rCBV values at the capillary level. A similar result was observed when using the SI corrected pre-mask. In order to avoid rapid signal changes between the first image (with full

magnetization) and subsequent (saturated) baseline images, the DSC sequence used in our study included several dummy scans resulting in mainly proton density weighted baseline images. In spite of the GRE-EPI sequence being strongly T2*-weighted in steady state, the image contrast in the very first image (following a single 90 degree RF pulse) will have a significant T2-weighting. For the purpose of obtaining a SI pre-mask based on differences in T2 relaxation times, it would therefore be a clear advantage to include this unsaturated EPI image in the analysis. Nevertheless, in DSC-based imaging of patients suggestive of non-neovascular diseases, such as stroke or multiple sclerosis, the radical vessel removal of the uncorrected or SI corrected vessel masking procedure may be attractive.

In our study, the vessel segmentation procedure was based on GRE-EPI imaging. Alternatively, SE-EPI imaging alone, or as a combination of both, have been proposed in order to reduce the sensitivity to susceptibility effects and thereby reducing the influence of macroscopic vessels in the resulting perfusion maps[11,28-31,91]. The disadvantage, however, as discussed in this thesis, is lower contrast agent sensitivity (effective contrast relaxivity) especially in normal appearing brain tissue structures and tumor tissue without contrast leakage and limited brain coverage for a given temporal resolution. Nevertheless, as shown in Figure 4, the bolus shape of arteries, veins and brain tissue from both techniques convey fairly similar characteristics and the proposed vessel segmentation technique may provide similar results in DSC-based glioma grading using SE-EPI. Thus, a study applying the vessel segmentation procedure in SE-EPI glioma grading is warranted.

Furthermore, in order to obtain optimal results, the proposed vessel segmentation procedure should be applied to DSC images with high spatial and temporal resolution. In particular, because separation of vessels from vascular tissue is mainly based on small differences in contrast dynamics, the temporal resolution should be as low as possible in order to increase the sampling steps. Although this would imply a compromise with the temporal resolution, the same principle applies for the spatial resolution. As the size of the capillaries is typically below 1mm, the image resolution should be as high as possible.

7.5 Predictive modeling (Paper IV)

An attractive potential of any diagnostic method is the ability to use prior knowledge from previous patients to predict status or outcome of new patients or to compare MR images of a

single patient from different examination dates. However, the use of predictive modeling in glioma grading from DSC imaging has not received much attention. A reason for this may be that current grading methods rely on threshold values between patient groups that are difficult to generalize into an accurate predictive model. As shown in numerous studies using the hot-spot method from DSC imaging, current threshold values show large variations[40,42,54]. The histogram analysis method however, may provide a more attractive starting point for predictive modeling because each new case is described by the normalized histogram *signature curve* of the rCBV distribution within the glioma volume. By using this signature curve rather than a single threshold, a more robust base for prediction is created, which can readily be compared to a database of histologically confirmed gliomas with signatures representative for each glioma grade.

One potential feature of the SVM model is that multiple channels of information can be included. In addition to MR perfusion parameters described in this thesis, a standard MR imaging protocol in patients suspected of a glioma may consist of conventional MR imaging[89], MR diffusion imaging[92] and MR spectroscopy[41] all with diagnostic parameters that can be included in the analysis. Also, measures of tumor growth and location as well as information on patient age, gender and treatment strategies are all important predictors in the continued patient care and in the decision making of further therapy

The perfect accuracy of 100% obtained in Paper IV when using the SVM model to retrospectively grade 53 gliomas as either high-grade or low-grade, and the corresponding lower accuracy values obtained when using the SVM model on new data obtained by a new observer (TPR = 0.76 and TNR 0.82), illustrate the problem of over-adaptation which is inherent in any retrospective glioma grading study. Thus, although this problem is reduced with increasing sample size, the development of a robust glioma grading method should not rely solely on retrospective analysis. Furthermore, the availability of the proposed SVM model to third-part users should be investigated across institutions with different MR machine vendors, at different MR field strengths and using different imaging protocols. For the SVM model to become independent of these parameters and to have clinical utility, the diagnostic accuracy of the model should be as high as possible. Thus, the accuracy of about 0.8 obtained with the proposed SVM model may not be sufficient for extensive clinical use. However, TPR was shown to significantly increase with increasing sample size suggesting that the diagnostic accuracy will reach an acceptable level with a sufficient increase in the

sample size. In addition, the inter-observer agreement between the four observers increased significantly with increasing sample size suggesting higher generalization and robustness of the SVM model with increasing sample size of the training dataset.

Although the standard RBF kernel used in our study has previously been shown to fit histogram data[70], it is likely that the model could be further optimized by a more careful selection of the kernel function. Especially, when including additional MR image metrics in the SVM model, the choice of kernel function will influence the final result. When the number of feature vectors is much larger than the number of subjects in the training and test dataset, mapping the data to a higher dimensional space using a non-linear kernel may reduce the performance and speed of the SVM model[69]. Thus, a final SVM model applied to the clinical routine should include an iterative procedure that not only updates the support vector parameters (ν, γ) as more patients and imaging feature vectors are included in the training database, but also iteratively derive the optimal kernel function with respect to glioma grade, tumor progression or patient survival.

8. Conclusion and future aspects

8.1 Conclusion

Current methods for DSC-based glioma grading suffer from high user-dependence, a clear limitation towards assessment of whether the technique may actually have an impact on patient care[93]. To address this issue, this thesis has focused on introducing and validating alternative approaches with the aim of minimizing user-bias in glioma grading from DSC images. One of these approaches is the analysis of cerebral blood volume heterogeneity as shown in Paper I. This method is based on the hypothesis that blood volume heterogeneity is related to vascular proliferation secondary to tumor malignancy. While the selection of multiple, small region-of-interests in the tumor area (hot-spot method) is inherently prone to user-bias, the large number of pixels included in the whole-tumor histogram analysis results in a procedure relatively insensitive to suboptimal tumor area selection. The histogram approach may, therefore, be less dependent on expert users, and more amenable to automatic tumor segmentation approaches as shown in Paper II. Here, the use of an automated tumor segmentation procedure provides a further reduction in user-dependency. Thus, in combination with the automatic vessel segmentation procedure shown in Paper III, implementing these methods in the clinical routine is expected to improve the diagnostic efficacy of DSC-based glioma imaging. Furthermore, the complete distribution of blood volume values in terms of a histogram signature provides a strong measure for comparing MR perfusion metrics from different patients and examination dates. From this, large databases of histogram signatures can form the basis for predictive models as shown in Paper IV.

In conclusion, our results suggest that the proposed methods provide objective measures for tumor vascularity and malignancy with minimal user-bias, an important step towards a fully automated MR-based glioma characterization regime. Furthermore, the methods described in this thesis can readily be implemented in clinical routine; thereby, providing fully automated methods amendable for third-part users with little experience with DSC-based glioma grading.

8.2 Future aspects

The studies presented in this thesis included patient data from Rikshospitalet University Hospital only. However, in order to validate the stability of the proposed methods, measures of reproducibility and diagnostic efficacy should be assessed through large, multi-institutional studies. In this, the methods proposed in this thesis may also be applied to other MR perfusion imaging techniques, such as SE-EPI, T1-weighted perfusion and potentially ASL.

Furthermore, with the current advents of automatic glioma grading methods from DSC-imaging with minimal user-dependency, the need for a systematic framework encompassing the complete chain of techniques is warranted. This may be achieved by developing a computer aided diagnostics (CAD) software program, involving every step from MR examination to a final report summarizing the results from the DSC-based glioma analysis. In theory, this report can be sent back to the hospital-wide picture archiving and communication system (PACS) as an additional source of information for the clinician. It is hypothesized that the availability of the proposed methods directly in PACS will increase their acceptability and clinical utility substantially. Thus, the methods may also be capitalized in new commercial software for the benefit of health care.

Tables

Table 1. Summary of patients with histopathological results included in papers I-IV

	Patient population	Histopathological diagnosis*
Paper I:	53 patients (24 females, 29 males) mean age 48 yrs, range 14-76 yrs	LGGs: 2 grade I, 22 grade II HGGs: 10 grade III, 19 grade IV
Paper II:	50 patients (21 females, 29 males) mean age 49 yrs, range 14-76 yrs	LGGs: 1 grade I, 21 grade II HGGs: 9 grade III, 19 grade IV
Paper III:	77 patients (32 females, 45 males) mean age 52 yrs, range 19-78 yrs	LGGs: 32 grade II HGGs: 10 grade III, 35 grade IV
Paper IV:	86 patients (37 females, 49 males) mean age 49 yrs, range 9-78 yrs	LGGs: 4 grade I, 34 grade II HGGs: 11 grade III, 37 grade IV

LGGs = low-grade gliomas, HGGs = high-grade gliomas, yrs = years

*Using the WHO classification of the central nervous system

Table 2. Overview of pulse sequences used in papers I-IV*

Anatomcial MR sequences		
T2-w FSE	**T1-w SE**	**FLAIR** (Paper II only)
TR: 4000 ms	TR: 500 ms	TR: 9050 ms
TE: 104 ms	TE: 7.7 ms	TE: 114 ms
Flip: 150º	Flip: 90º	Flip: 150º
VS: 0.45x0.45x5 mm^3	VS: 0.45x0.45x5 mm^3	VS: 0.9x0.9x5 mm^3
19 axial slices	19 axial slices	25 coronal slices
	pre- and post-contrast	TI = 1500 ms

Dynamic Susceptibility Contrast (DSC) sequence
GRE-EPI**
TR: 1430 - 1720 ms
TE: 46 - 48 ms
Flip: 90º
VS: 1.80x1.80x5mm^3
12 - 14 axial slices X 50 images per slice
0.2 mmol/kg of gadobutrol (5 mL/sec)

TR = repetition time, TE = echo time, Flip = flip angle, VS = voxel size, TI = inversion time, SE = spin-echo,

FSE = fast spin-echo, GRE-EPI = gradient-echo echo-planar imaging

*All imaging was performed on a 1.5Tesla Siemens Sonata, Symphony or Avanto (Siemens AG, Erlangen, Germany) using an 8-channel- (Symphony/Sonata) or a 12-channel (Avanto) head-coil.

**GRE-EPI parameters values were not constant because of varying tumor sizes

References

1. Ohgaki H, Kleihues P. Population-based studies on incidence, survival rates, and genetic alterations in astrocytic and oligodendroglial gliomas. J Neuropathol Exp Neurol 2005 Jun;64(6):479-489.

2. Kleihues P, Cavenee W. Astrocytic tumors & Oligodendroglial tumors and mixed gliomas. *The WHO classification of tumors of the nervous system*. 2 ed. Lyon: International Agency for Research on Cancer; 2000. p 9-70.

3. Woodworth G, McGirt MJ, Samdani A, Garonzik I, Olivi A, Weingart JD. Accuracy of frameless and frame-based image-guided stereotactic brain biopsy in the diagnosis of glioma: comparison of biopsy and open resection specimen. Neurol Res 2005 Jun;27(4):358-362.

4. Lev MH, Rosen BR. Clinical applications of intracranial perfusion MR imaging. Neuroimaging Clin N Am 1999 May;9(2):309-331.

5. Essig M, Weber MA, von Tengg-Kobligk H, Knopp MV, Yuh WT, Giesel FL. Contrast-enhanced magnetic resonance imaging of central nervous system tumors: agents, mechanisms, and applications. Top Magn Reson Imaging 2006 Apr;17(2):89-106.

6. Weber MA, Giesel FL, Stieltjes B. MRI for identification of progression in brain tumors: from morphology to function. Expert Rev Neurother 2008 Oct;8(10):1507-1525.

7. Covarrubias DJ, Rosen BR, Lev MH. Dynamic magnetic resonance perfusion imaging of brain tumors. Oncologist 2004;9(5):528-537.

8. Cha S. Perfusion MR imaging: basic principles and clinical applications. Magn Reson Imaging Clin N Am 2003 Aug;11(3):403-413.

9. Law M, Oh S, Johnson G, Babb JS, Zagzag D, Golfinos J, Kelly PJ. Perfusion magnetic resonance imaging predicts patient outcome as an adjunct to histopathology: a second reference standard in the surgical and nonsurgical treatment of low-grade gliomas. Neurosurgery 2006 Jun;58(6):1099-1107.

10. Law M, Young RJ, Babb JS, Peccerelli N, Chheang S, Gruber ML, Miller DC, Golfinos JG, Zagzag D, Johnson G. Gliomas: predicting time to progression or survival with cerebral blood volume measurements at dynamic susceptibility-weighted contrast-enhanced perfusion MR imaging. Radiology 2008 May;247(2):490-498.

11. Lev MH, Ozsunar Y, Henson JW, Rasheed AA, Barest GD, Harsh GR, Fitzek MM, Chiocca EA, Rabinov JD, Csavoy AN, Rosen BR, Hochberg FH, Schaefer PW, Gonzalez RG. Glial tumor grading and outcome prediction using dynamic spin-echo MR susceptibility mapping compared with conventional contrast-enhanced MR: confounding effect of elevated rCBV of oligodendrogliomas [corrected]. AJNR Am J Neuroradiol 2004 Feb;25(2):214-221.

12. Maia AC, Jr., Malheiros SM, da Rocha AJ, Stavale JN, Guimaraes IF, Borges LR, Santos AJ, da Silva CJ, de Melo JG, Lanzoni OP, Gabbai AA, Ferraz FA. Stereotactic biopsy guidance in adults with supratentorial nonenhancing gliomas: role of perfusion-weighted magnetic resonance imaging. J Neurosurg 2004 Dec;101(6):970-976.

13. Vlaardingerbroek MT, den Boer JA. Magnetic Resonance Imaging. Theory and Practice. 2nd ed. Berlin: Springer; 1999.

14. Bloch F. Nuclear induction. Phys Rev 1946;70:460-474.

15. Purcell EM, Torrey HC, Pound R.V. Resonance absorption by nuclear magnetic moments in solid. Phys Rev 1946;69:37-38.

16. Lauterbur P. Image formation by induced local interactions: examples employing nuclear magnetic resonance. Nature 1973;242:190-191.

17. Mansfield P, Maudsley AA. Medical imaging by NMR. Br J Radiol 1977 Mar;50(591):188-194.

18. Barbier EL, Lamalle L, Decorps M. Methodology of brain perfusion imaging. J Magn Reson Imaging 2001 Apr;13(4):496-520.

19. Williams DS, Detre JA, Leigh JS, Koretsky AP. Magnetic resonance imaging of perfusion using spin inversion of arterial water. Proc Natl Acad Sci U S A 1992 Jan;89(1):212-216.

20. Meier P, Zierler KL. On the theory of the indicator-dilution method for measurement of blood flow and volume. J Appl Physiol 1954 Jun;6(12):731-744.

21. Tofts PS, Kermode AG. Measurement of the blood-brain barrier permeability and leakage space using dynamic MR imaging. 1. Fundamental concepts. Magn Reson Med 1991 Feb;17(2):357-367.

22. Rosen BR, Belliveau JW, Vevea JM, Brady TJ. Perfusion imaging with NMR contrast agents. Magn Reson Med 1990 May;14(2):249-265.

23. Bjornerud A, Johansson LO, Briley-Saebo K, Ahlstrom HK. Assessment of T1 and T2* effects in vivo and ex vivo using iron oxide nanoparticles in steady state--dependence on blood volume and water exchange. Magn Reson Med 2002 Mar;47(3):461-471.

24. Ostergaard L, Weisskoff RM, Chesler DA, Gyldensted C, Rosen BR. High resolution measurement of cerebral blood flow using intravascular tracer bolus passages. Part I: Mathematical approach and statistical analysis. Magn Reson Med 1996 Nov;36(5):715-725.

25. Weisskoff RM, Chesler D, Boxerman JL, Rosen BR. Pitfalls in MR measurement of tissue blood flow with intravascular tracers: which mean transit time? Magn Reson Med 1993 Apr;29(4):553-558.

26. Zierler KL. Equations for measuring blood flow by external monitoring of radioisotopes. Circ Res 1965 Apr;16:309-321.

27. Calamante F. Bolus dispersion issues related to the quantification of perfusion MRI data. J Magn Reson Imaging 2005 Dec;22(6):718-722.

28. Ogawa S, Menon RS, Tank DW, Kim SG, Merkle H, Ellermann JM, Ugurbil K. Functional brain mapping by blood oxygenation level-dependent contrast magnetic resonance imaging. A comparison of signal characteristics with a biophysical model. Biophys J 1993 Mar;64(3):803-812.

29. Kennan RP, Zhong J, Gore JC. Intravascular susceptibility contrast mechanisms in tissues. Magn Reson Med 1994 Jan;31(1):9-21.

30. Schmainda KM, Rand SD, Joseph AM, Lund R, Ward BD, Pathak AP, Ulmer JL, Badruddoja MA, Krouwer HG. Characterization of a first-pass gradient-echo spin-echo method to predict brain tumor grade and angiogenesis. AJNR Am J Neuroradiol 2004 Oct;25(9):1524-1532.

31. Sugahara T, Korogi Y, Kochi M, Ushio Y, Takahashi M. Perfusion-sensitive MR imaging of gliomas: comparison between gradient-echo and spin-echo echo-planar imaging techniques. AJNR Am J Neuroradiol 2001 Aug;22(7):1306-1315.

32. Aronen HJ, Perkio J. Dynamic susceptibility contrast MRI of gliomas. Neuroimaging Clin N Am 2002 Nov;12(4):501-523.

33. Paulson ES, Schmainda KM. Comparison of dynamic susceptibility-weighted contrast-enhanced MR methods: recommendations for measuring relative cerebral blood volume in brain tumors. Radiology 2008 Nov;249(2):601-613.

34. Knopp EA, Cha S, Johnson G, Mazumdar A, Golfinos JG, Zagzag D, Miller DC, Kelly PJ, Kricheff II. Glial neoplasms: dynamic contrast-enhanced T2*-weighted MR imaging. Radiology 1999 Jun;211(3):791-798.

35. Boxerman JL, Schmainda KM, Weisskoff RM. Relative cerebral blood volume maps corrected for contrast agent extravasation significantly correlate with glioma tumor grade, whereas uncorrected maps do not. AJNR Am J Neuroradiol 2006 Apr;27(4):859-867.

36. Edelman RR, Mattle HP, Atkinson DJ, Hill T, Finn JP, Mayman C, Ronthal M, Hoogewoud HM, Kleefield J. Cerebral blood flow: assessment with dynamic contrast-enhanced T2*-weighted MR imaging at 1.5 T. Radiology 1990 Jul;176(1):211-220.

37. Aronen HJ, Gazit IE, Louis DN, Buchbinder BR, Pardo FS, Weisskoff RM, Harsh GR, Cosgrove GR, Halpern EF, Hochberg FH, Rosen BR. Cerebral blood volume maps of gliomas: comparison with tumor grade and histologic findings. Radiology 1994 Apr;191(1):41-51.

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

39. Cha S, Knopp EA, Johnson G, Litt A, Glass J, Gruber ML, Lu S, Zagzag D. Dynamic contrast-enhanced T2-weighted MR imaging of recurrent malignant gliomas treated with thalidomide and carboplatin. AJNR Am J Neuroradiol 2000 May;21(5):881-890.

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

41. Law M, Yang S, Wang H, Babb JS, Johnson G, Cha S, Knopp EA, Zagzag D. Glioma grading: sensitivity, specificity, and predictive values of perfusion MR imaging and proton MR spectroscopic imaging compared with conventional MR imaging. AJNR Am J Neuroradiol 2003 Nov;24(10):1989-1998.

42. Shin JH, Lee HK, Kwun BD, Kim JS, Kang W, Choi CG, Suh DC. Using relative cerebral blood flow and volume to evaluate the histopathologic grade of cerebral gliomas: preliminary results. AJR Am J Roentgenol 2002 Sep;179(3):783-789.

43. Law M, Young R, Babb J, Rad M, Sasaki T, Zagzag D, Johnson G. Comparing perfusion metrics obtained from a single compartment versus pharmacokinetic modeling methods using dynamic susceptibility contrast-enhanced perfusion MR imaging with glioma grade. AJNR Am J Neuroradiol 2006 Oct;27(9):1975-1982.

44. Roberts HC, Roberts TP, Brasch RC, Dillon WP. Quantitative measurement of microvascular permeability in human brain tumors achieved using dynamic contrast-enhanced MR imaging: correlation with histologic grade. AJNR Am J Neuroradiol 2000 May;21(5):891-899.

45. Fryback DG, Thornbury JR. The efficacy of diagnostic imaging. Med Decis Making 1991 Apr;11(2):88-94.

46. Pirzkall A, Nelson SJ, McKnight TR, Takahashi MM, Li X, Graves EE, Verhey LJ, Wara WW, Larson DA, Sneed PK. Metabolic imaging of low-grade gliomas with three-dimensional magnetic resonance spectroscopy. Int J Radiat Oncol Biol Phys 2002 Aug;53(5):1254-1264.

47. Henson JW, Gaviani P, Gonzalez RG. MRI in treatment of adult gliomas. Lancet Oncol 2005 Mar;6(3):167-175.

48. Caseiras GB, Thornton JS, Yousry T, Benton C, Rees J, Waldman AD, Jager HR. Inclusion or exclusion of intratumoral vessels in relative cerebral blood volume characterization in low-grade gliomas: does it make a difference? AJNR Am J Neuroradiol 2008 Jun;29(6):1140-1141.

49. van Osch MJ, Vonken EJ, Bakker CJ, Viergever MA. Correcting partial volume artifacts of the arterial input function in quantitative cerebral perfusion MRI. Magn Reson Med 2001 Mar;45(3):477-485.

50. Boxerman JL, Hamberg LM, Rosen BR, Weisskoff RM. MR contrast due to intravascular magnetic susceptibility perturbations. Magn Reson Med 1995 Oct;34(4):555-566.

51. Cha S, Tihan T, Crawford F, Fischbein NJ, Chang S, Bollen A, Nelson SJ, Prados M, Berger MS, Dillon WP. Differentiation of low-grade oligodendrogliomas from low-grade astrocytomas by using quantitative blood-volume measurements derived from dynamic susceptibility contrast-enhanced MR imaging. AJNR Am J Neuroradiol 2005 Feb;26(2):266-273.

52. Jenkinson MD, Smith TS, Joyce KA, Fildes D, Broome J, du Plessis DG, Haylock B, Husband DJ, Warnke PC, Walker C. Cerebral blood volume, genotype and chemosensitivity in oligodendroglial tumours. Neuroradiology 2006 Oct;48(10):703-713.

53. Law M, Brodsky JE, Babb J, Rosenblum M, Miller DC, Zagzag D, Gruber ML, Johnson G. High cerebral blood volume in human gliomas predicts deletion of chromosome 1p: Preliminary results of molecular studies in gliomas with elevated perfusion. J Magn Reson Imaging 2007 Jun;25(6):1113-1119.

54. Emblem KE, Nedregaard B, Nome T, Due-Tonnessen P, Hald JK, Scheie D, Borota OC, Cvancarova M, Bjornerud A. Glioma grading by using histogram analysis of blood volume heterogeneity from MR-derived cerebral blood volume maps. Radiology 2008 Jun;247(3):808-817.

55. Seber GAF. Multivariate Observations. Hoboken, New York: John Wiley & Sons, Inc.; 1984.

56. Hartigan JA, Wong MA. A K-Means Clustering Algorithm. Applied Statistics 1979;28(1):100-108.

57. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977 Mar;33(1):159-174.

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

59. Bezdek JC, Hall LO, Clark MC, Goldgof DB, Clarke LP. Medical image analysis with fuzzy models. Stat Methods Med Res 1997 Sep;6(3):191-214.

60. Clark MC, Hall LO, Goldgof DB, Velthuizen R, Murtagh FR, Silbiger MS. Automatic tumor segmentation using knowledge-based techniques. IEEE Trans Med Imaging 1998 Apr;17(2):187-201.

61. Fletcher-Heath LM, Hall LO, Goldgof DB, Murtagh FR. Automatic segmentation of non-enhancing brain tumors in magnetic resonance images. Artif Intell Med 2001 Jan;21(1-3):43-63.

62. Zuiderveld K. Contrast Limited Adaptive Histogram Equalization. In: Heckbert PS, editor. Graphics Gems IV. AP Professional, Boston; 1994. p 474-85.

63. Ashburner J, Friston K. Multimodal image coregistration and partitioning--a unified framework. Neuroimage 1997 Oct;6(3):209-217.

64. Hadjiprocopis A, Rashid W, Tofts PS. Unbiased segmentation of diffusion-weighted magnetic resonance images of the brain using iterative clustering. Magn Reson Imaging 2005 Oct;23(8):877-885.

65. Mouridsen K, Christensen S, Gyldensted L, Ostergaard L. Automatic selection of arterial input function using cluster analysis. Magn Reson Med 2006 Mar;55(3):524-531.

66. Ostergaard L, Sorensen AG, Kwong KK, Weisskoff RM, Gyldensted C, Rosen BR. High resolution measurement of cerebral blood flow using intravascular tracer bolus passages. Part II: Experimental comparison and preliminary results. Magn Reson Med 1996 Nov;36(5):726-736.

67. Boser BE, Guyon I.M., Vapnik V.N. A training algorithm for optimal margin classifiers. Proceedings of the Fifth Annual Workshop on Computational Learning Theory 1992;pp:144-152.

68. Vapnik V.N. Universal Learning Technology: Support Vector Machines. NEC Journal of Advanced Technology 2005;2(2):137-144.

69. Chang CC, Lin CJ. Training nu-support vector classifiers: theory and algorithms. Neural Comput 2001 Sep;13(9):2119-2147.

70. Chapelle O, Haffner P, Vapnik VN. Support vector machines for histogram-based image classification. IEEE Trans Neural Netw 1999;10(5):1055-1064.

71. Fan RE, Chen PH, Lin CJ. Working set selection using second order information for training SVM. Journal of Machine Learning Research 2005;6:1889-1918.

72. Scholkopf B, Smola AJ, Williamson RC, Bartlett PL. New support vector algorithms. Neural Comput 2000 May;12(5):1207-1245.

73. Batchelor TT, Sorensen AG, di TE, Zhang WT, Duda DG, Cohen KS, Kozak KR, Cahill DP, Chen PJ, Zhu M, Ancukiewicz M, Mrugala MM, Plotkin S, Drappatz J, Louis DN, Ivy P, Scadden DT, Benner T, Loeffler JS, Wen PY, Jain RK. AZD2171, a pan-VEGF receptor tyrosine kinase inhibitor, normalizes tumor vasculature and alleviates edema in glioblastoma patients. Cancer Cell 2007 Jan;11(1):83-95.

74. Jain RK, di Tomaso E, Duda DG, Loeffler JS, Sorensen AG, Batchelor TT. Angiogenesis in brain tumours. Nat Rev Neurosci 2007 Aug;8(8):610-622.

75. Danchaivijitr N, Waldman AD, Tozer DJ, Benton CE, Brasil CG, Tofts PS, Rees JH, Jager HR. Low-grade gliomas: do changes in rCBV measurements at longitudinal perfusion-weighted MR imaging predict malignant transformation? Radiology 2008 Apr;247(1):170-178.

76. Folkerth RD. Histologic measures of angiogenesis in human primary brain tumors. Cancer Treat Res 2004;117:79-95.

77. Lupo JM, Cha S, Chang SM, Nelson SJ. Dynamic susceptibility-weighted perfusion imaging of high-grade gliomas: characterization of spatial heterogeneity. AJNR Am J Neuroradiol 2005 Jun;26(6):1446-1454.

78. Law M, Young R, Babb J, Pollack E, Johnson G. Histogram analysis versus region of interest analysis of dynamic susceptibility contrast perfusion MR imaging data in the grading of cerebral gliomas. AJNR Am J Neuroradiol 2007 Apr;28(4):761-766.

79. Slotboom J, Schaer R, Ozdoba C, Reinert M, Vajtai I, El-Koussy M, Kiefer C, Zbinden M, Schroth G, Wiest R. A novel method for analyzing DSCE-images with an application to tumor grading. Invest Radiol 2008 Dec;43(12):843-853.

80. Emblem KE, Scheie D, Due-Tonnessen P, Nedregaard B, Nome T, Hald JK, Beiske K, Meling TR, Bjornerud A. Histogram analysis of MR imaging-derived cerebral blood volume maps: combined glioma grading and identification of low-grade oligodendroglial subtypes. AJNR Am J Neuroradiol 2008 Oct;29(9):1664-1670.

81. Kaus MR, Warfield SK, Nabavi A, Black PM, Jolesz FA, Kikinis R. Automated segmentation of MR images of brain tumors. Radiology 2001 Feb;218(2):586-591.

82. Itskovich VV, Samber DD, Mani V, Aguinaldo JG, Fallon JT, Tang CY, Fuster V, Fayad ZA. Quantification of human atherosclerotic plaques using spatially enhanced cluster analysis of multicontrast-weighted magnetic resonance images. Magn Reson Med 2004 Sep;52(3):515-523.

83. Vannier MW, Butterfield RL, Jordan D, Murphy WA, Levitt RG, Gado M. Multispectral analysis of magnetic resonance images. Radiology 1985 Jan;154(1):221-224.

84. Dickson S, Thomas BT, Goddard P. Using neural networks to automatically detect brain tumours in MR images. Int J Neural Syst 1997 Feb;8(1):91-99.

85. Glotsos D, Spyridonos P, Cavouras D, Ravazoula P, Dadioti PA, Nikiforidis G. An image-analysis system based on support vector machines for automatic grade diagnosis of brain-tumour astrocytomas in clinical routine. Med Inform Internet Med 2005 Sep;30(3):179-193.

86. Mazzara GP, Velthuizen RP, Pearlman JL, Greenberg HM, Wagner H. Brain tumor target volume determination for radiation treatment planning through automated MRI segmentation. Int J Radiat Oncol Biol Phys 2004 May;59(1):300-312.

87. Young R, Babb J, Law M, Pollack E, Johnson G. Comparison of region-of-interest analysis with three different histogram analysis methods in the determination of perfusion metrics in patients with brain gliomas. J Magn Reson Imaging 2007 Oct;26(4):1053-1063.

88. Grier JT, Batchelor T. Low-grade gliomas in adults. Oncologist 2006 Jun;11(6):681-693.

89. Talos IF, Zou KH, Ohno-Machado L, Bhagwat JG, Kikinis R, Black PM, Jolesz FA. Supratentorial low-grade glioma resectability: statistical predictive analysis based on anatomic MR features and tumor characteristics. Radiology 2006 May;239(2):506-513.

90. Walker DG, Kaye AH. Low grade glial neoplasms. J Clin Neurosci 2003 Jan;10(1):1-13.

91. Donahue KM, Krouwer HG, Rand SD, Pathak AP, Marszalkowski CS, Censky SC, Prost RW. Utility of simultaneously acquired gradient-echo and spin-echo cerebral blood volume and morphology maps in brain tumor patients. Magn Reson Med 2000 Jun;43(6):845-853.

92. Lam WW, Poon WS, Metreweli C. Diffusion MR imaging in glioma: does it have any role in the pre-operation determination of grading of glioma? Clin Radiol 2002 Mar;57(3):219-225.

93. Sorensen AG. Perfusion MR imaging: moving forward. Radiology 2008 Nov;249(2):416-417.

Papers I-IV

I

Glioma Grading by Using Histogram Analysis of Blood Volume Heterogeneity from MR-derived Cerebral Blood Volume Maps[1]

Kyrre E. Emblem, MSc
Baard Nedregaard, MD
Terje Nome, MD
Paulina Due-Tonnessen, MD
John K. Hald, MD, PhD
David Scheie, MD
Olivera Casar Borota, MD
Milada Cvancarova, MSc
Atle Bjornerud, PhD

Purpose: To retrospectively compare the diagnostic accuracy of an alternative method used to grade gliomas that is based on histogram analysis of normalized cerebral blood volume (CBV) values from the entire tumor volume (obtained with the histogram method) with that of the hot-spot method, with histologic analysis as the reference standard.

Materials and Methods: The medical ethics committee approved this study, and all patients provided informed consent. Fifty-three patients (24 female, 29 male; mean age, 48 years; age range, 14–76 years) with histologically confirmed gliomas were examined with dynamic contrast material–enhanced 1.5-T magnetic resonance (MR) imaging. CBV maps were created and normalized to unaffected white matter (normalized CBV maps). Four neuroradiologists independently measured the distribution of whole-tumor normalized CBVs and analyzed this distribution by classifying the values into area-normalized bins. Glioma grading was performed by assessing the normalized peak height of the histogram distributions. Logistic regression analysis and interobserver agreement were used to compare the proposed method with a hot-spot method in which only the maximum normalized CBV was used.

Results: For the histogram method, diagnostic accuracy was independent of the observer. Interobserver agreement was almost perfect for the histogram method ($\kappa = 0.923$) and moderate for the hot-spot method ($\kappa = 0.559$). For all observers, sensitivity was higher with the histogram method (90%) than with the hot-spot method (55%–76%).

Conclusion: Glioma grading based on histogram analysis of normalized CBV heterogeneity is an alternative to the established hot-spot method, as it offers increased diagnostic accuracy and interobserver agreement.

Supplemental material: *http://radiology.rsnajnls.org/cgi/content/full/247/3/808/DC1*

[1] From the Departments of Medical Physics (K.E.E., A.B.), Neuroradiology (B.N., T.N., P.D., J.K.H.), Pathology (D.S., O.C.B.), and Biostatistics (M.C.), and the Intervention Center (K.E.E.), Rikshospitalet-Radiumhospitalet Medical Centre, Sognsvannsveien 20, N-0027 Oslo, Norway; and Department of Physics, University of Oslo, Oslo, Norway (A.B.). Received March 28, 2007; revision requested May 23; revision received July 2; accepted August 1; final version accepted November 13. **Address correspondence to** K.E.E. (e-mail: *kyrre.eeg.emblem@rikshospitalet.no*).

Magnetic resonance (MR) imaging is the imaging method of choice for characterization of brain tumors prior to treatment. Although conventional contrast material–enhanced MR imaging may indicate the degree of tumor malignancy, studies have shown that the degree of contrast enhancement is not a reliable indicator of the tumor grade (1,2). Consequently, authors have suggested that contrast-enhanced dynamic perfusion imaging can improve the accuracy of MR-based glioma grading (3). Perfusion MR imaging involves the use of first-pass bolus-tracking analysis to derive relative cerebral blood volume (CBV) maps, and studies have shown that the maximal relative CBV of gliomas correlates with the glioma grade (4–7).

Differentiation of high-grade gliomas (HGGs) and low-grade gliomas (LGGs) with MR-derived relative CBV maps is based on measurement of the ratio between the most elevated relative CBV area within the glioma (ie, the hot-spot method) and the relative CBV of unaffected tissue. This value is often referred to as the normalized CBV, and HGGs tend to have higher normalized CBVs than do LGGs (7). It should be noted, however, that this approach has some inherent limitations. First, the selection of a glioma hot spot is highly user dependent because differentiation between vessels and the tumor region of true blood volume elevation can be challenging and a source of error. Second, since only a few image pixels are typically used to determine the relative CBV hot spot, the method is inherently sensitive to image noise and other sources of spurious pixel values (eg, spikes introduced by the algorithms used to generate the normalized CBV maps). Third, unaffected white matter relative CBV is generally used to derive the normalized CBV. This is based on the assumption that most gliomas are located in the white matter. However, incorrect selection of reference relative CBVs might result in either under- or overestimation of normalized CBVs. Fourth, oligodendrogliomas tend to have high normalized CBVs regardless of the glioma grade (1). As a result, cutoff normalized CBVs between HGG and LGG might be harder to establish if oligodendrogliomas are included.

In view of these facts, the purpose of our study was to retrospectively compare the diagnostic accuracy of an alternative method used to grade gliomas that is based on histogram analysis of normalized CBVs from the entire tumor volume (obtained with the histogram method) with that of the hot-spot method, with histologic analysis as the reference standard.

Advances in Knowledge

- Use of histogram analysis, compared with use of the current hot-spot glioma grading method, can increase diagnostic accuracy when grading gliomas.
- Use of histogram analysis, compared with use of the current hot-spot glioma grading method, can increase interobserver reproducibility when grading gliomas.

Implication for Patient Care

- The improved diagnostic accuracy and interobserver reproducibility of the histogram analysis method could potentially improve the care of patients with gliomas.

Materials and Methods

Patient Selection

One author (A.B.) is a consultant for Nordic Imaging Lab (Bergen, Norway). Authors without a financial interest controlled data and information that could have caused a conflict of interest. The regional medical ethics committee approved this study, and patients were included only if they provided written informed consent. Between June 2005 and March 2007, primary glioma was diagnosed at histologic analysis in 75 patients after MR perfusion imaging and subsequent surgery. Fifty-three of these patients (24 female, 29 male; mean age, 48 years; age range, 14–76 years) allowed us to use their data in our study. Two experienced neuropathologists (D.S., O.C.B.) performed histologic evaluation based on examination of tissue obtained via resection ($n = 42$) or stereotactic image-guided biopsy ($n = 11$), with use of the World Health Organization classification of central nervous system tumors (8).

Observers

Four experienced neuroradiologists (B.N., T.N., P.D., J.K.H.) with 4–5 years of experience with brain perfusion MR imaging independently performed all measurements. Patient-related information was removed from all images, and observers were blinded to the histopathologic diagnosis.

MR Imaging and Postprocessing

Imaging had been performed at 1.5 T (Sonata, Symphony, or Avanto; Siemens, Erlangen, Germany) with an eight-channel (Sonata or Symphony imagers) or 12-channel (Avanto imager) head coil. The imaging protocol included an axial T2-weighted fast spin-echo sequence (repetition time msec/echo time msec, 4000/104) and an axial T1-weighted spin-echo sequence (500/7.7) performed before and after intravenous administration of gadobutrol (Gadovist; Schering, Berlin, Germany). The voxel size was $0.45 \times 0.45 \times 5$ mm, with 19 sections acquired with both sequences.

Published online
10.1148/radiol.2473070571

Radiology 2008; 247:808–817

Abbreviations:
A_z = area under the receiver operating characteristic curve
CBV = cerebral blood volume
HGG = high-grade glioma
LGG = low-grade glioma

Author contributions:
Guarantor of integrity of entire study, K.E.E.; study concepts/study design or data acquisition or data analysis/interpretation, all authors; manuscript drafting or manuscript revision for important intellectual content, all authors; manuscript final version approval, all authors; literature research, K.E.E., B.N., T.N., P.D., J.K.H., D.S., O.C.B., A.B.; clinical studies, K.E.E., B.N., T.N., P.D., J.K.H., D.S., O.C.B., A.B.; statistical analysis, K.E.E., M.C., A.B.; and manuscript editing, all authors

See Materials and Methods for pertinent disclosures.

Dynamic contrast-enhanced perfusion MR imaging was performed with gradient-echo echo-planar imaging during contrast agent administration. The imaging parameters were as follows: 1430/46 and 1345 Hz/pixel bandwidth for acquisition of 12 axial sections and 1720/48 and 1500 Hz/pixel bandwidth for acquisition of 14 axial sections. We also used a 230 × 230-mm field of view, 1.80 × 1.80 × 5-mm voxel size, and 1.5-mm intersection gap in these examinations. For each section, 50 images were obtained at intervals equal to the repetition time. After eight to 10 time points, 0.2 mmol of gadobutrol per kilogram of body weight was injected at a rate of 5 mL/sec and immediately followed by a 20-mL bolus of saline (B. Braun Melsungen, Melsungen, Germany) injected at a rate of 5 mL/sec.

The images were postprocessed with a dedicated software package (Nordic ICE; Nordic Imaging Lab). The relative CBV (measured in milliliters per 100 g) maps were generated by using established tracer kinetic models applied to the first-pass data (9,10). To reduce the effects of recirculation, the ΔR2* (change in 1/T2*) curves were fitted to a gamma-variate function, which is an approximation of the first-pass response as it would appear in the absence of recirculation or leakage. Although potentially more rigorous correction methods exist (11), the gamma-variate approach was used to conform to the reference method described later in this article. Normalized CBV maps were calculated on a pixel-by-pixel basis by dividing every relative CBV value by a contralateral unaffected white matter relative CBV value defined by a neuroradiologist (B.N.) (12). The normalized CBV maps were displayed as color overlays on the structural images. Coregistration between the conventional MR images and the normalized CBV maps was performed on the basis of geometric information stored in the respective data sets (13).

Image Analysis

The four observers performed image analysis independently over a 3-month period. Two observers (B.N., T.N.) reviewed conventional MR findings for each patient. As described in previous studies (11,14), regions of interest that contained the complete tumor were drawn in each section according to the combined overlay and underlay information, with care taken to avoid areas of necrosis, cysts, or nontumor macrovessels evident on the postcontrast T1-weighted images (Fig 1). High-signal-intensity areas thought to represent tumor tissue on the T2-weighted images were used to define the outermost tumor margin, and areas of contrast enhancement seen on the postcontrast T1-weighted images were always included.

Figure 1

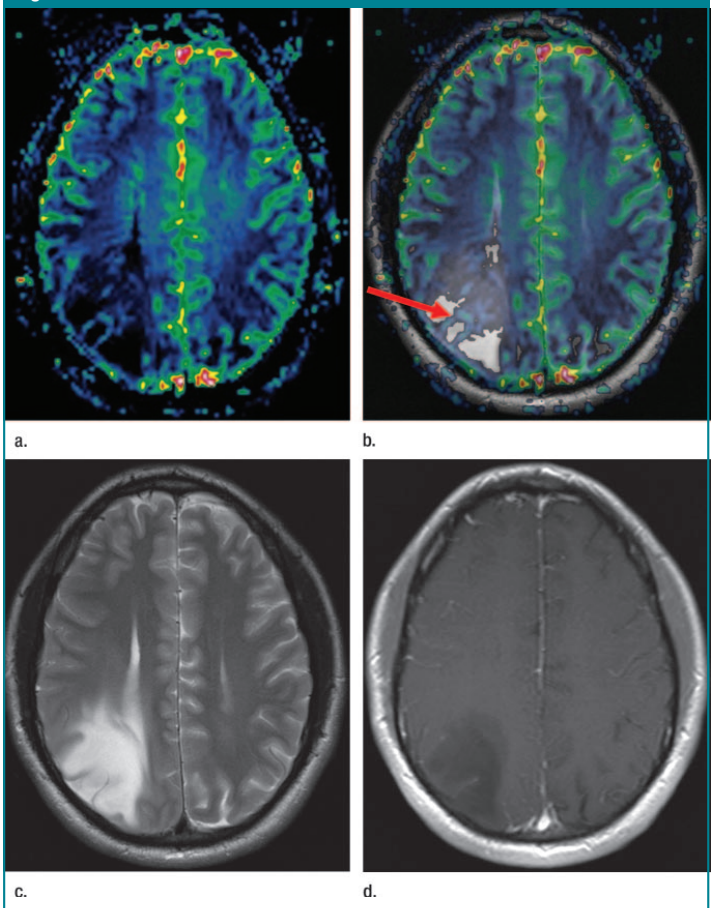


Figure 1: MR images of grade II diffuse astrocytoma in patient 44 (Table E1, *http://radiology.rsnajnls.org/cgi/content/full/247/3/808/DC1*) show how normalized CBV overlay maps are used to identify vessels within the tumor region. **(a)** Axial normalized CBV map. **(b)** Coregistered normalized CBV map overlaid on an axial T2-weighted fast spin-echo image (4000/104). **(c)** Axial T2-weighted fast spin-echo image (4000/104). **(d)** Axial postcontrast T1-weighted spin-echo image (500/7.7). In **b**, the arrow indicates a potential hot-spot area, as seen on the normalized CBV map. However, the underlying vessel-like structure identified in both **c** and **d** might indicate that this is not a hot spot.

The observers recorded the time needed to perform the analysis and evaluated how difficult the methods were to perform (easy, intermediate, or difficult).

Histogram analysis was performed as follows: Histograms were generated by classifying the normalized CBVs in each region of interest into a predefined number of bins (one to 1000 bins). The area under the resulting histogram curve was normalized to the value of one. The range of normalized CBVs along the x-axis was kept constant (between zero and 20). Glioma malignancy was assessed by measuring the maximum normalized peak height of distribution (ie, the relative frequency of normalized CBVs in a given histogram bin), with the hypothesis that normalized CBV heterogeneity is related to tumor malignancy and is inversely proportional to the peak height of the normalized CBV distribution. In the reference-standard hot-spot method (12), which was shown to have the highest intra- and interobserver reproducibility among a number of reported hot-spot methods, each observer selected a minimum of four regions of interest that were believed to represent high normalized CBV regions, and the maximum value was used. The size of the tumor regions of interest remained constant (radius, 1.8 mm). In the case of multiple lesions, the largest lesion was

Figure 2

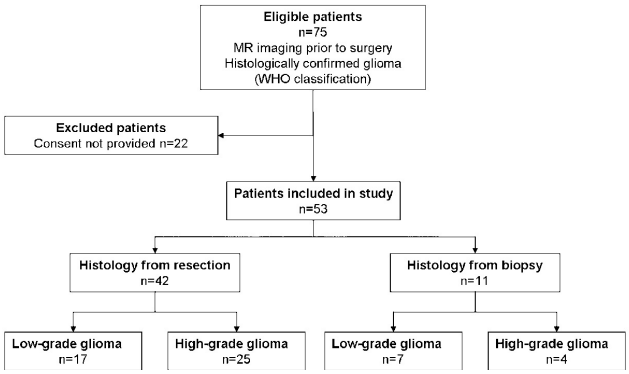


Figure 2: Flowchart shows 75 eligible patients received a histologic diagnosis of primary glioma after MR perfusion imaging and subsequent surgery over a 21-month period (June 2005 to March 2007). Only patients who agreed to participate in the study were included in the analysis.

Table 1

Normalized CBV Values and Optimal Cutoffs between LGG and HGG for Hot-Spot and Histogram Methods

Technique	Observer 1	Observer 2	Observer 3	Observer 4
Hot-spot method				
Mean normalized CBV values for LGG (mL/100 g)	3.48 ± 2.29	3.50 ± 2.32	4.29 ± 2.18	3.89 ± 3.95
Mean normalized CBV values for HGG (mL/100 g)	6.58 ± 3.31	6.14 ± 4.67	8.46 ± 4.07	8.30 ± 4.33
Optimal cutoff between LGG and HGG (mL/100 g)*	4.36	3.75	5.58	4.96
Histogram method				
Mean LGG histogram peak height (mL/100 g)	0.15 ± 0.08	0.16 ± 0.07	0.16 ± 0.08	0.17 ± 0.08
Mean HGG histogram peak height (mL/100 g)	0.07 ± 0.03	0.08 ± 0.03	0.08 ± 0.03	0.08 ± 0.05
Optimal cutoff between LGG and HGG (mL/100 g)*	0.10	0.11	0.11	0.12
Differentiation between grade III and grade IV gliomas†				
Hot-spot method	.060	.271	.551	.927
Histogram method	.002	.035	.090	.060
Differentiation between grade II oligodendroglial tumors and grade II astrocytomas†				
Hot-spot method	.036	.346	.247	.128
Histogram method	.311	.247	.277	.128
Differentiation between LGG and HGG†				
Hot-spot method	>.001	.014	>.001	>.001
Hot-spot method with grade II oligodendroglial tumors excluded	.012	.269	.007	.002
Histogram method	>.001	>.001	>.001	>.001
Histogram method with grade II oligodendroglial tumors excluded	>.001	>.001	>.001	>.001

Note.—Unless otherwise indicated, data are mean CBV values ± standard deviations. Histogram analysis was performed with 108 histogram bins.

* Data are cutoff points and were obtained by minimizing the number of glioma grade misclassifications and maximizing the average sensitivity and specificity.

† Data are *P* values and were calculated with the Mann-Whitney test.

Figure 3

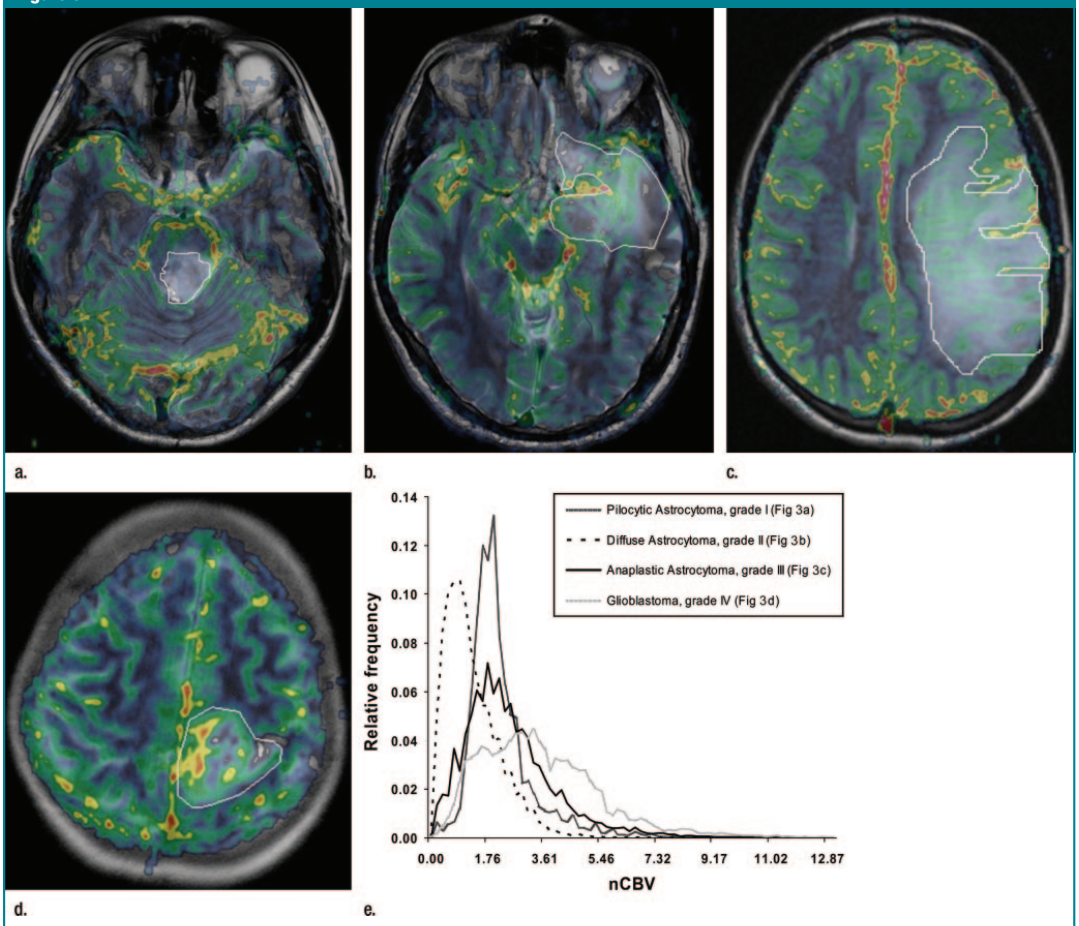


Figure 3: **(a–d)** Normalized CBV maps overlaid on axial T2-weighted fast spin-echo MR images (4000/104) in patients with **(a)** grade I pilocytic astrocytoma, **(b)** grade II diffuse astrocytoma, **(c)** grade III anaplastic astrocytoma, and **(d)** grade IV glioblastoma. The patients are subjects 28, 2, 18 and 7, respectively (Table E1, *http://radiology.rsnajnls.org/cgi/content/full/247/3/808/DC1*). Note the middle cerebral artery encased by the tumor volume in **b**. **(e)** The corresponding histogram signatures derived from the total tumor volume of these patients is shown. The histogram signatures for each patient were derived by using the mean normalized CBVs *(nCBV)* obtained by all four observers. Note the low maximum peak height and wide distribution in **c** and **d** compared with that in **a** and **b**.

chosen. For the hot-spot method, mean normalized CBV values and standard deviations were recorded for LGG and HGG. For the histogram method, mean histogram peak heights and standard deviations were recorded.

Statistical Analysis

The diagnostic accuracy of the two methods was evaluated by using binary logistic regression to derive sensitivity, specificity, and positive and negative predictive values for LGG versus HGG. A glioma classified as an HGG or an LGG with both observer data and histologic analysis was considered a true-positive finding or a true-negative finding, respectively. As described previously, optimal cutoff values between LGG and HGG for each observer were obtained by minimizing the number of glioma grade misclassifications and maximizing the average sensitivity and specificity (7,15,19). To compare our results with the results of previous studies, the diagnostic accuracy of the hot-spot method was also calculated by using a previously published cutoff normalized CBV of 1.75 (15). The sensitivity and specificity for each observer were compared by using the McNemar test and a

pairwise comparison of the area under the receiver operating characteristic curve (A_z). The number of histogram bins that yielded the highest A_z was derived by using an in-house developed Matlab routine (R2006a; MathWorks, Natick, Mass) that was used to calculate the A_z values for all bin numbers between one and 1000.

Mann-Whitney tests were used to assess the ability to differentiate *(a)* between grade II oligodendroglial tumors (oligodendrogliomas or oligoastrocytomas) and grade II diffuse astrocytomas and *(b)* between grade III gliomas and grade IV gliomas with the two methods. Mann-Whitney tests were also used to assess whether excluding grade II oligodendroglial tumors affected glioma grading. To account for multiple-comparison effects, a significance level of $P = .01$ was used. This value was obtained by dividing a default P value of .05 by the number of observers and applying the Bonferroni correction.

Interobserver reproducibility between the four observers was assessed by using Fleiss κ statistics based on whether the observers graded a glioma as HGG or LGG. A κ value of less than zero indicated poor agreement; a κ value of 0.00–0.20, slight agreement; a κ value of 0.21–0.40, fair agreement; a κ value of 0.41–0.60, moderate agreement; a κ value of 0.61–0.80, substantial agreement; and a κ value of 0.81–1.00, almost perfect agreement (16). Statistical analysis was performed by using SPSS13 (SPSS, Chicago, Ill).

Table 2

Diagnostic Accuracy of Hot-Spot and Histogram Methods

Observer and Method	Sensitivity (%)	Specificity (%)	Positive Predictive Value (%)	Negative Predictive Value (%)
Observer 1				
Hot-spot method	76 (56, 89)	83 (62, 89)	85 (65, 95)	74 (53, 88)
Hot-spot method with 1.75 mL/100 g cutoff	97 (82, 99)	21 (7, 42)	60 (44, 73)	83 (35, 99)
Histogram method with 108 bins	90 (72, 97)	83 (62, 95)	87 (69, 96)	87 (66, 97)
Observer 2				
Hot-spot method	55 (35, 73)	63 (40, 81)	64 (42, 82)	54 (33, 72)
Hot-spot method with 1.75 mL/100 g cutoff	100 (88, 100)	29 (12, 51)	63 (47, 76)	100 (59, 100)
Histogram method with 108 bins	90 (72, 97)	83 (62, 95)	87 (69, 96)	87 (66, 97)
Observer 3				
Hot-spot method	72 (52, 87)	88 (67, 97)	88 (67, 97)	72 (52, 87)
Hot-spot method with 1.75 mL/100 g cutoff	97 (82, 99)	8 (1, 27)	56 (41, 70)	67 (9, 99)
Histogram method with 108 bins	90 (72, 97)	83 (62, 95)	87 (69, 96)	87 (66, 97)
Observer 4				
Hot-spot method	76 (56, 89)	83 (62, 95)	85 (65, 95)	74 (53, 88)
Hot-spot method with 1.75 mL/100 g cutoff	97 (82, 99)	21 (7, 42)	60 (44, 73)	83 (35, 99)
Histogram method with 108 bins	90 (72, 97)	83 (62, 95)	87 (69, 96)	87 (66, 97)

Note.—Data in parentheses are 95% confidence intervals.

Figure 4

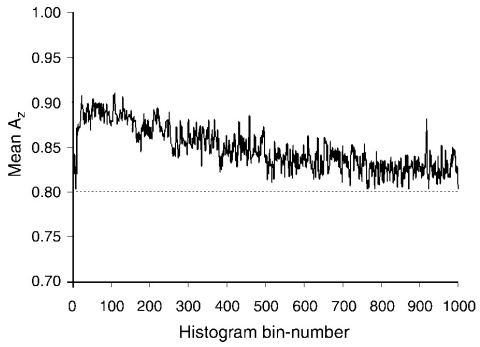


Figure 4: Curve shows the mean A_z for all four observers. The histogram method yielded larger mean A_z values than did the hot-spot method (0.801 ± 0.063, straight line), regardless of the bin number. Averaged over the four observers, the maximum A_z value was found at 108 histogram bins (0.909 ± 0.004).

Results

Of the 53 gliomas investigated, 24 were histologically confirmed to be LGGs (World Health Organization grade I or II) and 29 were histologically confirmed to be HGGs (World Health Organization grade III or IV) (Fig 2) (Table E1, *http://radiology.rsnajnls.org/cgi/content/full/247/3/808/DC1*). Signs of necrosis were seen on conventional MR images in six patients with LGGs and 17 with HGGs (one with grade III glioma and 16 with grade IV glioma). On average, the four observers reported examination times of 7 minutes per patient when using the hot-spot method and 11 minutes per patient when using the histogram method. All observers reported that the two methods were equally difficult to perform (intermediate difficulty).

Glioma Grading

Optimal normalized CBV cutoff values between LGG and HGG ranged from 3.75 to 5.58 mL/100 g (Table 1). Optimal histogram peak values between LGG and HGG ranged from 0.10 to 0.12 mL/100 g. One observer did not ob-

serve a significant difference between LGG and HGG with the hot-spot method. Excluding 13 grade II oligodendroglial tumors from glioma grading led to reduced *P* values for all observers when they used the hot-spot method, whereas *P* values obtained with the histogram method remained unchanged. One observer (observer 1) was able to differentiate ($P < .002$) between grade III ($n = 10$) and grade IV ($n = 19$) gliomas with the histogram method (Fig 3). Neither method enabled us to differentiate between grade II oligodendroglial tumors ($n = 13$) and grade II diffuse astrocytomas ($n = 8$).

For the histogram method, diagnostic accuracy was observer independent (Table 2). For all observers, sensitivity (90%, 26 of 29 HGG patients) and negative predictive value (87%, 20 of 23 patients) increased with use of the histogram method compared with the sensitivity (55%–76%, 16–22 of 29 patients) and the negative predictive value (54%–74%, 15 of 28 patients to 20 of 27 patients) obtained with the hot-spot method. For observer 2 (Table 2), sensitivity values obtained with the histogram method were significantly different from those obtained with the hot-spot method (McNemar test, $P = .002$). Specificity was 83% (20 of 24 patients with LGG) with the histogram method and 63%–88% (13–20 of 24 patients with LGG) with the hot-spot method. The McNemar tests did not reveal a significant difference between specificity values obtained with the histogram method and those obtained with the hot-spot method. The positive predictive value was 87% (26 of 30 patients) with the histogram method and 64%–88% (16 of 25 patients to 22 of 26 patients) with the hot-spot method. Compared with the optimal hot-spot cutoff value, the 1.75 mL/100 g cutoff value resulted in improved sensitivity (97%–100%, 28–29 of 29 patients with HGG) in all observers and improved negative predictive value in three of four observers (67%–100%, between two and seven of seven patients). Both the specificity (8%–29%, between two and seven of 24 patients with LGG) and the positive predictive value (56%–63%, between 28 of 50 and 29 of 46 patients) were reduced in all observers.

Figure 5

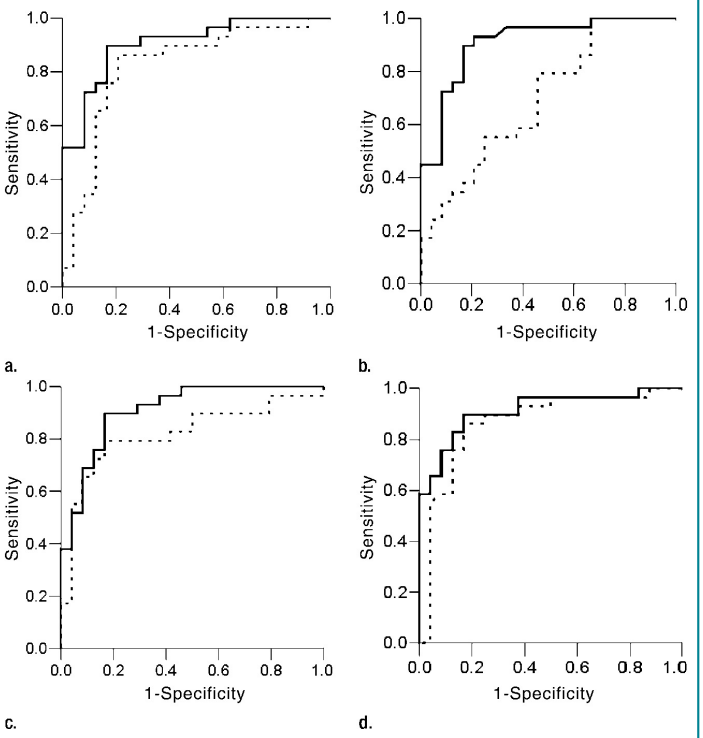


Figure 5: Receiver operating characteristic curves for the histogram method obtained by using 108 histogram bins (solid line) and the hot-spot method (dotted line) for observers **(a)** 1, **(b)** 2, **(c)** 3, and **(d)** 4. For all observers, the mean A_z values obtained with the histogram method were larger than those obtained with the hot-spot method. In **a**, the mean A_z obtained with the histogram method was 0.905 ± 0.041, whereas the mean A_z obtained with the hot-spot method was 0.822 ± 0.062. In **b**, the mean A_z obtained with the histogram method was 0.909 ± 0.041, whereas the mean A_z obtained with the hot-spot method was 0.698 ± 0.072. In **c**, the mean A_z obtained with the histogram method was 0.909 ± 0.040, whereas the mean A_z obtained with the hot-spot method was 0.818 ± 0.061. In **d**, the mean A_z obtained with the histogram method was 0.914 ± 0.039, whereas the mean A_z obtained with the hot-spot method was 0.867 ± 0.055.

The mean A_z values (± standard errors) were larger for all observers when they used the histogram method (range, 0.905 ± 0.041 to 0.914 ± 0.039) than when they used the hot-spot method (range, 0.698 ± 0.072 to 0.867 ± 0.055). For one observer, A_z was significantly higher when the histogram method was used than when the hot-spot method was used ($P < .001$). Averaged over the four observers, the maximum A_z was found at 108 histogram bins (mean $A_z = 0.909 \pm 0.004$) (Figs 4, 5).

Interobserver Reproducibility

For the hot-spot method, there was moderate interobserver agreement between the four observers with use of the optimal cutoff value ($\kappa = .559$) and the 1.75 mL/100 g cutoff value ($\kappa = .459$). Although the size and shape of the resulting tumor regions of interest varied between the four observers (Fig 6), the interobserver agreement between the four observers was almost perfect ($\kappa =$

.923) when the histogram method was used.

Discussion

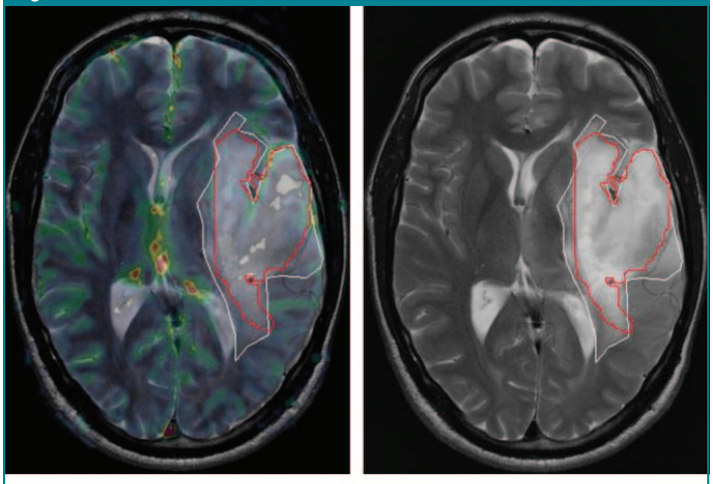
In our study, we evaluated an alternative method with which to differentiate HGG from LGG on the basis of the normalized CBV heterogeneity of the entire tumor volume. Our results suggest that the histogram method has higher interobserver agreement and yields higher sensitivity and negative predictive values and equal specificity when compared with the hot-spot method. The influence of increased diagnostic accuracy on clinical outcome is difficult to establish and was not investigated in our study. However, high diagnostic accuracy combined with high interobserver reproducibility are critical criteria for any diagnostic test (17). One potential advantage of the histogram method is that the results are independent of the choice of reference tissue, as long as the reference is kept the same throughout the cohort. For example, the effect of changing reference tissue from white to gray matter is simply a shift in the position of the peak distribution bin; the actual peak value does not change. However, the hot-spot method is critically dependent on correct selection of reference tissue since determination of normalized CBV is based solely on this parameter. Arguably, the objective of an optimal grading method should be identification of the most malignant part of the tumor, which should favor the hot-spot method. However, our results suggest that observers are not able to consistently identify the most malignant tumor region with current hot-spot methods and that there is a consequent loss in sensitivity or specificity depending on the cutoff value used.

The results we obtained with the hot-spot method are consistent with previously published data (6,15,18,19). However, the optimal cutoff values between HGG and LGG in our study (3.75–5.58 mL/100 g) were higher than those in previous studies (1.5–1.98 mL/100 g) (7,15,19). In a study in which the hot-spot method was used in 160 patients, authors reported a sensitivity of 95.0% and a specificity of 57.5% when they used a CBV cutoff of 1.75 mL/100 g (15). When we applied a normalized CBV cutoff of 1.75 mL/100 g to our data, we obtained similar values but with wide confidence intervals. Hence, our results suggest that to obtain maximum diagnostic accuracy, the choice of an optimal cutoff value should be based on perfusion data generated at a given site. One further observation in the current work was the trade-off between high sensitivity and high specificity with both methods investigated. Lowering the cutoff values increased sensitivity at the cost of reduced specificity. It could be argued that a low false-negative rate is more important than a low false-positive rate because of the serious consequences of false-negative findings. However, both types of errors are potentially critical, given the different treatment strategies for LGG and HGG (20).

The only difference between grade III gliomas and grade IV gliomas was seen by one observer who used the histogram method. In previous studies in which the hot-spot method was used in 26 (6) and 120 (15) patients with HGGs, no difference between these groups was reported. Furthermore, as described previously (21), necrosis was a specific marker for distinguishing grade III gliomas from grade IV gliomas, but it was not a sensitive one.

We were unable to differentiate between grade II astrocytomas and grade II oligodendroglial tumors with either method. Excluding grade II oligodendroglial tumors affected glioma grading only when the hot-spot method was used; this finding suggests that the diagnostic accuracy of the hot-spot method is more dependent on the number of patients than is the histogram method. When we compared the use of different histogram bin numbers, the maximum A_z was found at 108 bins. The reduced diagnostic accuracy at lower bin num-

Figure 6



a. b.

Figure 6: MR images obtained in a patient with grade II oligodendroglioma (patient 9) (Table E1, *http://radiology.rsnajnls.org/cgi/content/full/247/3/808/DC1*) show the manual glioma volume delineation for two observers (red and white regions of interest). **(a)** Coregistered relative CBV map overlaid on a T2-weighted fast spin-echo image (4000/104). **(b)** T2-weighted fast spin-echo image (4000/104). Although the variation between observers is evident, resulting histograms correctly depicted LGG in both cases. Almost perfect interobserver agreement ($\kappa = .923$) obtained with the histogram method suggests that the variations between observers caused by imperfect tumor delineation are relatively unimportant, given the large number of data points included in the histogram.

bers can be explained by the large range of relative CBVs contained in each of the resulting bins, which tends to mask small hypervascular regions in patients with HGGs. Also, the reduced diagnostic accuracy at higher bin numbers may be explained by increasing noise in the resulting histogram, since each bin contains fewer pixel averages.

In our study, we used gradient-echo echo-planar imaging rather than spin-echo echo-planar imaging because a higher temporal resolution can be achieved with this sequence. Also, previous studies have shown a stronger correlation between tumor grade and CBV with use of gradient-echo techniques (14,22,23). Gradient-echo echo-planar imaging sequences have also been shown to be more sensitive to macrovascular structures, aiding in the differentiation between infiltrating vessels and true tumor CBV elevation (24).

Our study had limitations. It would have been preferable to include more patients to strengthen the statistical power. However, the number of patients included in our study ($n = 53$) is similar to that in other studies on glioma grading (3,5–7,12–14,18,19,23). Furthermore, only one observer obtained a significant difference between the sensitivity values of the two methods. However, the mean A_z values were higher for all observers when they used the histogram method compared with when they used the hot-spot method. Therefore, our data suggest that the histogram method has higher diagnostic accuracy. A further limitation of both methods is the need for an optimal coregistration between normalized CBV maps and conventional MR images. Hence, the increasing availability of intramodal image coregistration methods in clinical image software will be of benefit for the clinical utility of both methods. An obvious challenge with the histogram method is that of identifying the appropriate tumor region. Optimal operational definition of tumor volume is complicated because gliomas are infiltrating tumors with indistinct borders beyond the radiologic margins (25,26). However, the high interobserver agreement of the histogram method suggests that variations between observers caused by imperfect tumor delineation are relatively unimportant, given the large number of data points included in the histogram. The observers reported that they spent more time per patient for the histogram method than for the hot-spot method because the tumor volume had to be identified in every section. However, with both methods, the observer had to exclude vessels that infiltrated the glioma region, and, consequently, the methods were considered equally difficult to perform. On the basis of these observations, there was a clear need for more user-independent and automated methods with which to identify total tumor volume and regions representing unaffected reference tissue. Studies have shown that cluster analysis techniques can be used to quantify and classify similar tissue components on MR images (27), and such methods are currently being implemented as part of our histogram analysis software.

Histogram analysis was performed by assessing the peak height of the normalized histogram distribution of normalized CBVs in the tumor. This approach was chosen because the resulting height was directly determined on the basis of the underlying heterogeneity of the normalized CBV distribution. It is hypothesized that histogram-based analysis can be further improved with parametric analysis of the histogram shape rather than just the peak value.

In conclusion, our results suggest that the proposed histogram method is a diagnostically accurate and reproducible method with which to grade gliomas on the basis of MR-derived blood volume maps. Compared with the hot-spot method, the histogram method had higher interobserver agreement, sensitivity, and negative predictive value and equal specificity. Future developments in cluster methods for automated segmentation of tumor volume may further enhance the clinical utility of this method.

References

1. Lev MH, Rosen BR. Clinical applications of intracranial perfusion MR imaging. Neuroimaging Clin N Am 1999;9(2):309–331.
2. Law M, Oh S, Babb JS, et al. Low-grade gliomas: dynamic susceptibility-weighted contrast-enhanced perfusion MR imaging—prediction of patient clinical response. Radiology 2006;238(2):658–667.
3. Covarrubias DJ, Rosen BR, Lev MH. Dynamic magnetic resonance perfusion imaging of brain tumors. Oncologist 2004;9(5):528–537.
4. Edelman RR, Mattle HP, Atkinson DJ, et al. Cerebral blood flow: assessment with dynamic contrast-enhanced T2*-weighted MR imaging at 1.5 T. Radiology 1990;176(1):211–220.
5. Aronen HJ, Gazit IE, Louis DN, et al. Cerebral blood volume maps of gliomas: comparison with tumor grade and histologic findings. Radiology 1994;191(1):41–51.
6. Knopp EA, Cha S, Johnson G, et al. Glial neoplasms: dynamic contrast-enhanced T2*-weighted MR imaging. Radiology 1999;211(3):791–798.
7. Lev MH, Ozsunar Y, Henson JW, et al. Glial tumor grading and outcome prediction using dynamic spin-echo MR susceptibility mapping compared with conventional contrast-enhanced MR: confounding effect of elevated rCBV of oligodendrogliomas. AJNR Am J Neuroradiol 2004;25(2):214–221.
8. Kleihues P, Cavenee WK. Astrocytic tumors and oligodendroglial tumors and mixed gliomas. In: Kleihues P, Cavenee WK, eds. The WHO classification of tumors of the nervous system. Lyon, France: International Agency for Research on Cancer, 2000; 9–70.
9. Rosen BR, Belliveau JW, Vevea JM, Brady TJ. Perfusion imaging with NMR contrast agents. Magn Reson Med 1990;14(2):249–265.
10. Ostergaard L, Weisskoff RM, Chesler DA, Gyldensted C, Rosen BR. High resolution measurement of cerebral blood flow using intravascular tracer bolus passages. I. Mathematical approach and statistical analysis. Magn Reson Med 1996;36(5):715–725.
11. Boxerman JL, Schmainda KM, Weisskoff RM. Relative cerebral blood volume maps corrected for contrast agent extravasation significantly correlate with glioma tumor grade, whereas uncorrected maps do not. AJNR Am J Neuroradiol 2006;27(4):859–867.
12. Wetzel SG, Cha S, Johnson G, et al. Relative cerebral blood volume measurements in intracranial mass lesions: interobserver and intraobserver reproducibility study. Radiology 2002;224(3):797–803.
13. Bjornerud A. The ICE software package: direct co-registration of anatomical and func-

tional datasets using DICOM image geometry information. Proc Hum Brain Mapping 2003;19(2):1018p.

14. Schmainda KM, Rand SD, Joseph AM, et al. Characterization of a first-pass gradient-echo spin-echo method to predict brain tumor grade and angiogenesis. AJNR Am J Neuroradiol 2004;25(9):1524–1532.

15. Law M, Yang S, Wang H, et al. Glioma grading: sensitivity, specificity, and predictive values of perfusion MR imaging and proton MR spectroscopic imaging compared with conventional MR imaging. AJNR Am J Neuroradiol 2003;24(10):1989–1998.

16. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33(1):159–174.

17. Fryback DG, Thornbury JR. The efficacy of diagnostic imaging. Med Decis Making 1991; 11(2):88–94.

18. Sugahara T, Korogi Y, Kochi M, et al. Correlation of MR imaging-determined cerebral blood volume maps with histologic and angiographic determination of vascularity of gliomas. AJR Am J Roentgenol 1998;171(6): 1479–1486.

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

20. Claus EB, Black PM. Survival rates and patterns of care for patients diagnosed with supratentorial low-grade gliomas: data from the SEER program, 1973–2001. Cancer 2006;106(6):1358–1363.

21. Nelson JS, Tsukada Y, Schoenfeld D, Fulling K, Lamarche J, Peress N. Necrosis as a prognostic criterion in malignant supratentorial, astrocytic gliomas. Cancer 1983;52(3):550–554.

22. Donahue KM, Krouwer HG, Rand SD, et al. Utility of simultaneously acquired gradient-echo and spin-echo cerebral blood volume and morphology maps in brain tumor patients. Magn Reson Med 2000;43(6):845–853.

23. Sugahara T, Korogi Y, Kochi M, Ushio Y, Takahashi M. Perfusion-sensitive MR imaging of gliomas: comparison between gradient-echo and spin-echo echo-planar imaging techniques. AJNR Am J Neuroradiol 2001; 22(7):1306–1315.

24. Aronen HJ, Perkio J. Dynamic susceptibility contrast MRI of gliomas. Neuroimaging Clin N Am 2002;12(4):501–523.

25. Price SJ, Jena R, Burnet NG, et al. Improved delineation of glioma margins and regions of infiltration with the use of diffusion tensor imaging: an image-guided biopsy study. AJNR Am J Neuroradiol 2006;27(9):1969–1974.

26. Grier JT, Batchelor T. Low-grade gliomas in adults. Oncologist 2006;11(6):681–693.

27. Itskovich VV, Samber DD, Mani V, et al. Quantification of human atherosclerotic plaques using spatially enhanced cluster analysis of multicontrast-weighted magnetic resonance images. Magn Reson Med 2004; 52(3):515–523.

II

Automated Glioma Characterization from Dynamic Susceptibility Contrast Imaging. Brain Tumor Segmentation using Knowledge-based Fuzzy Clustering

Kyrre E. Emblem[1,3], M.Sc, Baard Nedregaard[2], MD, John K. Hald[2], MD, PhD, Terje Nome[2], MD, Paulina Due-Tonnessen[2], MD, Atle Bjornerud[1,4], PhD

From the Departments of Medical Physics[1], Clinic for Imaging and Intervention[2] and The Interventional Center[3], Rikshospitalet University Hospital, Sognsvannsveien 20, N-0027 Oslo, Norway. Department of Physics[4], University of Oslo, Sem Saelandsvei 24, N-0371 Oslo, Norway.

Corresponding author:

Kyrre E. Emblem

The Interventional Center, Gaustad

Rikshospitalet University Hospital

Sognsvannsveien 20

N-0027 Oslo

Norway

Telephone number: +47 23070145

Fax number: +47 23070110

E-mail address: kyrre.eeg.emblem@rikshospitalet.no

Presented in part at the International Society for Magnetic Resonance in Medicine, 1458p, 2007; May 19-25, Berlin, Germany

ABSTRACT:

Purpose: Current methods for pre-surgical MR-based glioma characterization from dynamic susceptibility contrast (DSC) imaging are time consuming and user-dependent. We have assessed whether glioma volumes from knowledge-based fuzzy c-means (FCM) clustering of multiple MR image classes can provide similar diagnostic efficacy values as manually defined tumor volumes when characterizing gliomas from DSC imaging.

Materials and Methods: Fifty patients with newly diagnosed gliomas were imaged using DSC MR imaging at 1.5 Tesla. To compare our results with manual tumor definitions, glioma volumes were also defined independently by four neuroradiologists. Using a histogram analysis method, diagnostic efficacy values for glioma grade and expected patient survival were assessed.

Results: The areas under the receiver operator characteristics (ROC) curves were similar when using manual and automated tumor volumes to grade gliomas (P=.576-.970). When identifying a high-risk patient group (expected survival <2 years) and low-risk patient group (expected survival >2 years), a higher log-rank value from Kaplan-Meier survival analysis was observed when using automatic tumor volumes (14.984, P<.001) compared to the manual volumes (9.441-12,022, P=.001-.002).

Conclusion: Our results suggest that knowledge-based FCM clustering of multiple MR image classes holds promise as a user-independent approach to selecting the target region for pre-surgical glioma characterization.

Key words: DSC MR imaging, tumor segmentation, pre-surgical glioma grading, histogram method, fuzzy clustering.

INTRODUCTION:

Magnetic Resonance (MR) is the imaging modality of choice to characterize brain tumors prior to surgical intervention. The high level of soft tissue contrast makes MR images an important aid in the identification and treatment of gliomas(1-3). Further, in combination with conventional MR imaging, several studies have shown that dynamic susceptibility contrast (DSC) MR imaging of gliomas can be used as a pre-surgical indicator for glioma grade(4-6) and patient outcome(7;8). Typically, glioma grading from DSC is based on analysis of cerebral blood volume (CBV) values within the glioma area, using either a hot-spot method(4;9) or a histogram analysis method(10;11). Correct identification of glioma tissue however, may be dependent on experienced operators with good anatomical knowledge, making current grading methods inherently operator-dependent and time consuming(6;10;11). Recently, both supervised and unsupervised segmentation methods for identification of brain tissue structures have been proposed. Automatic tissue or tumor segmentation based on multi-spectral data analysis(12;13), neural networking(14;15), support vector machines(16;17) and knowledge-based fussy c-means (FCM) clustering techniques(18;19) all show great promise. The potential advantages of automatic tumor segmentation include removal of inter-observer variations, time efficiency and standardized criteria's for tumor characterization(18). To the author's knowledge, no study has investigated whether automatically segmented glioma volumes can substitute manually selected glioma volumes when characterizing gliomas from DSC imaging pre-surgically. A reason for this might be that traditional brain tumor MR imaging protocols consist of 2 dimensional (2D) anatomical MR images which are sub-optimal for tumor segmentation compared to 3D MR images(20;21). Also, glioma segmentation from MR images is inherently difficult because gliomas are known to be highly infiltrative and might mimic both unaffected brain tissue and cerebrospinal fluid (CSF) depending on image parameters. Finally, an automatic tumor segmentation model must

include tumor areas with contrast enhancement without including contrast enhanced blood vessels. Because glioma appearance on MR images and anatomical localization may differ between glioma grades and patients(1-3), a supervised training of a brain atlas from a sub-sample of the data is dependent on a very large sample size in order to derive a robust template(14;20). An unsupervised segmentation procedure combining FCM clustering of anatomical MR images and knowledge-based operations(18;19;22) however, is not dependent on sample size. Here, FCM differ from a standard k-means clustering technique in that each data element (single pixel) can belong to more than one cluster. The strength of the association between a data element and a class is indicated using a value between zero and one(19). The knowledge-based operation consists of a linear sequence of low-level image processing operations based on known MR image properties secondary to brain structures or pathology. A drawback with knowledge-based FCM is that a perfect delineation between the resulting tissue classes is difficult, especially in heterogeneous tissue. However, it has been shown that imperfect glioma delineation is relatively unimportant in pre-surgical glioma grading by histogram analysis of DSC images(10). Thus, using a histogram analysis method on DSC values from automatically segmented tumor volumes should favor the unsupervised knowledge-based FCM technique(18).

In view of the above, the purpose of our study was to compare the diagnostic efficacy values of DSC MRI, in terms of pre-surgical diagnostic accuracy and expected patient survival, using automatically segmented and manually selected glioma volumes.

MATERIALS AND METHODS:

Patient Selection

Study approval was obtained from the Regional Medical Ethics Committee and patients were included only if informed consent was signed. Between June 2005 and June 2007, fifty previously untreated patients (21 females, 29 males, mean age 49 years, range 14-76 years) were imaged using an MR imaging protocol with relevant sequences for tumor segmentation. All patients received a diagnosis of glioma after MR perfusion imaging and subsequent surgery (resection or biopsy). The histological evaluation was based on the World Health Organization (WHO) classification system(23).

Observers

Manual glioma identification was performed independently by four neuroradiologists with several years of experience with brain MR perfusion imaging. The observers were blinded to patient related information and the histopathological diagnosis.

MR Imaging and Manual Post-Processing

Imaging was performed at 1.5 Tesla (Siemens Sonata, Symphony or Avanto, Siemens AG, Erlangen, Germany), using an 8-channel- (Symphony/Sonata) or a 12-channel (Avanto) head-coil. The protocol included a 19 slice axial T2-weighted (T2-w) fast spin-echo sequence with 4000/104 (repetition time msec /echo time msec), a 25 slice coronal FLAIR sequence (9050/114) with inversion time of 1500ms and a 19 slice axial T1-weighted (T1-w) spin-echo sequence (500/7.7) obtained before and after i.v. contrast agent injection. The voxel size of the axial images was 0.45x0.45x5mm^3 and 0.9x0.9x5 mm^3 for the coronal images. DSC MR imaging was performed using a gradient-echo echo-planar imaging sequence acquired during contrast agent administration. The imaging parameters were: 1430/46, bandwidth 1345 Hz/pixel (12 axial slices) or 1720/48, bandwidth 1500 Hz/pixel (14 axial slices), field of view 230x230mm, voxel size 1.80x1.80x5 mm^3, inter-slice gap 1.5mm. For each slice, 50 images

were recorded at intervals equal to the repetition time. After approximately 8 time-points, 0.2 mmol/kg of gadobutrol (Bayer Schering Pharma AG, Berlin, Germany) was injected at a rate of 5mL/sec, immediately followed by a 20mL bolus of saline (B. Braun Melsungen AG, Melsungen, Germany) also at 5 mL/sec.

The images were post-processed using a dedicated software package (nordicICE, NordicImagingLab, Bergen, Norway) and Matlab R2007a (MathWorks, Natick, US). The CBV maps were generated from the area under the 1/T2* converted first-pass curves(24;25) and corrected for possible extra-vascular contrast agent leakage(26). Relative (r)CBV maps were derived by dividing every CBV value in each slice with an unaffected white matter CBV value in the same slice defined by a neuroradiologist(9) and displayed using a black-blue-green-yellow-red-white colormap. All images were automatically coregistered using a normalized mutual information algorithm(27).

Manual Glioma Definition

As described elsewhere(10;26;28), the four observers determined the glioma area in each slice based on the rCBV overlay / MR image underlay information taking care to avoid areas of necrosis, cysts or non-tumor macro-vessels readily evident on the post-contrast T1-w images and the rCBV images. Although gliomas may extend beyond hyper-intensities as seen on the T2-w images, this parameter was used to define the outmost glioma margins as studies have shown that the metabolically active tumor area is restricted mainly to the T2 tumor border(29;30).

Automatic Glioma Segmentation

Prior to glioma segmentation, the image intensity values were standardized across MR image types, MR image slices and patients using adaptive histogram equalization(31). Also, using

pre-contrast T1-w images only, an intra-cranial brain mask procedure in Statistical Parametric Mapping (SPM5) was applied to isolate brain tissue pixels from the remaining non-brain pixels(32). Here, brain tissue pixels were identified using a standard T1-w image template from SPM5. After applying an affine transformation that best matched the patient MR image with the template image, a maximum likelihood mixture model cluster algorithm(32) was applied to identify each image pixel as either gray matter, white matter or CSF. Using a-priori information, the final binary brain mask was derived from the combination of image pixels most likely to be either gray or white matter. Image pixels most likely to be CSF was not included in the brain mask. In the tumor region, intensity values similar to CSF were thought to be non-tumor edema or cysts, and were subsequently removed from the brain mask.

Glioma volumes were automatically segmented from the T2-w and FLAIR images using a previously published knowledge-based FCM cluster technique shown to correspond well with manually defined tumor volumes(18;19;22). For the T2-w images, a 3-class FCM algorithm was applied on the complete image stack to identify the cluster class representative of the highest mean pixel values thought to represent glioma. The other two cluster classes were thought to represent (a) non-brain areas (i.e. zero-valued pixels as defined by the brain mask) and (b) normal appearing gray and white matter. In order to identify the correct glioma class, Otsu's method(33) was used to reduce the continuously scaled FCM cluster images (with values ranging from 0 to 1) to binary images. Here, the optimal threshold separating classes is found by minimizing the within-class variance.

For the FLAIR images, a 4-class FCM algorithm was applied on the complete image stack to identify the cluster class representative of the highest pixel values thought to represent glioma tissue only. The other three classes were thought to represent (a) non-brain areas, (b) normal appearing gray or white matter and (c) fluids or vessels. Similar to the T2-w image clustering

routine, the Otsu method was applied a second time to derive binary images from the continuously scaled FCM cluster images. The final glioma image was a combination of the binary T2-w and FLAIR glioma cluster classes.

A set of 2D and 3D binary morphological image operations was performed on the final glioma image to remove non-tumor pixels mimicking tumor tissue(18). First, on the 2D binary glioma images, H-connected pixels (i.e. single 1's connecting two areas of 1's) were removed. Second, a pixel was assigned a value of 1 if five or more pixels in its 3-by-3 neighborhood was 1's, otherwise, it was set to zero. Third, the most compact object, thought to represent the glioma only (called "First Tumor" (18)), was separated from remaining non-tumor regions. Here, First Tumor was assessed by deriving circularity values for all remaining objects across all slices:

$$Circularity: \frac{4\pi Area}{Perimeter^2} \qquad [1]$$

From this, a 3D seed growing algorithm was applied to the complete image stack to identify pixels connected to the First Tumor. In the case of multiple lesions, the largest lesion was chosen.

In addition, using a 3-class FCM algorithm after brain masking, areas of contrast enhancement as seen on post-contrast T1-w images were always included. Binary cluster classes were assessed using Otsu's method and presence of contrast enhanced was assumed if; (a) the mean intensity value minus the standard deviation of a cluster class was higher than the mean value plus standard deviations of the other two classes, (b) the variance of a cluster class was 4 times higher than the variance of the other two classes. In addition, as described above, the binary morphological image operations were used to clean the final binary contrast

enhancement mask. A schematic flow-diagram of the entire automatic segmentation procedure is shown in Figure 1.

Glioma Characterization

Histogram signatures were generated for each patient using rCBV values from the complete manual and automatic glioma volumes. In order to conform to the histogram method(10), the range of the rCBV values along the x-axis was kept constant (values; 0-20). The area under each histogram curve was normalized to one and the histograms were divided into 108 bins, a number shown to give the maximal diagnostic accuracy. Glioma malignancy was assessed by measuring the maximum peak height of the normalized histograms. Also, two observers in consensus determined the degree of T1-w contrast enhancement in each glioma as (a) none, (b) moderate or (c) extensive.

Statistical Analysis

To compare how well (on a pixel-by-pixel basis) the automatic glioma volumes corresponded with the manual glioma volumes, sensitivity values and positive predictive values (PPV) were derived for high- and low-grade gliomas separately. Specificity values and negative predictive values (NPV) were not assessed as these parameters were considered redundant as such an analysis would include all non-tumor pixels in the MR image stack. Because of variations between observers, pixels classified as glioma by at least three of four observers were thought to represent *true* glioma tissue. From this, pixels identified as glioma tissue by both manual and automatic tumor segmentation were considered true positives, whereas pixels identified as glioma tissue by manual or automatic tumor segmentation only were considered false negatives and false positives, respectively. The correlation between presence of contrast enhancement as identified by both observers and the FCM cluster algorithm was assessed

using Spearman correlation tests (Yes / No). Sensitivity and specificity values when using manual and automatic glioma volumes to pre-surgically grade gliomas were assessed by pairwise comparisons of the areas (A_z) under the Receiver Operating Characteristic (ROC) curves.

Finally, regardless of glioma grade, Kaplan-Meier survival curves with log-rank tests (Mantel-Cox) were used to compare the manual and automatic tumor volumes with respect to separating a "high-risk" patient group from a "low-risk" group. Here, the high-risk group was defined as those patients with an expected survival from MR examination date to death of less than 2 years, whereas the low-risk group was defined as those patients with an expected survival of more than 2 years. Based on available survival data, optimal histogram peak height cut-off values between the high- and low-risk patient groups were derived using binary logistic regression. Statistical analysis was performed using SPSS 15 (Apache Software Foundation, Chicago, US).

RESULTS:

Of the 50 patients investigated, 28 received a histological diagnosis of a high-grade glioma (WHO grade III-IV), and 22 patients received a diagnosis of a low-grade glioma (WHO grade I-II). There were 1 grade I pilocytic astrocytoma, 8 grade II diffuse astrocytomas, 5 grade II oligodendrogliomas and 8 grade II oligoastrocytomas, 2 grade III anaplastic astrocytomas, 3 grade III anaplastic oligodendrogliomas, 4 grade III anaplastic oligoastrocytomas and 19 grade IV glioblastomas.

Manual and Automatic Glioma Segmentation

Figure 2 shows results from the T1-w image brain masking of a patent diagnosed with a low-grade diffuse astrocytoma and a patient diagnosed with a high-grade glioblastoma. Figure 3

shows results from the FCM cluster analysis and binary morphological image operations, respectively, in the same patients. When comparing the manually defined glioma volumes on a pixel-by-pixel basis across the four observers, the mean sensitivity (± st.errors) when identifying *true* glioma pixels was 59% (±2%) for the low-grade gliomas and 57% (±2%) for the high-grade gliomas. The corresponding pixel-by-pixel PPV for the manually defined tumor volumes was 89% (±1%) for the low-grade gliomas and 87% (±1%) for the high-grade gliomas.

For the automatic tumor volumes, the pixel-by-pixel sensitivity values for the low-grade gliomas (83%±2%) and high-grade gliomas (69%±4%) were significantly higher than the sensitivity values of the manual method (Mann-Whitney; $P<.001$ and $P=.005$, respectively). The corresponding pixel-by-pixel PPVs for the low-grade gliomas (66%±3%) and high-grade gliomas (73%±4%) were significantly lower than the PPVs of the manual method ($P<.001$ and $P=.004$, respectively). Examples of manual and automatic glioma volume are shown in Figure 4. Compared to the observer readings, the FCM cluster algorithm was able to identify 20/20 gliomas with no reported contrast enhancement (i.e. no enhancement detected), 6/7 gliomas with moderate enhancement and 23/23 gliomas with extensive enhancement. The correlations between (a) moderate- and no enhancement and (b) extensive- and no enhancement were significant at the .01 level ($R_s=.904$ and $R_s=1.00$, respectively).

Glioma Characterization:

Figure 5 show the resulting histogram rCBV distribution signatures when using the manual and automatic glioma volumes on the rCBV maps of the patients shown in Figures 2-4. Here, the histogram signatures from the manual method were mean signatures across the observers. For all observers, there was no significant difference ($P=.576$-$.970$) between the A_z values (± st.errors) from the glioma grading using manual glioma volumes ($A_z=.875±.049$-$.908±.040$)

and automatic glioma volumes (A_z=.890±.046). The ROC curves for the manual glioma volume (based on mean histogram peak values across observers) and the automatic glioma volumes are shown in Figure 6.

For the Kaplan-Meier survival analysis, the log-rank value between the low-risk group and high-risk group (Mantel-Cox Chi-Square = 14.984, *P*<.001) for the automatic method was higher than the log-rank values from the manual method (9.441-12,022, *P*=.001-.002). In Table 1, log-rank comparisons between actual survival data (i.e. survival status from last two years) and the survival curves of the manual and automatic method is shown. The corresponding Kaplan-Meier survival curves for (a) actual survival data, (b) the manual method (based on mean histogram peak values across observers) and (c) the automatic glioma volumes are shown in Figure 7.

DISCUSSION:

Although current methods for pre-surgical glioma characterization from DSC imaging show promise(4-9), the availability of these methods is mainly restricted to large institutions with research focus. This may be partially due to the relatively complex image analysis with high inherent user-dependence. Recently, histogram analysis methods have been proposed in order to address this issue(10;11). Based on distribution analysis of rCBV values within the glioma volume, results suggest that histogram analysis methods provide reproducible data with high inter-observer agreement. However, although promising, the current utility of the histogram method in clinical practice is relatively ineffective. It has been reported that observers use longer time per patient for the histogram method compared to other glioma grading methods as the tumor volume has to be manually identified across several MR image slices(10). Also, a study reported an average of 28% ± 12% in inter-observer variations and 20% ± 15% in intra-observer variations when performing manual tumor volume definition over a period of

one month(34). Thus, there is a clear need for more time-efficient, user-independent methods for identifying total tumor volume when characterizing gliomas from DSC imaging.

In our study, we have compared the diagnostic efficacy of pre-surgical glioma characterization from DSC imaging using manually defined and automatically segmented glioma volumes. In agreement with other studies related to automatic tumor segmentation(15;17), our results suggest that the diagnostic efficacy of DSC glioma characterization when using a knowledge-based FCM glioma segmentation method correspond well with the results from manually defined tumor volumes. The diagnostic accuracy values obtained in our study were similar to those reported in the literature(5;11;34;35) and there was no significant difference between the A_z values of the automatic and manual method. Further, the highest log-rank value from the Kaplan-Meier survival analysis was observed when using the automatic method. Also, compared to actual survival data, the log-rank values for the histogram analysis using the automatic segmentation method was similar, or better, than the manual method. Although the correlation to other diagnostic methods was not within the scope of our study, high diagnostic efficacy is critical for any diagnostic method(36). As shown by others(8), the use of DSC imaging for identification of high- and low-risk patient groups with respect to time to progression of glioma grade or patient survival provides important information concerning treatment planning. As sampling error and inaccessible tumors may limit stereotactic biopsy sampling, alternative non-invasive methods are desirable.

Compared to the manually defined tumor volumes, significantly higher sensitivity values were observed for both low-grade and high-grade gliomas when using automatic tumor segmentation. This result suggests that the automated segmentation method produces less

false negative tumor pixels than the manual method. However, although our values are consistent with FCM methods reported in the literature(22), the significantly lower pixel-by-pixel PPVs of the automatic method suggest that the proposed segmentation routine is less conservative than the manually defined glioma volumes, consequently including more hyper-intense areas as seen on the T2-w and FLAIR images. Although the metabolically active tumor area is mainly restricted to the visible T2 tumor border(29;30), it is well know that gliomas are infiltrating tumors with indistinct borders on anatomical MR images. This argument should in turn favor a less conservative method for tumor delineation such as the automatic segmentation method, and that a high sensitivity value is more important than a high PPV. Furthermore, although not a focus of our study, this result also indicates that both the manual and automatic segmentation routine used in our study may not be an adequate measure of tumor volume for quantitative assessment of tumor growth and for aiding neurosurgeons intra-operatively. For quantification of tumor volume, growth and invasion, alternative methods such as threshold-based, semi-automated methods(37) or adaptive template–moderated classification routines(20) may be more suitable. Nevertheless, when grading gliomas pre-surgically, it has been shown that potentially imperfect tumor volume definition has little influence on the diagnostic accuracy and reproducibility of the histogram method(10). Alternatively, brain lesion delineation may be improved by multi-spectral segmentation of both anatomical and perfusion images(38) or by diffusion tensor imaging (DTI)(39). However, in our experience, the relatively low spatial resolution of current perfusion and diffusion MR sequences do not improve the proposed segmentation routine with respect to pixel-by-pixel tumor area comparisons or pre-surgical glioma characterization from DSC imaging. Also, in our study, higher CBV values in the tumor penumbra outside the tumor borders as suggested by the T2-w images was not observed on a consistent basis.

The FCM cluster algorithm detected presence of contrast enhancement in all 23 gliomas with extensive enhancement and no false positives in the 20 gliomas with no contrast enhancement. However, 1 of 7 gliomas with moderate contrast enhancement was not identified by the FCM clustering routine as having contrast enhancement. Although these areas are typically within the hyper-intense areas seen on the T2-w and FLAIR images, and thus included in the final glioma mask, our results suggest that the current clustering scheme might not be sensitive enough to consistently identify vague post-contrast T1-w image intensity variations. Hence, manual selection of image slices with confirmed presence of contrast enhancement prior to FCM clustering might further improve the automatic segmentation routine.

The segmentation procedure used in our study was unsupervised(18;19). Compared to a supervised segmentation procedure(14-17), an unsupervised procedure does not use manually defined training data. A drawback with unsupervised tumor segmentation is that the final glioma mask may be erroneous if initial steps of the knowledge-based operations are violated. This is especially a problem in areas with heterogeneous MR image intensity values or indistinct borders between tissue types. In addition to not being dependent on a large training data sample, an advantage of using unsupervised methods is the lack of subjective human interactions associated with user variability. Hence, the optimal segmentation procedure on our data may be a combination of the two. We hypothesized that manual inspection of the different segmentation steps may further improve the diagnostic efficacy of the automatic segmentation method. In this, the operator will be able to approve or disapprove the results of each step and potentially remove non-tumor slices prior to image analysis. However, in order to minimize user variability, the operator will not be able to interfere with the FCM clustering

procedure. This procedure is currently being implemented as part of our segmentation routine and will be evaluated in future studies.

A limitation to our study is that segmentation techniques based on pixel intensities are inherently prone to partial volume effects and intra- / inter-slice intensity variations due to inhomogeneities in the MR imaging field. This problem may be prominent if the glioma volume as seen on the anatomical MR images is relatively small compared to non-tumor areas mimicking tumor characteristics. To reduce the influence of these confounding factors, an adaptive histogram equalization procedure was used. Furthermore, the final binary glioma mask was a combination of coregistered 2D axial T1-w and T2-w images and 2D coronal FLAIR images with a slice thickness of 5mm and an inter-slice gap of 1.5mm which may result in suboptimal glioma volume definitions. Although the high diagnostic efficacy values obtained in our study suggest that these confounding factors have relatively little influence on pre-surgical glioma characterization., the additional gain of replacing the 2D sequences with isotropic 3D sequences is currently been investigated at our institution. Finally, no special care was taken to exclude rCBV pixels from large vessels within the segmented glioma volume. Although large vessels can be appreciated as dark areas on T2-w images and thus should be excluded during the knowledge-based operations(1;3;29), it is well known that the T2*-effect from large vessels in gradient-echo echo-planar perfusion imaging result in an over-estimation of the intravascular susceptibility effect in pixels adjacent to large vessels. The consequence of this is over-estimated rCBV values in pixels not recognized as vessel tissue by the segmentation procedure. To address this problem, we are currently developing routines for automatic vessel removal based on cluster analysis of the DSC dynamic first-pass parameters(40). This approach is attractive in that it provides a mask which covers all pixels

affected by the intravascular susceptibility effect as seen on the DSC images and not vessel-like pixels as seen on anatomical MR images.

In conclusion, we have shown that using automatically segmented glioma volumes from knowledge-based FCM clustering provides similar values for diagnostic accuracy and patient outcome as using manually selected volumes when characterizing gliomas from DSC MR imaging pre-surgically. This simple segmentation technique holds promise as a user independent approach to selecting the tumor area used in pre-surgical glioma characterization, a vital step towards a fully automated MR-based analysis tool.

ABBREVIATIONS:

DSC = dynamic susceptibility-contrast, MR = magnetic resonance, CBV = cerebral blood volume, 2D = 2 dimensional, 3D = 3 dimensional, FCM = fuzzy c-means, WHO = world health organization, T2-w = T2-weighted, T1-w = T1-weighted, CSF = cerebrospinal fluid, PPV = positive predictive values, NPV = negative predictive values, ROC = receiver operating characteristic, DTI = diffusion tensor imaging.

ACKNOWLEDGEMENTS:

We thank David Scheie, MD, from the Pathology Clinic, Rikshospitalet University Hospital, Oslo, Norway for providing the histopathological results.

REFERENCES:

(1) Talos IF, Zou KH, Ohno-Machado L et al. Supratentorial low-grade glioma resectability: statistical predictive analysis based on anatomic MR features and tumor characteristics. Radiology 2006;239(2):506-513.

(2) Dean BL, Drayer BP, Bird CR et al. Gliomas: classification with MR imaging. Radiology 1990;174(2):411-415.

(3) Johnson PC, Hunt SJ, Drayer BP. Human cerebral gliomas: correlation of postmortem MR imaging and neuropathologic findings. Radiology 1989;170(1 Pt 1):211-217.

(4) Covarrubias DJ, Rosen BR, Lev MH. Dynamic magnetic resonance perfusion imaging of brain tumors. Oncologist 2004;9(5):528-537.

(5) Knopp EA, Cha S, Johnson G et al. Glial neoplasms: dynamic contrast-enhanced T2*-weighted MR imaging. Radiology 1999;211(3):791-798.

(6) Lev MH, Ozsunar Y, Henson JW et al. Glial tumor grading and outcome prediction using dynamic spin-echo MR susceptibility mapping compared with conventional contrast-enhanced MR: confounding effect of elevated rCBV of oligodendrogliomas [corrected]. AJNR Am J Neuroradiol 2004;25(2):214-221.

(7) Law M, Oh S, Babb JS et al. Low-grade gliomas: dynamic susceptibility-weighted contrast-enhanced perfusion MR imaging--prediction of patient clinical response. Radiology 2006;238(2):658-667.

(8) Law M, Oh S, Johnson G et al. Perfusion magnetic resonance imaging predicts patient outcome as an adjunct to histopathology: a second reference standard in the surgical and nonsurgical treatment of low-grade gliomas. Neurosurgery 2006;58(6):1099-1107.

(9) Wetzel SG, Cha S, Johnson G et al. Relative cerebral blood volume measurements in intracranial mass lesions: interobserver and intraobserver reproducibility study. Radiology 2002;224(3):797-803.

(10) Emblem KE, Nedregaard B, Nome T et al. Glioma grading by using histogram analysis of blood volume heterogeneity from MR-derived cerebral blood volume maps. Radiology 2008;247(3):808-817.

(11) Young R, Babb J, Law M, Pollack E, Johnson G. Comparison of region-of-interest analysis with three different histogram analysis methods in the determination of perfusion metrics in patients with brain gliomas. J Magn Reson Imaging 2007;26(4):1053-1063.

(12) Taxt T, Lundervold A. Multispectral analysis of the brain using magnetic resonance imaging. IEEE Trans Med Imaging 1994;13(3):470-481.

(13) Vannier MW, Butterfield RL, Jordan D, Murphy WA, Levitt RG, Gado M. Multispectral analysis of magnetic resonance images. Radiology 1985;154(1):221-224.

(14) Dickson S, Thomas BT, Goddard P. Using neural networks to automatically detect brain tumours in MR images. Int J Neural Syst 1997;8(1):91-99.

(15) Abdolmaleki P, Mihara F, Masuda K, Buadu LD. Neural networks analysis of astrocytic gliomas from MRI appearances. Cancer Lett 1997;118(1):69-78.

(16) Zhou J, Chan K, Chong H, Krishnan S. Extraction of Brain Tumor from MR Images Using One-Class Support Vector Machine. Conf Proc IEEE Eng Med Biol Soc 2005;6:6411-6414.

(17) Glotsos D, Spyridonos P, Cavouras D, Ravazoula P, Dadioti PA, Nikiforidis G. An image-analysis system based on support vector machines for automatic grade diagnosis of brain-tumour astrocytomas in clinical routine. Med Inform Internet Med 2005;30(3):179-193.

(18) Clark MC, Hall LO, Goldgof DB, Velthuizen R, Murtagh FR, Silbiger MS. Automatic tumor segmentation using knowledge-based techniques. IEEE Trans Med Imaging 1998;17(2):187-201.

(19) Bezdek JC, Hall LO, Clark MC, Goldgof DB, Clarke LP. Medical image analysis with fuzzy models. Stat Methods Med Res 1997;6(3):191-214.

(20) Kaus MR, Warfield SK, Nabavi A, Black PM, Jolesz FA, Kikinis R. Automated segmentation of MR images of brain tumors. Radiology 2001;218(2):586-591.

(21) Itskovich VV, Samber DD, Mani V et al. Quantification of human atherosclerotic plaques using spatially enhanced cluster analysis of multicontrast-weighted magnetic resonance images. Magn Reson Med 2004;52(3):515-523.

(22) Fletcher-Heath LM, Hall LO, Goldgof DB, Murtagh FR. Automatic segmentation of non-enhancing brain tumors in magnetic resonance images. Artif Intell Med 2001;21(1-3):43-63.

(23) Kleihues P, Cavenee W. Astrocytic tumors & Oligodendroglial tumors and mixed gliomas. *The WHO classification of tumors of the nervous system*. 2 ed. Lyon: International Agency for Research on Cancer; 2000. p. 9-70.

(24) Rosen BR, Belliveau JW, Vevea JM, Brady TJ. Perfusion imaging with NMR contrast agents. Magn Reson Med 1990;14(2):249-265.

(25) Ostergaard L, Weisskoff RM, Chesler DA, Gyldensted C, Rosen BR. High resolution measurement of cerebral blood flow using intravascular tracer bolus passages. Part I: Mathematical approach and statistical analysis. Magn Reson Med 1996;36(5):715-725.

(26) Boxerman JL, Schmainda KM, Weisskoff RM. Relative cerebral blood volume maps corrected for contrast agent extravasation significantly correlate with glioma tumor grade, whereas uncorrected maps do not. AJNR Am J Neuroradiol 2006;27(4):859-867.

(27) Maes F, Collignon A, Vandermeulen D, Marchal G, Suetens P. Multimodality image registration by maximization of mutual information. IEEE Trans Med Imaging 1997;16(2):187-198.

(28) Schmainda KM, Rand SD, Joseph AM et al. Characterization of a first-pass gradient-echo spin-echo method to predict brain tumor grade and angiogenesis. AJNR Am J Neuroradiol 2004;25(9):1524-1532.

(29) Grier JT, Batchelor T. Low-grade gliomas in adults. Oncologist 2006;11(6):681-693.

(30) Pirzkall A, Nelson SJ, McKnight TR et al. Metabolic imaging of low-grade gliomas with three-dimensional magnetic resonance spectroscopy. Int J Radiat Oncol Biol Phys 2002;53(5):1254-1264.

(31) Zuiderveld K. Contrast Limited Adaptive Histogram Equalization. In: Heckbert PS, editor. Graphics Gems IV. AP Professional, Boston; 1994. p. 474-85.

(32) Ashburner J, Friston K. Multimodal image coregistration and partitioning--a unified framework. Neuroimage 1997;6(3):209-217.

(33) Otsu NA. A Threshold Selection Method from Gray-Level Histograms. IEEE Transactions on Systems, Man, and Cybernetics 1979;9(1):62-66.

(34) Mazzara GP, Velthuizen RP, Pearlman JL, Greenberg HM, Wagner H. Brain tumor target volume determination for radiation treatment planning through automated MRI segmentation. Int J Radiat Oncol Biol Phys 2004;59(1):300-312.

(35) Law M, Yang S, Wang H et al. Glioma grading: sensitivity, specificity, and predictive values of perfusion MR imaging and proton MR spectroscopic imaging compared with conventional MR imaging. AJNR Am J Neuroradiol 2003;24(10):1989-1998.

(36) Fryback DG, Thornbury JR. The efficacy of diagnostic imaging. Med Decis Making 1991;11(2):88-94.

(37) Joe BN, Fukui MB, Meltzer CC et al. Brain tumor volume measurement: comparison of manual and semiautomated methods. Radiology 1999;212(3):811-816.

(38) Kvinnsland Y, Brekke N, Taxt TM, Gruner R. Multispectral analysis of multimodal images. Acta Oncol 2008;1-8.

(39) Price SJ, Jena R, Burnet NG et al. Improved delineation of glioma margins and regions of infiltration with the use of diffusion tensor imaging: an image-guided biopsy study. AJNR Am J Neuroradiol 2006;27(9):1969-1974.

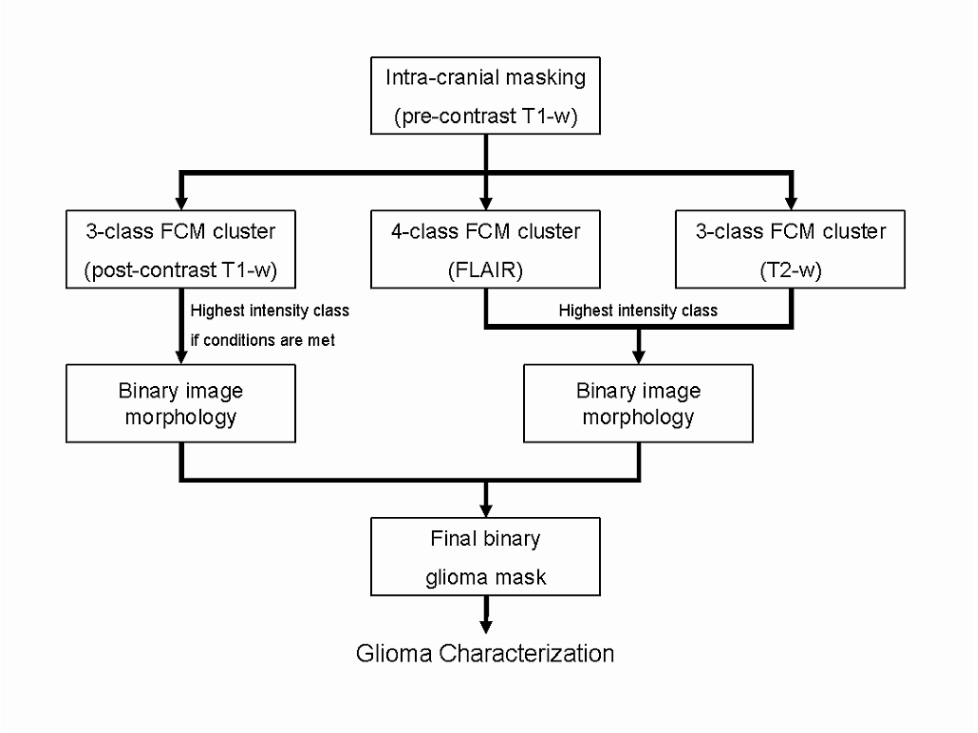
(40) Mouridsen K, Christensen S, Gyldensted L, Ostergaard L. Automatic selection of arterial input function using cluster analysis. Magn Reson Med 2006;55(3):524-531.

Table 1: Log-rank values from Kaplan-Meier survival curves when trying to identify a low-risk patient group (expected survival >2 years) and a high-risk patient group (expected survival <2 years). The log-rank values with *P*-values illustrate the difference between actual survival data and (a) the manual (Observers 1-4) and (b) the automated method. A lower log-rank value and a higher *P*-value indicate stronger correlations with actual survival data.

	Observer 1 Log-rank (*P*-value)	**Observer 2** Log-rank (*P*-value)	**Observer 3** Log-rank (*P*-value)	**Observer 4** Log-rank (*P*-value)	**Automated** Log-rank (*P*-value)
Low-risk group	2.976 (.085)	4.487 (.034)	4.487 (.034)	4.617 (.032)	5.589 (.018)
High-risk group	9.795 (.002)	8.485 (.004)	9.013 (.003)	9.877 (.002)	5.566 (.018)

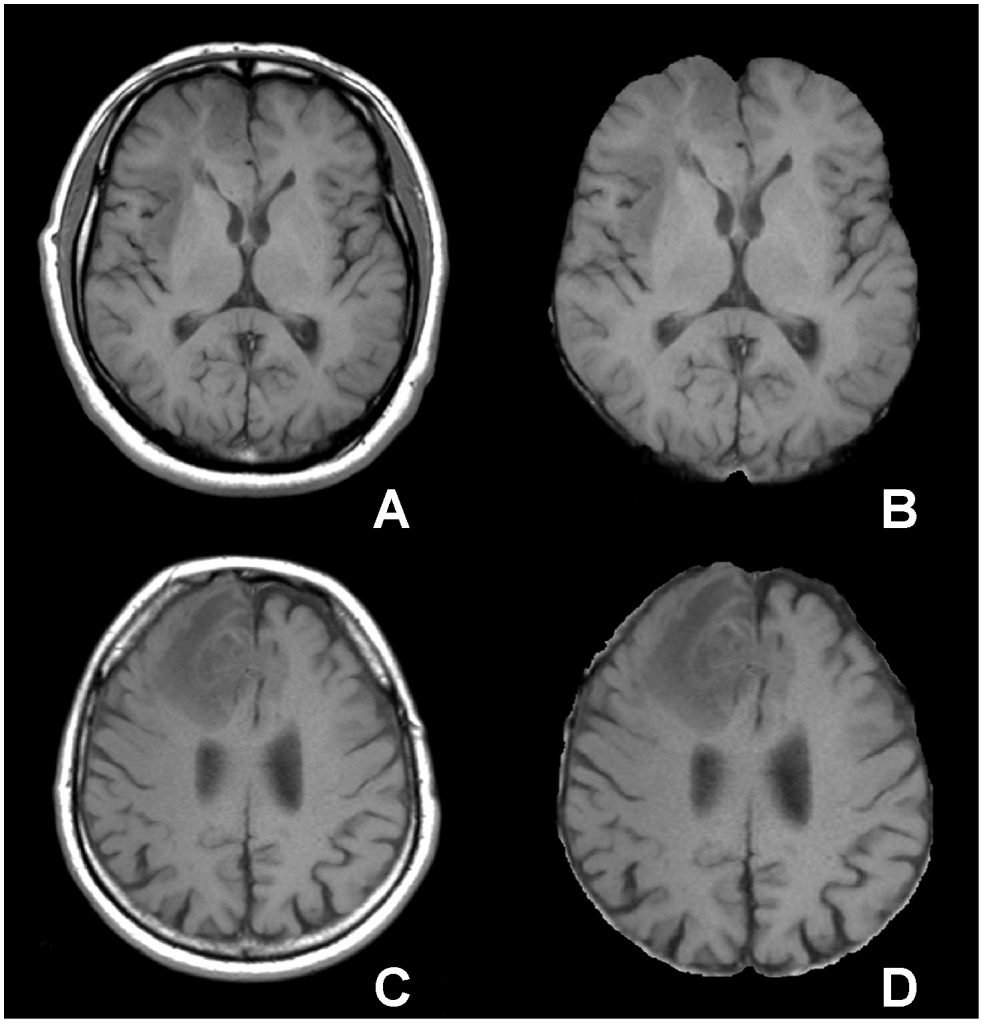
FIGURES AND LEGENDS:

FIGURE 1



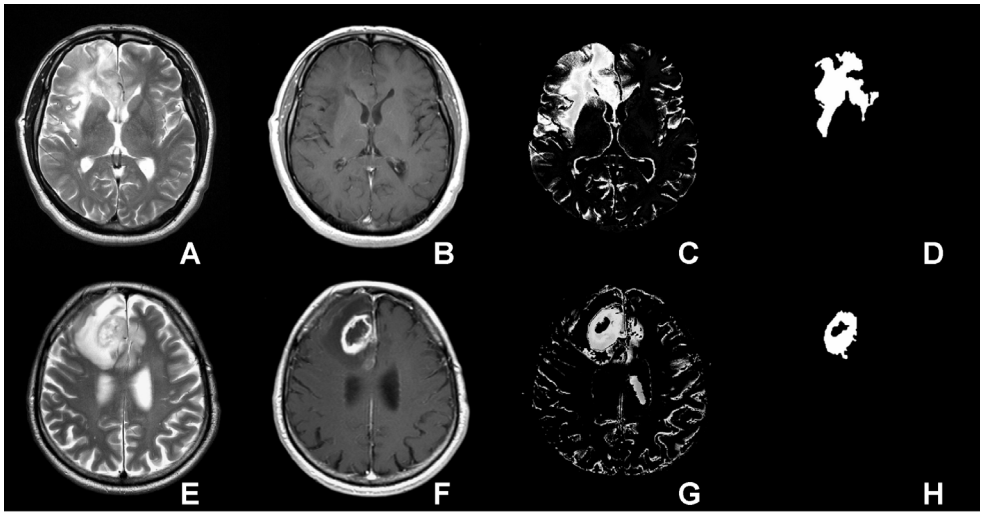
Schematic flow-diagram of the entire automatic segmentation procedure performed in our study. The method utilizes anatomical images (in brackets) part of a typical brain tumor MR imaging protocol. The entire segmentation procedure took approximately 4 minutes using Matlab 2007a and a standard desktop computer.

FIGURE 2



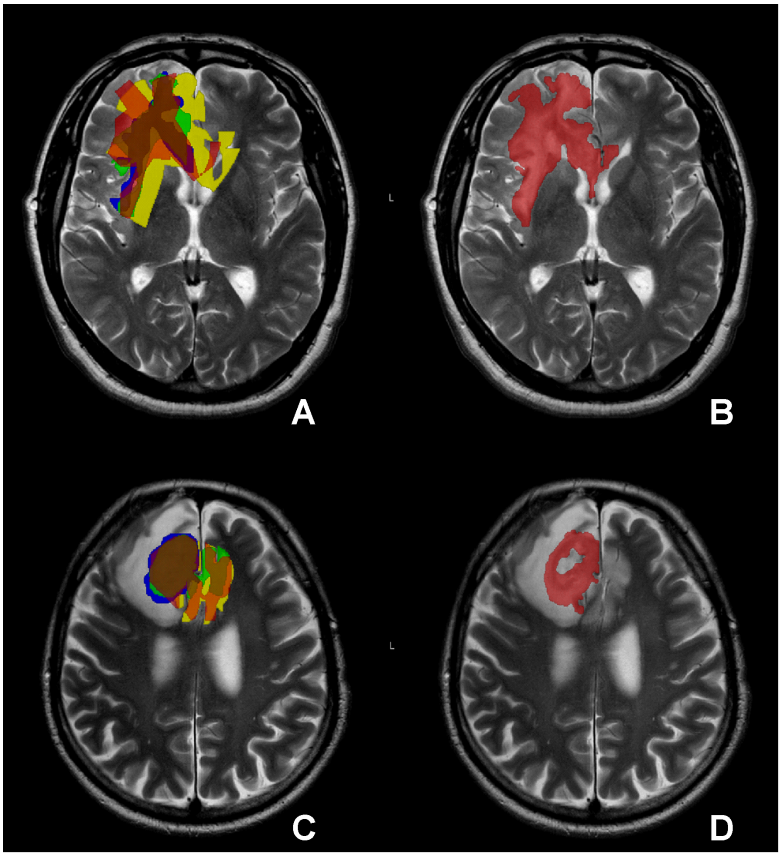
Axial pre-contrast T1-w images of patients diagnosed with a grade II oligodendroglioma **(A)** and a grade IV glioblastoma **(C)**. The resulting images after applying intra-cranial brain masking are shown in **(B)** and **(D)**, respectively. Non-brain pixels are removed using a template-based, intra-cranial brain mask procedure in Statistical Parametric Mapping (SPM5). The binary brain mask is used on all MR images prior to the FCM clustering. The tumor region is recognized as brain tissue and is not removed.

FIGURE 3



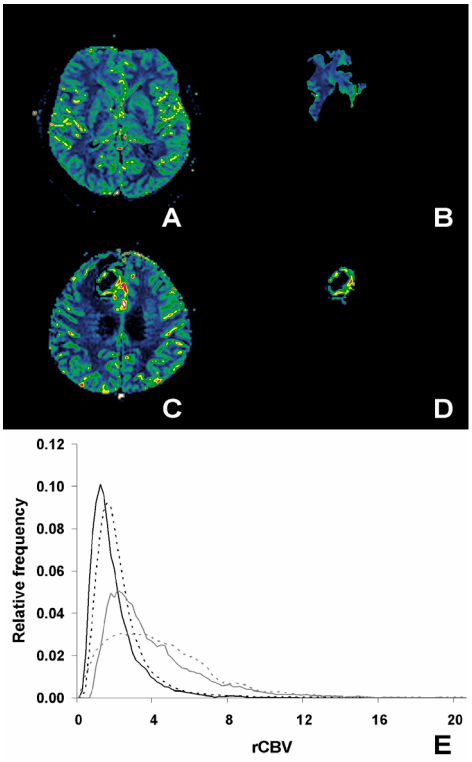
Results of the FCM clustering procedure in the low-grade glioma patient **(A-D)** and high-grade glioma patient **(E-H)** shown in Figure 2. Using anatomical MR images, such as T2-w **(A, E)** and T1-w post contrast **(B, F)** images, the resulting cluster class representative of the highest mean pixel values is thought to represent the tumor class. Examples of the resulting 'tumor class' images from the 3-class FCM clustering of the T2-w images is shown in **(C, G)**. These images are continuously scaled with values ranging from [0-1], illustrating the strength of the association between a given pixel and the cluster class. From this, Otsu's method is used to reduce the continuous cluster images into binary tumor masks and a set of knowledge-based operations is applied to "clean" the binary tumor mask in order to remove spur non-tumor pixels mimicking tumor tissue **(D, H)**. Finally, a 3D seed growing algorithm is applied on the complete image stack to identify tumor regions in neighboring image slices.

FIGURE 4



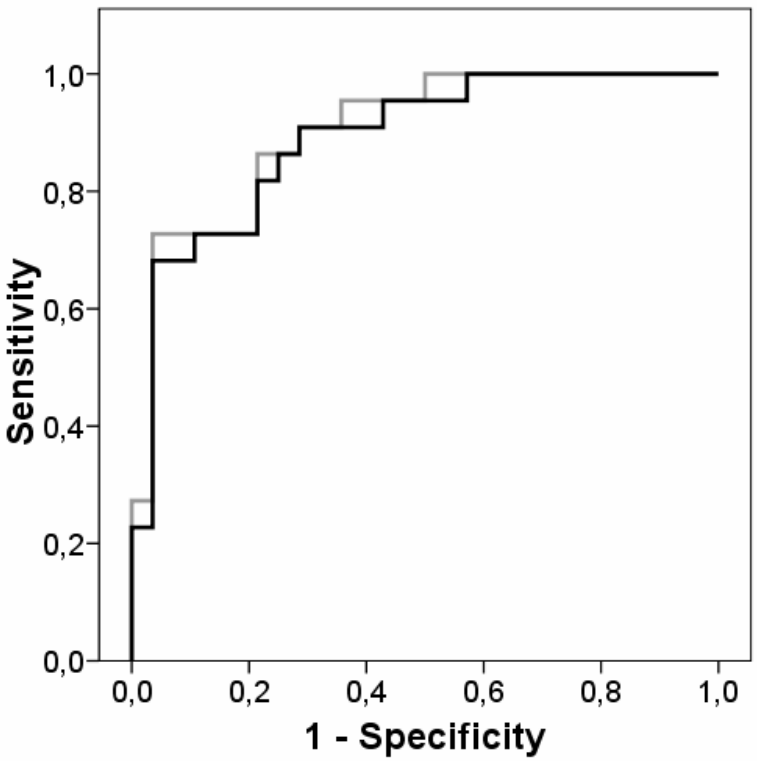
Manual- and automatic binary tumor masks superimposed on T2-w images of the low-grade **(A, B)** and high-grade **(C, D)** glioma patients shown in Figures 2 and 3. For the manual tumor definitions **(A, C)**, the sensitivity values of the four observers (in colors yellow, blue, green and red) were significantly lower and the PPV were significantly higher than the automatic method (in red) **(B, D)**, respectively. Despite the large variations between the manual observers shown in **(A, C)**, the diagnostic accuracy of the pre-surgical glioma grading was high for all observers.

FIGURE 5



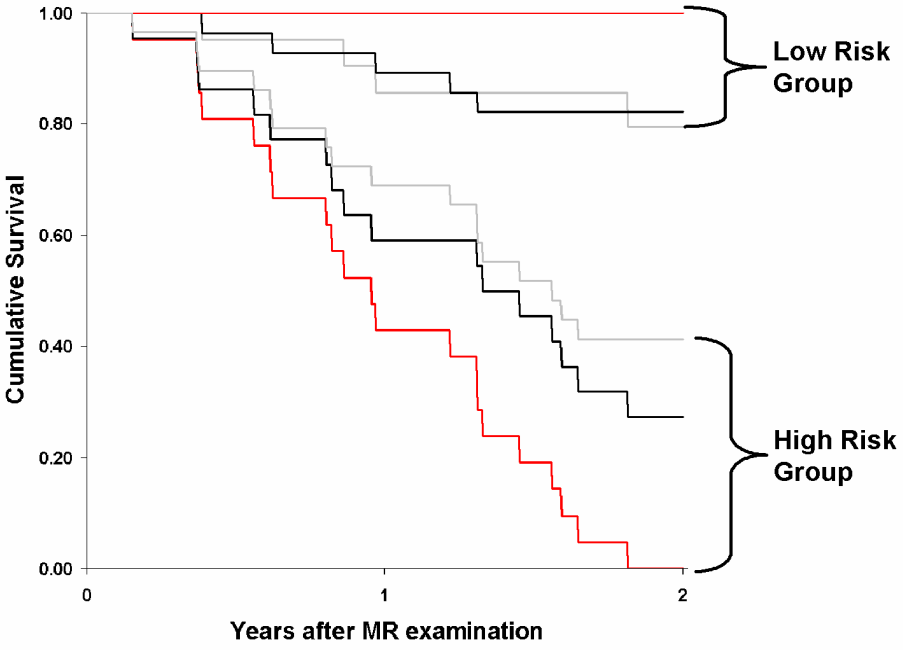
Axial rCBV maps of the low-grade **(A)** and high-grade **(C)** glioma patients shown in Figures 2-4. The resulting areas from the FCM clustering used in the pre-surgical glioma characterization are shown in **(B)** and **(D)**, respectively. After deriving rCBV values from all tumor areas in all slices, normalized histogram signatures of the complete distribution of rCBV values from the manual (dotted lines) and automatic (solid lines) tumor volumes of the high-grade (gray lines) and low-grade (black lines) glioma patients can be assessed **(E)**. The manual histogram signatures are average signatures across the four observers. Note the reduced histogram peak height in the high-grade histogram signatures attributed to increased vascular heterogeneity.

FIGURE 6



Using histology as reference (high-grade/low-grade gliomas), the plot shows receiver operator characteristics (ROC) curves for the manual (gray line) and the automatic method (black line). The manual ROC curve is based on mean histogram peak values across observers. For any observer, the A_z values (± st.errors) from on the manual tumor volumes (A_z=.875±.049-.908±.040) were not significantly different (*P*=.576-.970) from the automatic method (A_z=.890±.046).

FIGURE 7



Kaplan-Meier survival curves from actual survival data (red lines), the manual method across observers (gray lines) and the automatic method (black lines). Based on actual survival data over a two year period, logistic regression was used to derive optimal histogram peak height cut-off values between a high-risk and a low-risk patient group. Here, the high-risk group was defined as those patients with an expected survival from MR examination date to death of less than 2 years, whereas the low-risk group was defined as those patients with an expected survival of more than 2 years. All patients were included in the analysis and a higher log-rank value between the high- and low-risk group was observed when using the automatic method (14.984, $P<.001$) compared to the manual method (9.441-12,022, $P=.001$-.002).

III

Automatic Vessel Removal in Gliomas from Dynamic Susceptibility Contrast Imaging

Kyrre E. Emblem, M.Sc, , Paulina Due-Tonnessen, MD, John K. Hald, MD, PhD, Atle Bjornerud, PhD

From the Departments of Medical Physics (K.E.E., A.B.), Clinic for Imaging and Intervention (P.D-T., J.K.H.) and The Interventional Center (K.E.E), Rikshospitalet University Hospital, Sognsvannsveien 20, N-0027 Oslo, Norway. Department of Physics (A.B), University of Oslo, Sem Saelandsvei 24, N-0371 Oslo, Norway.

Corresponding author: Kyrre E. Emblem

Full address:
The Interventional Center, Gaustad
Rikshospitalet University Hospital
Sognsvannsveien 20
N-0027 Oslo
Norway

Telephone number: +47 23070145
Fax number: +47 23070110
E-mail address: kyrre.eeg.emblem@rikshospitalet.no

Presented in part at the International Society for Magnetic Resonance in Medicine, 630p, 2008; May 3-9, Toronto, Canada

ABSTRACT:

The presence of macroscopic vessels within the tumor region is potentially a confounding factor in MR-based dynamic susceptibility contrast (DSC) enhanced glioma grading. In order to distinguish between such vessels and the elevated cerebral blood volume (CBV) of brain tumors, we propose a vessel segmentation technique based on clustering of multiple parameters derived from the dynamic contrast enhanced first-pass curve. Seventy-seven adult patients with histologically confirmed gliomas were imaged at 1.5 Tesla and glioma region-of-interests (ROIs) were derived from the conventional MR images by a neuroradiologist. The diagnostic accuracy of applying vessel exclusion by segmentation of glioma ROIs with vessels included was assessed using a histogram analysis method and compared to glioma ROIs with vessels included. For all measures of diagnostic efficacy investigated, the highest values were observed when the glioma diagnosis was based on vessel segmentation in combination with an initial mean transit time (MTT) mask. Our results suggest that vessel segmentation based on DSC parameters may improve the diagnostic efficacy of glioma grading. The proposed vessel segmentation is attractive because it provides a mask which covers all pixels affected by the intravascular susceptibility effect.

Key words: DSC MRI, glioma grading, histogram analysis, vessel segmentation, k-means clustering

INTRODUCTION:

Dynamic susceptibility contrast (DSC) enhanced MR imaging is a widely used technique for assessing the microvascularity of brain tumors[1]. Several studies have shown high correlation between histopathological glioma grade and cerebral blood volume (CBV) derived from DSC imaging[2-8]. However, presence of elevated CBV values from large vessels within the tumor area may be a confounding factor[8-10]. Low-grade gliomas (LGGs) are typically located in the frontal- or temporal lobes[11-13], in close proximity to the anterior- or middle cerebral arteries, which might result in an overestimation of glioma grade[9]. To correct for this, large vessels within the tumor area is manually excluded from the glioma region-of-interest (ROI) prior to glioma grading[10,14]. Hence, current glioma grading techniques are inherently operator-dependent, time consuming and reliable results may only be obtained by experienced operators with good anatomical knowledge[8,14,15]. This problem becomes increasingly relevant as automated tumor segmentation techniques are introduced as an alternative to manual tumor identification[16]. One important consideration when attempting to identify vessels in DSC images is that the susceptibility effect – and consequent elevation in CBV values – due to vessels containing contrast agent will expand outside the vessel itself[17]. Therefore, the anatomical images cannot be used to identify and remove vessels since such an approach will not address the extravascular effects.

To date, little focus has been placed on automatic routines for identification and removal of large vessels from CBV maps. However, studies have reported using an automatic procedure for identification of the arterial input function (AIF) based on k-means cluster analysis[18,19]. In a k-means cluster algorithm, *n* objects are divided into a given number of cluster classes of which the objects in a specific class share a common set of attributes. One study reported excellent agreement between manual AIF selection and the automatic routine[18]. It is our hypothesis that a similar k-means cluster analysis approach[20] can be used to automatically identify all image pixels which are affected by large vessels (arteries and veins), and hence eliminate these from the final CBV maps. The separation of vessels from tissue is based on the temporal characteristics of the first-pass curve, including contrast arrival time (T_0), area under the first-pass curve (AUC) and wash-out characteristics of the contrast agent which can be assessed by deriving the first moment of the AUC (fmAUC):

$$fmAUC = \int \Delta R_2(t) \cdot t \cdot dt \qquad [1]$$

Here, ΔR_2 is the change in transverse relaxation rate, assumed to have a linear relationship with contrast agent concentration. It is expected that both arteries and veins will exhibit a high AUC. Additionally, arteries (veins) will exhibit a short (long) T_0 and a low (high) fmAUC[18]. Combining these features, it is then hypothesized that arteries and veins can be distinguished from highly perfused tumor tissue.

In view of the above, the purpose of our study was to assess whether macroscopic vessels could reliably be removed from DSC-generated CBV maps and whether the vessel removal can improve the diagnostic efficacy of glioma grading, using histology as reference.

MATERIALS AND METHODS:

Patient Selection

Study approval was obtained from the Regional Medical Ethics Committee and patients were included only if informed consent was signed. Between July 2005 and November 2007, seventy-seven adult patients (aged 19-78 yrs, mean age 52; 45 males, 32 females) received a histological diagnosis of primary glioma after MR perfusion imaging and subsequent surgery, and agreed to allow the use of their data in our study. The histological evaluation was performed by a neuropathologist using the World Health Organization (WHO) classification of central nervous system (CNS) tumors[13]. As WHO grade I pilocytic astrocytomas have been reported to have elevated CBV values irrespective of tumor grade[21], and these tumors can usually be differentiated based on additional criteria[22], only WHO glioma grades II (low-grade gliomas; LGGs) and grades III-IV (high-grade gliomas; HGGs) were included.

MR Imaging

Imaging was performed at 1.5 Tesla (Siemens Sonata, Symphony or Avanto, Siemens AG, Erlangen, Germany), using an 8-channel- (Symphony/Sonata) or a 12-channel (Avanto) head-coil. The protocol included a 19 slice axial T2-weighted fast spin-echo sequence with 4000/104 (repetition time msec /echo time msec) and a 19 slice axial T1-weighted spin-echo sequence (500/7.7) obtained before and after i.v. contrast agent injection. The voxel size of the axial images was $0.45 \times 0.45 \times 5 mm^3$.

DSC MR imaging was performed using a single-shot gradient-echo echo-planar imaging (GRE-EPI) sequence acquired during contrast agent administration. The imaging parameters were: 1430/46, bandwidth 1345 Hz/pixel (12 axial slices) up to 1720/48, bandwidth 1500

Hz/pixel (14 axial slices), field of view 230x230mm, voxel size 1.80x1.80x5 mm^3, inter-slice gap 1.5mm. For each slice, 50 images were recorded at intervals equal to the repetition time. After approximately 8 time-points, 0.2 mmol/kg of gadobutrol (Bayer Schering Pharma AG, Berlin, Germany) was injected at a rate of 5 mL/sec, immediately followed by a 20 mL bolus of saline (B. Braun Melsungen AG, Melsungen, Germany) also at 5 mL/sec.

Post-processing and Vessel Segmentation

The images were post-processed using a dedicated software package (nordicICE, NordicImagingLab, Bergen, Norway) and Matlab R2007a (MathWorks, Natick, US). Relative (r)CBV maps were generated from the area under the 1/T2* converted first-pass curves and corrected for possible extra-vascular contrast agent leakage effects[23]. Based on the first-pass curves, fmAUC, T_0 and relative mean transit time (rMTT) were estimated for each pixel. rMTT maps were estimated as the ratio $rCBV/\Delta R2_{max}$ where $\Delta R2_{max}$ is the peak height of the first-pass curve. Normalized (n)CBV maps were created from the original rCBV maps by voxel-vise division of the rCBV values with an unaffected slice-specific white matter rCBV value defined by a neuroradiologist[10]. The images were displayed using a black-blue-green-yellow-red-white colormap. All nCBV maps were automatically coregistered with the conventional MR images using a normalized mutual information algorithm[24]. Since malignant tumor tissue can mimic arterial signal with similarly short T_0 and large rCBV, three different methods for automatic vessel segmentation were implemented;

Method 1: Uncorrected vessel masks - an iterative 5-class k-means cluster analysis[20] was performed to identify arteries and veins from the estimated parameters, rCBV, fmAUC and T_0. Arterial pixels were identified as the pixel class with the shortest T_0 combined with the highest rCBV and lowest fmAUC values. Venous pixels were identified as the pixel class with the largest fmAUC combined with the highest rCBV and longest T_0 values. The remaining cluster classes were thought to represent gray- and white matter and cerebrospinal fluid. The resulting binary masks consisted of arterial and venous pixels only and vessel segmented nCBV maps were created by removing pixels in the nCBV maps corresponding to the binary masks. Since malignant tumor tissue can mimic arterial signal with similarly short T_0 and large rCBV, two different additional pre-cluster steps were tested as follows:

Method 2: MTT corrected vessel masks - Prior to identification of arterial and venous pixels (Method 1), an initial iterative 5-class k-means cluster analysis was performed on the standard

rMTT maps to identify and remove the cluster class with the highest mean rMTT value. This was based on the hypothesis that highly vascular tumor tissue would exhibit increased rMTT values relative to unaffected tissue or blood due to a more complex vascular structure[25]. A recent study reported rMTT values of 3.2±0.3s and 3.7±0.6s in unaffected white matter tissue and tumor tissue, respectively[26]. Then, on the four remaining cluster classes, the iterative 5-class k-means cluster algorithm was applied to identify arteries and veins as described above.

Method 3: SI corrected vessel masks - The initial cluster step using the rMTT cluster approach (Method 2) was replaced by a similar 5-class k-means cluster routine removing the cluster class with the highest mean signal intensities (SI) as seen on the T2*-w baseline DSC images prior to the bolus arrival. This was based on the hypothesis that, high intensities on T2-w images are secondary to pathology whereas large blood vessels appear dark.

Manual glioma definition

Based on previously published criteria[14,23,27], a board-certified neuroradiologist with several years experience with MR perfusion imaging defined freehand glioma ROIs based on the anatomical MR image information. All patient-related information and histopathological results were removed from the images. As the diagnostic accuracy of the histogram method is relatively unaffected by imperfect glioma definitions[14], the outmost glioma margins was defined based on the hyper-intensities as seen on the T2-weighted images. For each patient, two sets of glioma ROIs were defined in all MR image slices suggestive of tumor; one with vessels excluded (as identified on the post-contrast T1-weighted images) and one with vessels included.

Comparison of automated and manual vessel identification

The vessel segmentation masks were applied to the glioma ROIs with vessels included, yielding vessel-removed nCBV maps. These maps were then compared to the CBV maps from the glioma ROIs with vessels excluded by the neuroradiologist. To compare the manual ROI definition with results in the literature[14], the neuroradiologist recorded the time used to perform the analysis and how difficult it was to perform (easy, intermediate or difficult).

Glioma Grading

The grading was performed using a previously reported histogram method[14]. Histograms signatures were generated for each patient using nCBV values from the complete glioma

volumes. The range of the nCBV values along the x-axis was kept constant (values; 0-20) and the area under each histogram curve was normalized to one. Glioma malignancy was assessed by measuring the maximum normalized peak height of the distributions. The histograms were divided into 108 bins, a number shown to give the maximal diagnostic accuracy[14].

Statistical analysis

Mean histogram peak heights with standard errors for glioma grade II, III and IV were assessed from the glioma ROIs with and without vessels included and for the vessel segmentation. In order to conform to similar studies[14,15,28], logistic regression was used to derive sensitivity values, specificity values, negative predictive values (NPV) and positive predictive values (PPV) between LGGs and HGGs. A glioma classified as HGG or LGG by both observer data and histology was considered as a true-positive/true-negative finding, respectively. In addition, the areas A_z (with standard errors) under the Receiver Operating Characteristic (ROC) curves were assessed. Here, an A_z value of 1 suggest perfect discrimination between HGG and LGG whereas an A_z value closer to .5 suggest a weaker discrimination between the two groups. Kaplan-Meier survival curves were derived from the glioma ROIs with and without vessel included and for the vessel segmentation. Here, optimal histogram peak height cut-off values between a low risk and a high risk group were derived using binary logistic regression. To test for equality of the survival curves for the high- and low risk groups, log-rank tests (Mantel Cox) were performed. In the log-rank test, all time points were weighted equally. Statistical analysis was performed using SPSS 13 (SPSS, Inc., Chicago, IL, USA).

RESULTS:

Of the 77 patients investigated, 45 received a histological diagnosis of a HGG and 32 patients received a diagnosis of a LGG. Of the LGGs, there were 19 grade II diffuse astrocytomas, 5 grade II oligodendrogliomas and 8 grade II oligoastrocytomas. Of the HGGs, there were 3 grade III anaplastic astrocytomas, 3 grade III anaplastic oligodendrogliomas, 4 grade III anaplastic oligoastrocytomas and 35 grade IV glioblastomas.

Vessel segmentation

Figure 1 show the result of using the different vessel masks on nCBV maps of a patient diagnosed with a low-grade diffuse astrocytoma (WHO grade II). Figure 1 also shows the resulting binary vessel mask when using the uncorrected, MTT corrected and SI corrected

vessel masking, respectively. Figure 2 show the difference between the uncorrected- and MTT corrected vessel mask on a patient diagnosed with a high-grade glioblastoma (WHO grade IV). Figure 2 also shows the binary rMTT cluster mask representative of the cluster class with the highest rMTT values which was excluded from the MTT corrected vessel masks. The resulting histograms for the low-grade and high-grade gliomas with and without vessel segmentation are shown in Figure 3.

Manual glioma definition

The neuroradiologist recorded using approximately 10 minutes per patient when drawing the glioma ROIs with vessels excluded and 5 minutes when drawing glioma ROIs with vessels included. When excluding vessels, the degree of difficulty was intermediate. When including vessels, the degree of difficulty was easy to intermediate.

Glioma grading

Mean histogram peak heights with standard errors for glioma grades II, III and IV from the nCBV maps with and without vessel segmentation are shown in Figure 4. The biggest absolute difference between grades II and III and between grades II and IV were observed when using the MTT corrected vessel masks (.030 and .054, respectively). The corresponding sensitivity values, specificity values, NPV and PPV when grading the gliomas as LGGs or HGGs are shown in Table 1. For all measures of diagnostic accuracy, the highest value was obtained using the MTT corrected vessel mask. The lowest values were obtained when using the uncorrected vessel segmentation mask. For the 32 patients diagnosed with LGGs, using the MTT corrected vessel masks resulted in a change of glioma grade in 4 patients compared to the ROIs with vessels excluded. Three LGGs were correctly classified as LGGs and 1 HGG were misclassified as LGG. The 3 LGGs correctly classified as LGGs were all located within the M1 or M2 segments of the middle cerebral artery distribution, whereas the HGG misclassified as LGG was not located in close proximity to the main artery distributions. For the 45 patients diagnosed with HGGs, using the MTT corrected vessel masks resulted in a change of glioma grade in 5 patients compared to the ROIs with vessels excluded. Three HGGs were correctly classified as HGGs and 2 LGG were misclassified as HGG. None of the HGGs were located in proximity to one of the main artery distributions. The resulting ROC curves with A_z values (± standard error) for all methods investigated in our study are shown in Figure 5. In terms of manual ROI definition, a higher A_z value was observed when the

radiologist included vessels in the ROIs compared to the ROIs with vessels excluded (.915±.030 and .881±.038, respectively).

The log-rank (Mantel-Cox Chi-Square) values describing the difference between the survival curves for the low- and high risk group were as follows; glioma ROIs with vessels excluded = 11.484 (P=.001), glioma ROIs with vessels included = 14.806 (P<.001), uncorrected vessel masks = 16.603 (P<.001), MTT corrected vessel masks = 20.390 (P<.001) and SI corrected vessel masks = 15.825 (P<.001). Figure 6 show the resulting Kaplan-Meier survival curves for the glioma ROIs with vessels excluded versus (a) the glioma ROIs with vessels included, (b) uncorrected vessel masks, (c) MTT corrected vessel masks and (d) SI corrected vessel masks, respectively.

DISCUSSION:

It is well known that GRE-EPI sequences are sensitive to T2* effects from large vessels and this effect may cause significant T2* shortening outside the vessel lumen[17,29]. Hence, vessel identification from anatomical images only, is likely to underestimate the true extent of the intravascular susceptibility effect in the DSC images. In our study, we propose a novel vessel segmentation technique which can be applied to parameters derived directly from the DSC images. Although the importance of intra-tumoral macroscopic vessel exclusion has been reported in several studies[9,10,14,30], automatic and user-independent approaches for vessel removal in glioma characterization from MR perfusion imaging are, to the authors knowledge, not reported in the literature. Typically, exclusion of intra-tumoral macroscopic vessels is performed manually by an experienced operator[9,10,14,30] or by applying an upper cut-off value to reduce the effect of very high CBV values related to noise, blood vessels and severe blood-brain-barrier breakdown[28]. However, our results suggest that manual removal of macroscopic vessels from static MR images does not provide the same diagnostic accuracy values as automatic vessel removal based on the dynamic DSC information. Further, a cut-off CBV value is unable to discriminate between arterial, venous and tissue image pixels as the AUC can be similar even though the shapes of the first-pass curves are different. The vessel segmentation technique proposed in our study is fast and provides a direct estimate of pixels which are actually affected by the vascular susceptibility effect, thereby offering a better correction for vessel-induced elevation in tumor CBV values. Further, the method does not require an expert user with good anatomical knowledge to mask out vessel regions prior to glioma area identification. For automated tumor area segmentation routines[16], this feature is

attractive as optimal tumor area identification may rely on hyper-intensities as seen on T2-weighted images or similar MR image types, only. With respect to the manual glioma definition, the ROI drawings with vessels excluded took approximately twice as long to perform as when the vessels were included[14]. Also, the ROI drawings were easier to perform when the vessels were included. In clinical routine, this feature is attractive as it reduces the time spent on post-processing. Also, as the ROI drawing becomes easier to perform and thus the reproducibility of the method may potentially increase.

The diagnostic accuracy obtained in our study is similar to other studies using the histogram method on gliomas[14,15,28]. Compared to the glioma ROIs with vessels excluded, the sensitivity and NPV of the glioma grading increased and the specificity and PPV decreased when using the glioma ROIs with vessels included. This result suggests that correct manual vessel removal is difficult as the border between tumor- and vessel induced elevations in CBV is diffuse and difficult to distinguish in the static images. Interestingly, the A_z value and log-rank value was higher when using the glioma ROIs with vessels included. As high-grade glioblastomas account for the majority of gliomas in adults[31], this result suggest that a higher diagnostic accuracy may be obtained by using less conservative criteria when drawing the glioma volumes. However, not correcting for the inherent overestimation of nCBV values in low-grade gliomas due to the inclusion of vessels may result in an over-aggressive treatment plan in these patients[12,22].

For the vessel segmentation, the uncorrected vessel masking resulted in reduced sensitivity, NPV and PPV whereas the specificity increased. The lowest A_z value in our study was also obtained when using the uncorrected vessel masking. The increased number of false negatives and reduced number of false positives suggest that the uncorrected vessel masking routine is too radical in removing nCBV values in the tumor bed. Similarly, SI corrected masking also resulted in a lower sensitivity and equal specificity compared to the glioma ROIs with vessels excluded, suggesting to radical segmentation. It should be noted that our DSC sequence includes several dummy scans prior to dynamic acquisition. These are used to avoid rapid signal changes between the first image (with full magnetization) and subsequent (saturated) baseline images. The image contrast in the baseline images is therefore mainly proton density weighted. In spite of the GRE-EPI sequence being strongly T2*-weighted in steady state, the image contrast in the very first image (following a single 90 degree RF pulse) will have a significant T2-weighting. For the purpose of obtaining a SI tumor mask based on differences

in T2 relaxation times, it would therefore be a clear advantage to include this unsaturated EPI image in the analysis. We are in the process of modifying our DSC sequence for this purpose. Nevertheless, our results show that the main challenge of the vessel segmentation routine is to differentiate vessels from elevated nCBV values at the capillary level. As the purpose of vessel segmentation in DSC based glioma grading is to remove false positives without introducing false negatives, the use of a pre-mask to exclude areas of elevated MTT values, secondary to glioma malignancy, was found to improve the overall diagnostic accuracy of the vessel segmentation method. The MTT corrected vessel masking resulted in equal sensitivity to the ROIs with vessels included and the same specificity as the uncorrected vessel mask. Consequently, the biggest difference in mean peak height between glioma grades and the highest diagnostic accuracy values were observed when using the MTT corrected vessel masks. At an individual patient level, the uncertainty of these measures is illustrated by NPV and PPV of which higher values indicate less false negatives and false positives, respectively. Of the 9 gliomas with a change in glioma grade due to use of the MTT corrected vessel mask, the 3 gliomas located within the M1 or M2 segments of the middle cerebral artery distribution were correctly classified. The remaining 6 gliomas, of which 3 were correctly classified and 3 misclassified, were not located in close proximity to one of the main arteries. As confirmed by others, this result suggests that glioma ROI definition in areas close to a prominent artery is difficult[9,10]. In these regions, use of glioma ROIs with vessels included in combination with the vessel segmentation routine should be warranted.

The Kaplan-Meier survival curves with log-rank values for the high- and low-risk groups obtained in our study are similar to previously reported survival curves in similar patient groups using MR perfusion based grading methods[2]. Ideally, the difference between the high- and the low-risk group (log-rank value) should be as large as possible. In our study, the highest log-rank value was obtained when using the MTT corrected vessel masking. As for the logistic regression analysis, a higher log-rank value was obtained for the glioma ROIs with vessel included compared to the glioma ROIs with vessel excluded. This result supports the conclusion that a higher diagnostic efficacy in glioma characterization from manually defined ROIs is obtained when vessels are included.

The results obtained in our study may be limited by low temporal and spatial resolution. High temporal resolution is of particular importance for the proposed method since separation of vessels from vascular tissue is mainly based on small differences in contrast dynamics. In our

study, the temporal resolution was between 1.4 and 1.7 seconds. Although a correlation between the accuracy of the vessel segmentation and temporal resolution was not performed in our study, the temporal resolution should be as short as possible in order to increase the sample steps resulting in optimal separation between the arterial, venous and tissue first-pass curves. Limited spatial resolution results in partial volume effects in the DSC images. The voxel size of the DSC images was 1.80x1.80x5mm^3, potentially resulting in an underestimation of pixels identified as macrovessels. As the size of the capillaries is typically below 1mm, a higher image resolution would be desirable. Also, as shown in the literature[32], local variations in the dynamic DSC curves due to bolus delay or dispersions between tissue areas may confound the vessel segmentation. Hence, we are currently implementing methods for minimizing the effects of bolus dispersion. Further, the number of cluster classes used in our study was based on the number of tissue types in the DSC images. However, this may not be the optimal number of cluster classes with respect to glioma characterization and should be investigated in a future study. Also, both T2- and T1- weighted perfusion imaging have been proposed as alternatives to T2*-based perfusion imaging in the diagnosis of gliomas[27,33,34]. The advantage of both these methods over T2*-weighted methods is reduced sensitivity to susceptibility effects and consequently better delineation of actual vessel containing pixels in the resulting perfusion maps. The disadvantage of both these methods is lower contrast agent sensitivity (effective contrast relaxivity) and limited brain coverage for a given temporal resolution.

In our study, we have shown that a vessel segmentation routine based on automatic k-mean clustering of dynamic first-pass parameters can improve the diagnostic efficacy of DSC imaging for glioma grading. The automated MTT corrected vessel segmented CBV maps provided a higher diagnostic efficacy than a manual method where the tumor ROI was defined by a neuroradiologist. The proposed method is attractive in that it is fast, easy to perform and provides a mask which covers all pixels affected by the intravascular susceptibility effect. In addition, the method can readily be implemented in a clinical radiological setting without transferring images to computers outside the hospital-wide picture archiving and communication system (PACS) and thereby further improving the usefulness of the histogram method.

Abbreviations:

DSC = dynamic susceptibility contrast, MR = magnetic resonance, CBV = cerebral blood volume, LGG = low-grade glioma, AIF = arterial input function, MTT = mean transit time, fmAUC = first moment of the area under the curve, WHO = world health organization, CNS = central nervous system, HGG = high-grade glioma, ROI = region of interest, NPV = negative predictive values, PPV = positive predictive values, ROC = receiver operating characteristic, PACS = picture archiving and communication system

Acknowledgements:

We thank Bard Nedregaard, MD, and Bjorn Tennoe, MD, from the Clinic for Imaging and Intervention, Rikshospitalet University Hospital, Oslo, Norway for selecting white matter areas and David Scheie, MD, The Pathology Clinic, Rikshospitalet University Hospital, Oslo, Norway, for performing the histopathological analysis.

References

1. Covarrubias DJ, Rosen BR, Lev MH. Dynamic magnetic resonance perfusion imaging of brain tumors. Oncologist 2004;9(5):528-537.

2. Law M, Oh S, Babb JS, Wang E, Inglese M, Zagzag D, Knopp EA, Johnson G. Low-grade gliomas: dynamic susceptibility-weighted contrast-enhanced perfusion MR imaging--prediction of patient clinical response. Radiology 2006 Feb;238(2):658-667.

3. Edelman RR, Mattle HP, Atkinson DJ, Hill T, Finn JP, Mayman C, Ronthal M, Hoogewoud HM, Kleefield J. Cerebral blood flow: assessment with dynamic contrast-enhanced T2*-weighted MR imaging at 1.5 T. Radiology 1990 Jul;176(1):211-220.

4. Aronen HJ, Gazit IE, Louis DN, Buchbinder BR, Pardo FS, Weisskoff RM, Harsh GR, Cosgrove GR, Halpern EF, Hochberg FH, Rosen BR. Cerebral blood volume maps of gliomas: comparison with tumor grade and histologic findings. Radiology 1994 Apr;191(1):41-51.

5. Aronen HJ, Perkio J. Dynamic susceptibility contrast MRI of gliomas. Neuroimaging Clin N Am 2002 Nov;12(4):501-523.

6. Knopp EA, Cha S, Johnson G, Mazumdar A, Golfinos JG, Zagzag D, Miller DC, Kelly PJ, Kricheff II. Glial neoplasms: dynamic contrast-enhanced T2*-weighted MR imaging. Radiology 1999 Jun;211(3):791-798.

7. Lev MH, Rosen BR. Clinical applications of intracranial perfusion MR imaging. Neuroimaging Clin N Am 1999 May;9(2):309-331.

8. Lev MH, Ozsunar Y, Henson JW, Rasheed AA, Barest GD, Harsh GR, Fitzek MM, Chiocca EA, Rabinov JD, Csavoy AN, Rosen BR, Hochberg FH, Schaefer PW, Gonzalez RG. Glial tumor grading and outcome prediction using dynamic spin-echo MR susceptibility mapping compared with conventional contrast-enhanced MR: confounding effect of elevated rCBV of oligodendrogliomas [corrected]. AJNR Am J Neuroradiol 2004 Feb;25(2):214-221.

9. Caseiras GB, Thornton JS, Yousry T, Benton C, Rees J, Waldman AD, Jager HR. Inclusion or exclusion of intratumoral vessels in relative cerebral blood volume

characterization in low-grade gliomas: does it make a difference? AJNR Am J Neuroradiol 2008 Jun;29(6):1140-1141.

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

11. Walker DG, Kaye AH. Low grade glial neoplasms. J Clin Neurosci 2003 Jan;10(1):1-13.

12. Talos IF, Zou KH, Ohno-Machado L, Bhagwat JG, Kikinis R, Black PM, Jolesz FA. Supratentorial low-grade glioma resectability: statistical predictive analysis based on anatomic MR features and tumor characteristics. Radiology 2006 May;239(2):506-513.

13. Kleihues P, Cavenee W. Astrocytic tumors & Oligodendroglial tumors and mixed gliomas. *The WHO classification of tumors of the nervous system*. 2 ed. Lyon: International Agency for Research on Cancer; 2000. p 9-70.

14. Emblem KE, Nedregaard B, Nome T, Due-Tonnessen P, Hald JK, Scheie D, Borota OC, Cvancarova M, Bjornerud A. Glioma grading by using histogram analysis of blood volume heterogeneity from MR-derived cerebral blood volume maps. Radiology 2008 Jun;247(3):808-817.

15. Young R, Babb J, Law M, Pollack E, Johnson G. Comparison of region-of-interest analysis with three different histogram analysis methods in the determination of perfusion metrics in patients with brain gliomas. J Magn Reson Imaging 2007 Oct;26(4):1053-1063.

16. Kaus MR, Warfield SK, Nabavi A, Black PM, Jolesz FA, Kikinis R. Automated segmentation of MR images of brain tumors. Radiology 2001 Feb;218(2):586-591.

17. van Osch MJ, Vonken EJ, Bakker CJ, Viergever MA. Correcting partial volume artifacts of the arterial input function in quantitative cerebral perfusion MRI. Magn Reson Med 2001 Mar;45(3):477-485.

18. Mouridsen K, Christensen S, Gyldensted L, Ostergaard L. Automatic selection of arterial input function using cluster analysis. Magn Reson Med 2006 Mar;55(3):524-531.

19. Rausch M, Scheffler K, Rudin M, Radu EW. Analysis of input functions from different arterial branches with gamma variate functions and cluster analysis for quantitative blood volume measurements. Magn Reson Imaging 2000 Dec;18(10):1235-1243.

20. Hadjiprocopis A, Rashid W, Tofts PS. Unbiased segmentation of diffusion-weighted magnetic resonance images of the brain using iterative clustering. Magn Reson Imaging 2005 Oct;23(8):877-885.

21. Ball WS, Jr., Holland SK. Perfusion imaging in the pediatric patient. Magn Reson Imaging Clin N Am 2001 Feb;9(1):207-30, ix.

22. Grier JT, Batchelor T. Low-grade gliomas in adults. Oncologist 2006 Jun;11(6):681-693.

23. Boxerman JL, Schmainda KM, Weisskoff RM. Relative cerebral blood volume maps corrected for contrast agent extravasation significantly correlate with glioma tumor grade, whereas uncorrected maps do not. AJNR Am J Neuroradiol 2006 Apr;27(4):859-867.

24. Maes F, Collignon A, Vandermeulen D, Marchal G, Suetens P. Multimodality image registration by maximization of mutual information. IEEE Trans Med Imaging 1997 Apr;16(2):187-198.

25. Jain RK, di Tomaso E, Duda DG, Loeffler JS, Sorensen AG, Batchelor TT. Angiogenesis in brain tumours. Nat Rev Neurosci 2007 Aug;8(8):610-622.

26. Bastin ME, Carpenter TK, Armitage PA, Sinha S, Wardlaw JM, Whittle IR. Effects of dexamethasone on cerebral perfusion and water diffusion in patients with high-grade glioma. AJNR Am J Neuroradiol 2006 Feb;27(2):402-408.

27. Schmainda KM, Rand SD, Joseph AM, Lund R, Ward BD, Pathak AP, Ulmer JL, Badruddoja MA, Krouwer HG. Characterization of a first-pass gradient-echo spin-echo

method to predict brain tumor grade and angiogenesis. AJNR Am J Neuroradiol 2004 Oct;25(9):1524-1532.

28. Law M, Young R, Babb J, Pollack E, Johnson G. Histogram analysis versus region of interest analysis of dynamic susceptibility contrast perfusion MR imaging data in the grading of cerebral gliomas. AJNR Am J Neuroradiol 2007 Apr;28(4):761-766.

29. Boxerman JL, Hamberg LM, Rosen BR, Weisskoff RM. MR contrast due to intravascular magnetic susceptibility perturbations. Magn Reson Med 1995 Oct;34(4):555-566.

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

31. Hess KR, Broglio KR, Bondy ML. Adult glioma incidence trends in the United States, 1977-2000. Cancer 2004 Nov;101(10):2293-2299.

32. Calamante F. Bolus dispersion issues related to the quantification of perfusion MRI data. J Magn Reson Imaging 2005 Dec;22(6):718-722.

33. Cha S, Yang L, Johnson G, Lai A, Chen MH, Tihan T, Wendland M, Dillon WP. Comparison of microvascular permeability measurements, K(trans), determined with conventional steady-state T1-weighted and first-pass T2*-weighted MR imaging methods in gliomas and meningiomas. AJNR Am J Neuroradiol 2006 Feb;27(2):409-417.

34. Haroon HA, Buckley DL, Patankar TA, Dow GR, Rutherford SA, Baleriaux D, Jackson A. A comparison of Ktrans measurements obtained with conventional and first pass pharmacokinetic models in human gliomas. J Magn Reson Imaging 2004 May;19(5):527-536.

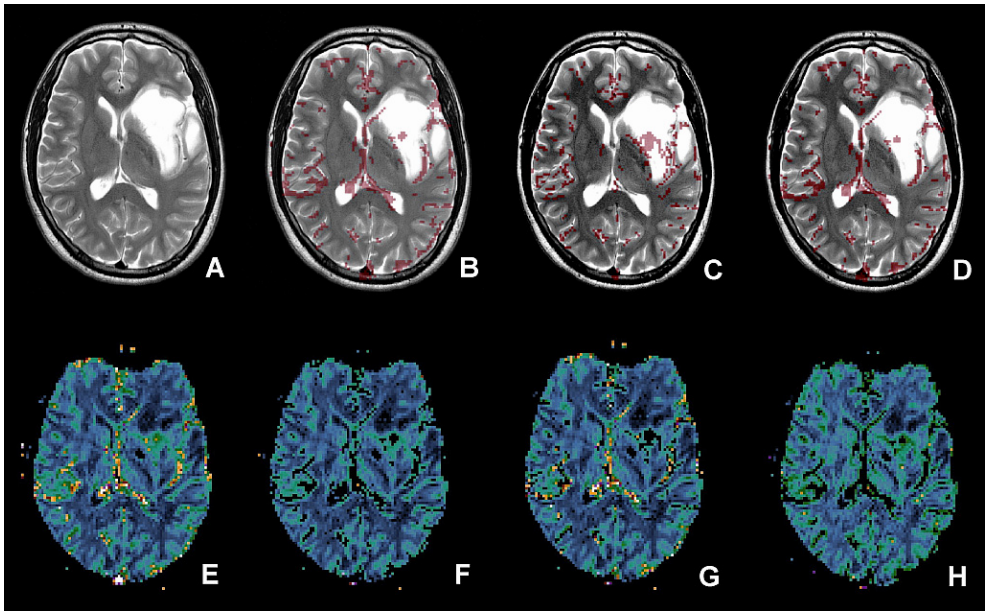
Table 1: Diagnostic accuracy values for the glioma grading using the nCBV maps with and without vessel segmentation. Absolute number of patients is shown in parentheses.

	Sensitivity	Specificity	NPV	PPV
	% (HGGs)	**%** (LGGs)	**%** (TN/TN+FN)	**%** (TP/TP+FP)
nCBV maps (-)	**87** (39/45)	**78** (25/32 LGGs)	**81** (25/31)	**85** (39/46)
nCBV maps (+)	**91** (41/45)	**72** (23/32 LGGs)	**85** (23/27)	**82** (41/50)
Vessel mask	**82** (37/45)	**81** (26/32 LGGs)	**77** (26/34)	**86** (37/43)
MTT	**91** (41/45)	**81** (26/32 LGGs)	**87** (26/30)	**87** (41/47)
SI	**82** (37/45)	**78** (25/32 LGGs)	**76** (25/33)	**84** (37/44)

NPV = Negative Predictive Value, PPV = Positive Predictive Value, HGG = high-grade gliomas, LGG = low-grade gliomas,
TN = true negatives, TP = true positives, FN = false negatives, FP = false positives
nCBV maps(-) = vessels removed in the glioma ROIs, nCBV maps(+) = vessels included in the glioma ROIs,
Vessel mask = uncorrected vessel segmentation, MTT = MTT corrected vessel segmentation, SI = SI corrected vessel segmentation

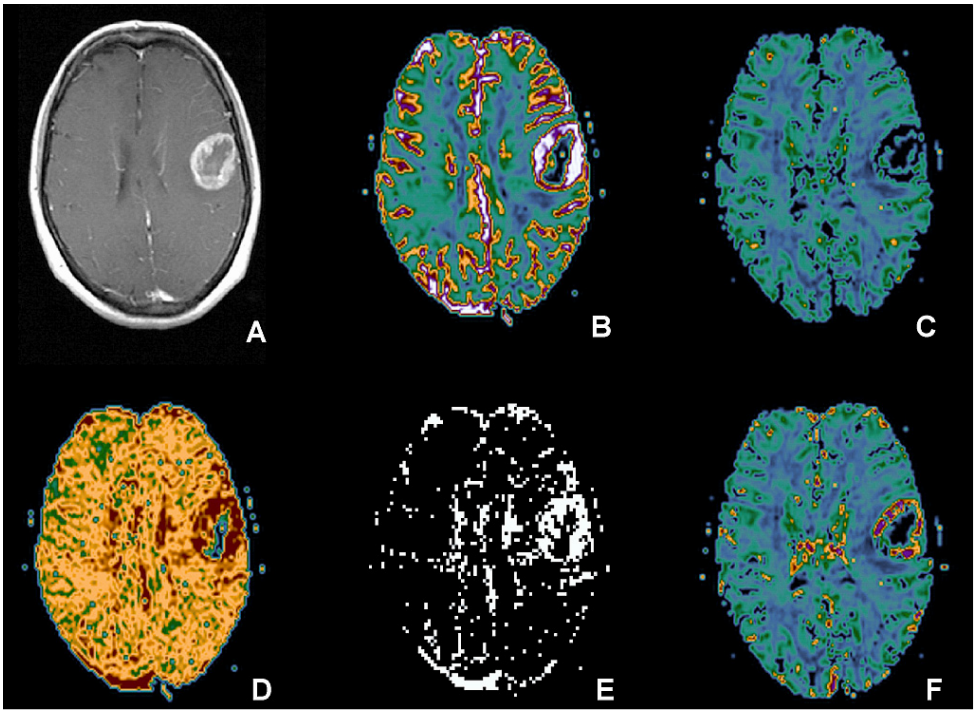
FIGURE CAPTIONS:

FIGURE 1



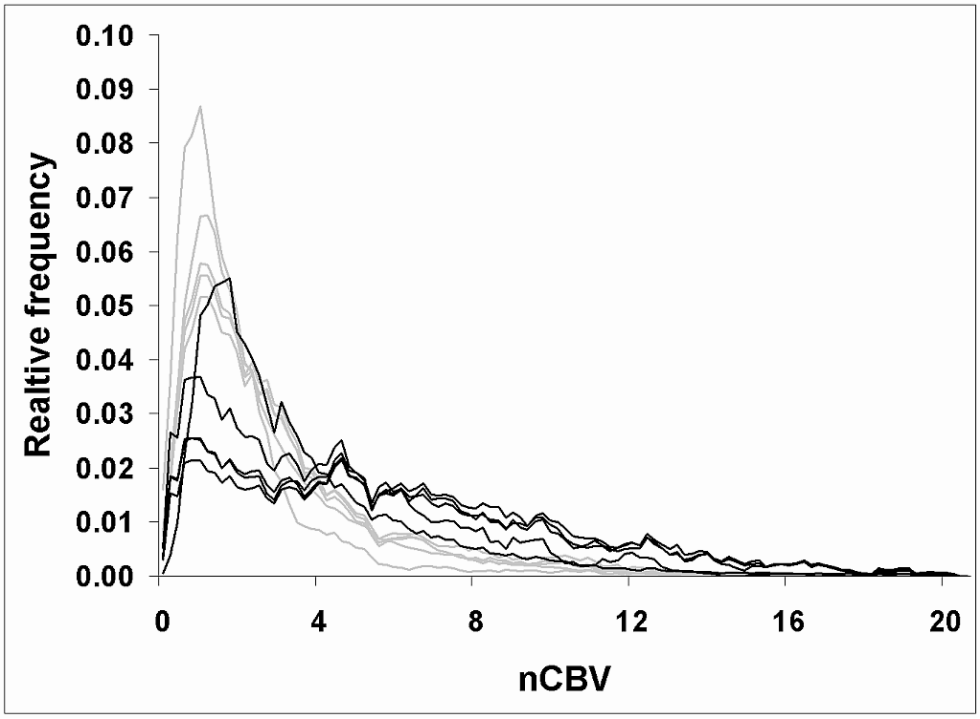
(A) Axial T2-weighted fast spin-echo image (TR=4000ms/TE=104ms) of a patient with a low-grade diffuse astrocytoma (WHO grade II). **(B-D)** Binary vessel mask overlaid on the T2-weighted image using **(B)** uncorrected vessel masks, **(C)** MTT corrected vessel masks and **(D)** binary SI corrected masks. **(E)** nCBV map with vessel included, coregistered to the T2-weighted image. **(F-H)** nCBV map with vessels removed using **(F)** uncorrected vessel masks, **(G)** MTT corrected vessel masks and **(H)** SI corrected masks. Note the identification of the main arteries and veins in **(B-D)**, also recognized on the structural T2-weighted image **(A)**. Compared to the uncorrected **(F)** and SI corrected **(H)** vessel masks, the MTT corrected vessel masks **(G)** are generally more conservative in masking out macroscopic vessels in normal appearing tissue regions.

FIGURE 2



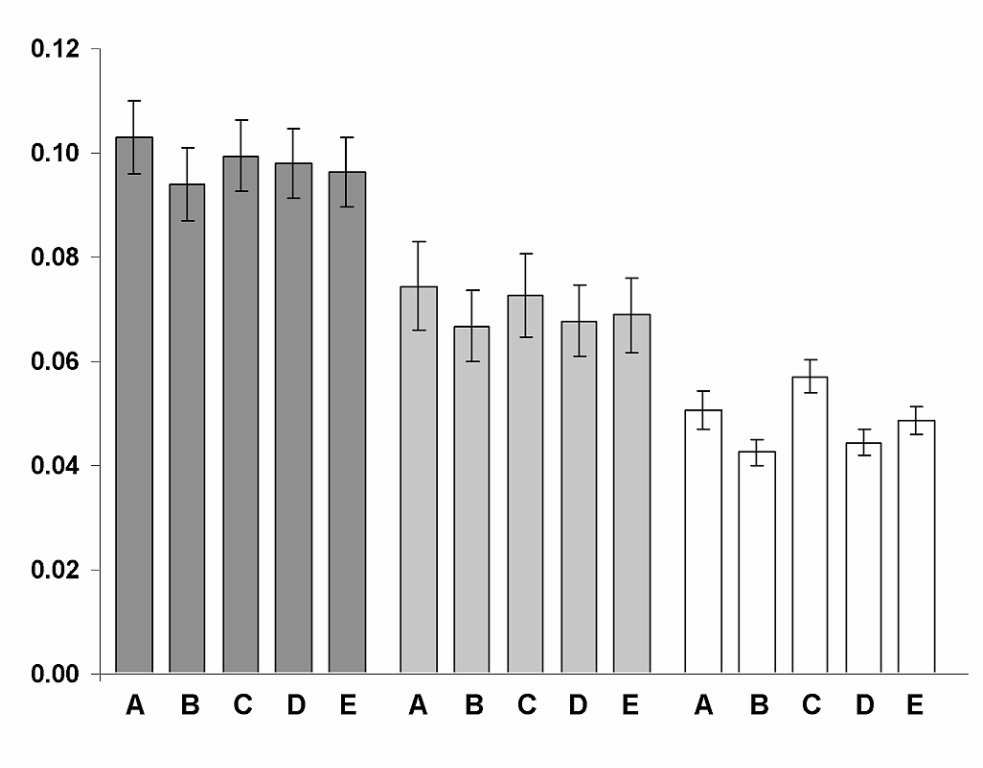
(A) Axial post-contrast T1-weighted spin-echo image (TR=500ms/TE=7.7ms) of a patient with a high-grade glioblastoma (WHO grade IV). **(B)** Corresponding nCBV map with vessel included. **(C)** nCBV map with vessels corresponding to the uncorrected vessel mask removed. **(D)** rMTT map with vessels included. **(E)** Binary rMTT cluster mask representative of the cluster class with the highest rMTT values. In the MTT corrected vessel masks, this cluster class was excluded from the vessel segmentation routine. **(F)** nCBV map with vessels corresponding to the MTT corrected vessel mask removed. Note the difference between the uncorrected- and MTT corrected vessel masks in **(C)** and **(F)**.

FIGURE 3



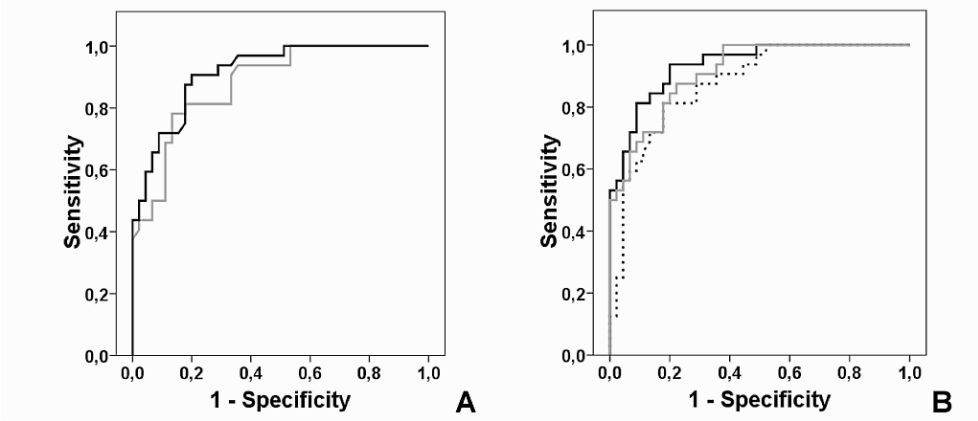
Resulting normalized histograms for the LGG (grey lines) and HGG (black lines) shown in Figures 1 and 2, respectively. The figure show histograms using nCBV maps with vessels included (LGG peak height=.0623, HGG peak height=.0305), manually excluded (.0517, .0214), excluded by uncorrected vessel masks (.0555, .0367), excluded by MTT corrected vessel masks (.087, .0254) and excluded by SI corrected vessel masks (.0578, .0255).

FIGURE 4



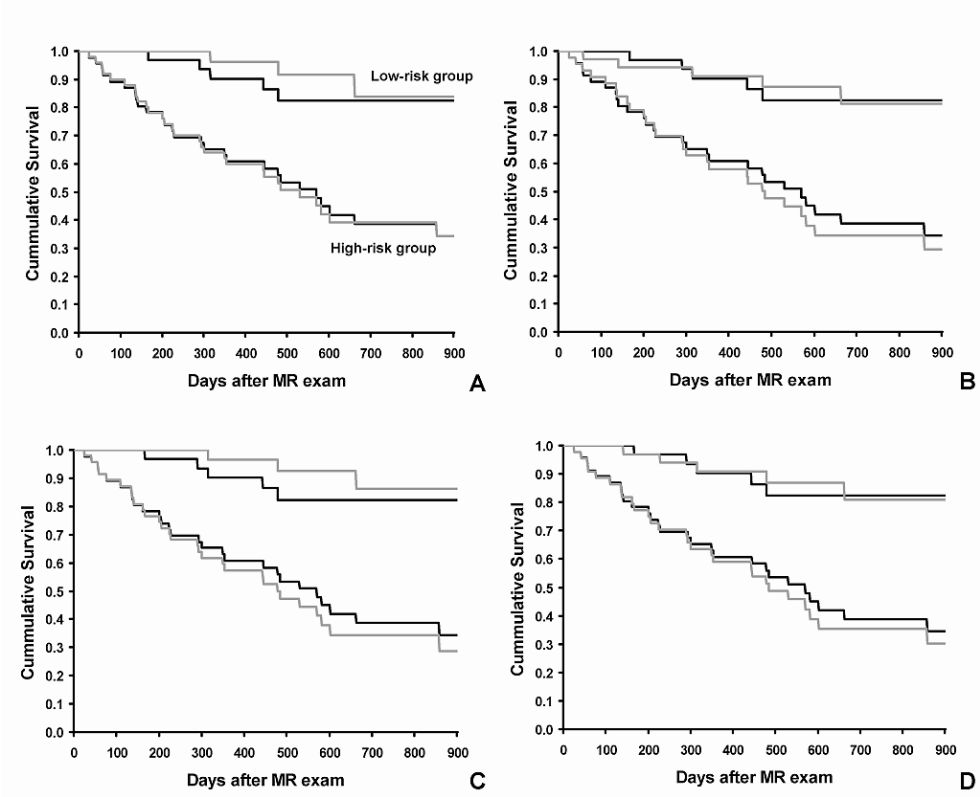
Mean histogram peak heights with standard error bars for glioma grade II (dark grey bars; 32 patients), grade III (light grey bars; 10 patients) and grade IV (white bars; 35 patients) when using **(A)** nCBV maps with vessels manually excluded, **(B)** nCBV maps with vessels included, **(C)** nCBV maps with vessels removed by the uncorrected vessel masks, **(D)** nCBV maps with vessels removed by the MTT corrected vessel masks and **(E)** nCBV maps with vessels removed by the SI corrected vessel masks, respectively.

FIGURE 5



(A) Receiver operator characteristics (ROC) curves for the histogram method when using the glioma ROIs with vessels manually excluded (gray line) and with vessels included (black line). The areas (± standard errors) under the ROC curve were .881 (±.038) and .915 (±.030), respectively. **(B)** ROC curves for the histogram method when using the vessels segmentation routine on the glioma ROIs with vessels included. The ROC curves represent the uncorrected vessel masking (dotted line), MTT corrected vessel masking (black line) and SI corrected vessel masking (gray line). The areas (± standard errors) under the ROC curve were .876 (±.039), .935 (±.026) and .912 (±.031), respectively.

FIGURE 6



Kaplan-Meier survival curves for the different methods investigated. Based on available survival data in all 77 patients, logistic regression was used to derive optimal histogram peak height cut-off values between a high- and a low-risk patient group. The figures show survival curves for the reference histogram method using glioma ROIs with vessels removed by the radiologist (black lines) versus survival curves for the (gray lines); **(A)** glioma ROIs with vessels included, **(B)** uncorrected vessel masks, **(C)** MTT corrected vessel masks and **(D)** SI corrected vessel masks. The highest log-rank value between the high- and low-risk group was observed when using the MTT corrected vessel masks (20.390, $P<.001$).

IV

Magnetic Resonance in Medicine 60:945–952 (2008)

Predictive Modeling in Glioma Grading From MR Perfusion Images Using Support Vector Machines

Kyrre E. Emblem,[1,6*] Frank G. Zoellner,[4] Bjorn Tennoe,[2] Baard Nedregaard,[2] Terje Nome,[2] Paulina Due-Tonnessen,[2] John K. Hald,[2] David Scheie,[3] and Atle Bjornerud[1,5]

The advantages of predictive modeling in glioma grading from MR perfusion images have not yet been explored. The aim of the current study was to implement a predictive model based on support vector machines (SVM) for glioma grading using tumor blood volume histogram signatures derived from MR perfusion images and to assess the diagnostic accuracy of the model and the sensitivity to sample size. A total of 86 patients with histologically-confirmed gliomas were imaged using dynamic susceptibility contrast (DSC) MRI at 1.5T. Histogram signatures from 53 of the 86 patients were analyzed independently by four neuroradiologists and used as a basis for the predictive SVM model. The resulting SVM model was tested on the remaining 33 patients and analyzed by a fifth neuroradiologist. At optimal SVM parameters, the true positive rate (TPR) and true negative rate (TNR) of the SVM model on the 33 patients was 0.76 and 0.82, respectively. The interobserver agreement and the TPR increased significantly when the SVM model was based on an increasing sample size ($P < 0.001$). This result suggests that a predictive SVM model can aid in the diagnosis of glioma grade from MR perfusion images and that the model improves with increasing sample size. Magn Reson Med 60: 945–952, 2008. © 2008 Wiley-Liss, Inc.

Key words: DSC MRI; histogram analysis; glioma grading; support vector machines; predictive modeling

Several studies have shown that cerebral blood volume (CBV) maps derived from dynamic susceptibility contrast (DSC) MRI can improve differentiation between high-grade (grades III–IV) and low-grade (grades I–II) gliomas, using the World Health Organization (WHO) classification system (1–4). Based on normalized CBV (nCBV) maps, viable malignant tumor tissue can be identified as regions of elevated microvascular blood volume (maximum nCBV; "hotspot method") compared to unaffected tissue (5). However, applying these grading methods prospectively requires some prior knowledge about the appropriate threshold values that provide optimal differentiation between high- and low-grade gliomas. In studies using the hotspot method on gradient-echo perfusion images, the reported maximum nCBV threshold for optimal differentiation show large variations (6–8). This suggests that the optimal nCBV threshold depends on several method-specific parameters, including contrast agent properties and dose, imaging technique, and postprocessing routines. This method-dependency on the critical nCBV threshold means that the threshold value must be determined specifically at each site, which complicates the comparison of data between sites and also places restrictions on modification of any of the model-sensitive parameters in a given institution.

Although common in the literature on tumor growth and invasion (9–11), application of predictive modeling to MR perfusion images with the aim of predicting glioma grade is, to our knowledge, not reported in the literature. One possible reason for this is that current nCBV threshold values are difficult to generalize into a useful model. A predictive model based on one single value per subject will inherently lack sufficient robustness. An alternative histogram-based method for analysis of nCBV maps has recently been proposed, which provides a measure of the distribution of nCBV values in the entire volume affected by the tumor; it has been shown that this approach may improve differentiation between low- and high-grade gliomas compared to the hotspot method (8,12). Histogram analysis also provides a more attractive starting point for predictive modeling since each new case is now described by a histogram "signature" curve of the nCBV distribution within the tumor volume, which can readily be compared to a database of histologically-confirmed gliomas with signatures representative for each glioma grade.

From this, we hypothesize that a predictive model based on support vector machines (SVM) can aid in the diagnosis of gliomas. Thus, the purpose of our study was to implement a predictive model based on SVM for glioma grading using tumor blood volume histogram signatures derived from MR perfusion images. Also, we hypothesize that the diagnostic accuracy of the SVM model will improve with sample size. Thus, a second purpose was to assess the diagnostic accuracy (13) of the SVM model and the sensitivity to sample size.

THEORY

SVMs are generalized linear classification algorithms that can be used to create a model for classification of objects based on a set of training data (14,15). An SVM model

[1]Department of Medical Physics, Rikshospitalet University Hospital, Oslo, Norway.

[2]Department of Radiology, Rikshospitalet University Hospital, Oslo, Norway.

[3]Department of Pathology, Rikshospitalet University Hospital, Oslo, Norway.

[4]Department of Computer Assisted Clinical Medicine, Faculty of Medicine Mannheim, University of Heidelberg, Mannheim, Germany.

[5]Department of Physics, University of Oslo, Oslo, Norway.

[6]The Interventional Center, Rikshospitalet University Hospital, Oslo, Norway.

Abbreviations: SVM = support vector machines; CBV = cerebral blood volume; DSC = dynamic susceptibility-contrast; MRI = magnetic resonance imaging; WHO = World Health Organization; RBF = radial basis function; ROI = region of interest; TPR = true-positive rate; TNR = true-negative rate.

*Correspondence to: Kyrre E. Emblem, The Interventional Center, Gaustad, Rikshospitalet University Hospital, Sognsvannsveien 20, N-0027 Oslo, Norway. E-mail: kyrre.eeg.emblem@rikshospitalet.no

Received 28 March 2008; revised 21 May 2008; accepted 30 May 2008.

DOI 10.1002/mrm.21736

Published online in Wiley InterScience (www.interscience.wiley.com).

creates a separating hyperplane, a higher-dimensional generalization, so that it optimally discriminates between two or more classes. During a minimization procedure (i.e., a learning process), the hyperplane is tuned so that the SVM model generalization error is minimized, thus achieving an optimal solution to the classification problem. In linear classifications, a binary decision is performed by a real-valued function:

$$f{:}X \subseteq R^n \rightarrow R. \quad [1]$$

The input $\vec{x} = (x_1, \ldots, x_n) \in X$ is assigned a positive class if $f(\vec{x}) > 0$ and a negative class otherwise. In this case f is a linear function and can be written as:

$$f(\vec{x}) = \langle \vec{w}, \vec{x} \rangle + \vec{b}. \quad [2]$$

Here, $f(\vec{x})$ defines a hyperplane with parameters $\vec{w}$ (the direction perpendicular to the hyperplane) and $\vec{b}$ (position vector). The SVM model is trained by a training dataset (of size M) and divides the input samples into a predefined number of classes. The training dataset can be written as $\{(\vec{x}_i, \vec{y}_i)\} i = 1, \ldots, M \text{ with } y_i \in \{-1,1\}$. Classification performed this way may be difficult, since real-world data are noisy, which limits the accuracy with which the resulting hyperplane can be determined in feature space (14). One approach to address this problem is to transform the original input space into a higher-dimensional space. Usually, a high-dimensional space is sparse. Mapping the input space into a higher-dimensional space thus simplifies the identification of separating hyperplanes. To map the data, so-called "kernel" functions are determined so that the hyperplane optimally discriminates between the two classes. Using a nonlinear kernel function ϕ that maps the input $\vec{x} \in R^n$ to a higher-dimensional space R^N, n $\ll$ N, where the input is linearly separable, the function separating the input can be written as:

$$f(\vec{x}) = \vec{w}\phi(\vec{x}) + \vec{b}. \quad [3]$$

Usually, a special transformation function (ϕ) to build a kernel can not be readily obtained. Instead already-known kernels are used. The choice of kernel function has to be done carefully to avoid poor discrimination power (16). In addition, to be able to handle a large dataset and to reduce processing time, this transformation is only implicitly executed; this is also known as the "kernel trick" (15). Maximal reduction of the generalization error is reached by optimizing the distance between the margin of the function separating the classes (functional margin) and the input examples during a training process. If $\vec{x}_i$ is classified correctly, the functional margin $\tilde{y}_i$ of the training example $(\vec{x}_i, y_i)$ is defined as:

$$\tilde{y}_i = y_i(\vec{w}^T\vec{x}_i + b) \Rightarrow \tilde{y}_i > 0. \quad [4]$$

Further, $\tilde{y}_i$ can be transformed into a so-called geometric margin:

$$y_i = \frac{\tilde{y}_i}{\|w\|}. \quad [5]$$

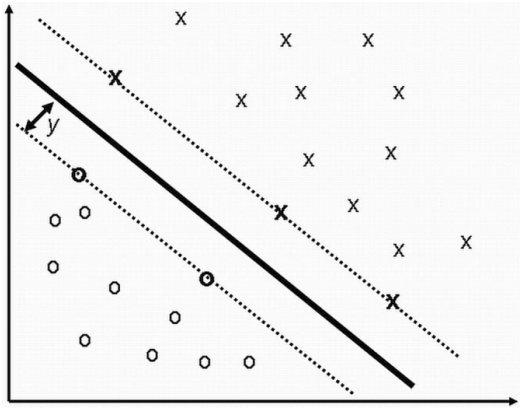


FIG. 1. Illustration of the optimal margins (dotted lines) and hyperplane y (solid line) for arbitrary binary data. The data points on the optimal margins are called support vectors, and define the position of the margins. When searching for the most generalized model, the distance between the hyperplane and the margins should be as large as possible.

Here, an optimization procedure (15) can be applied to find the optimal hyperplane. Figure 1 depicts an illustration of the optimal margin (dotted lines) and hyperplane (solid line) for a two-dimensional (2D) binary example. The data points defining the margins are called support vectors and their positions should be such as to maximize the distance between the support vectors and the hyperplane (y).

Although numerous SVM models exists (16), an SVM model known as υ-SVM is considered practical for our data (17,18). The υ-SVM has the advantage of using a parameter, υ, which is related to the number of support vectors used and the ratio of the training error. The range of the υ-parameters is between 0 and 1. A low υ value results in less softness of the classification margins (i.e., less generalized) but fewer misclassifications, whereas a high υ value result in more softness of the classification margins but with more misclassifications allowed. The optimal υ-SVM model has a υ value giving minimal misclassifications and maximal generalization.

If the input data are not linearly separable as in Eq. [3], a nonlinear kernel has to be applied. A study using SVM on histogram-based color image classifications suggested using a nonlinear radial basis function (RBF) kernel when analyzing low-dimension histogram data as in our study (19):

$$K(x_i, x_y) = e^{-\gamma\|x_i - x_j\|^2}. \quad [6]$$

Here, a Gaussian-based RBF kernel is used. The parameter γ is related to the standard deviation (SD); i.e., the width of the Gaussian distribution. By default, the γ parameter is defined as 1/M, where M is the number of samples in the training data. To illustrate the effects of changing the two parameters (υ,γ), Fig. 2 depicts a simulation on arbitrary data for a υ-SVM model at different settings of υ and γ.

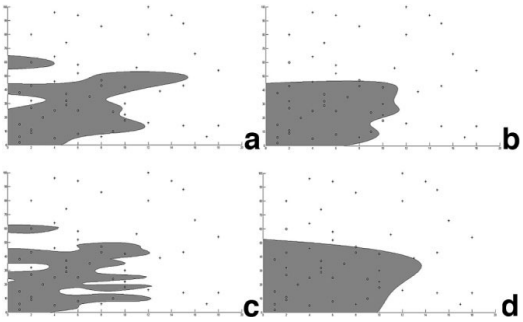


FIG. 2. Simulation of a υ-SVM model with a Gaussian RBF kernel on two-class ("+" and "o") arbitrary data, using different values of υ and γ. The parameter values are as follows: (**a**) υ = 0.1 and γ = 0.02; (**b**) υ = 0.9 and γ = 0.02; (**c**) υ = 0.5 and γ = 0.2; and (**d**) υ = 0.5 and γ = 0.002. In (a) and (b) the γ parameter is kept constant at default value (0.02 ≈1/53; 53 patients in training dataset), whereas in (c) and (d) the υ parameter is kept constant at a default value (υ = 0.5). Increasing the υ parameter results in softer margins (i.e., higher generalization) at the cost of more misclassifications. Reducing the γ parameter, affects the standard deviation in the Gaussian distribution, resulting in a wider distribution and softer margins.

MATERIALS AND METHODS

Patient Selection

Study approval was obtained from the Regional Medical Ethics Committee and patients were included only if informed consent was signed. Between July 2005 and November 2007, 86 patients (age = 9–78 years, mean age = 49 years; 49 males, 37 females) received a histological diagnosis of primary glioma after MR perfusion imaging and subsequent surgery. The histological evaluation was performed by a neuropathologist using the WHO classification of central nervous system tumors (20).

Observers

All image measurements were performed independently by five experienced neuroradiologists with several years of experience reading brain MR perfusion images. Patient-related information was removed from all images, and the observers were blinded to the histopathological diagnosis.

MRI

Imaging was performed at 1.5T (Siemens Sonata, Symphony, or Avanto; Siemens AG, Erlangen, Germany), using an eight-channel (Symphony/Sonata) or a 12-channel (Avanto) head coil. The protocol included an axial T_2-weighted fast spin-echo sequence with repetition time/echo time (TR/TE = 4000 ms/104 ms) and an axial T_1-weighted spin-echo sequence (TR/TE = 500 ms/7.7 ms) obtained before and after intravenous (i.v.) contrast agent injection. The voxel size = 0.45 × 0.45 × 5 mm^3 with 19 slices in both sequences. DSC MRI was performed using a gradient-echo echo-planar imaging sequence acquired during contrast agent administration. The imaging parameters were: TR/TE = 1430 ms/46 ms (12 axial slices) to TR/TE = 1590 ms/52 ms (14 axial slices), bandwidth = 1345 Hz/pixel, field of view = 230 × 230 mm, voxel size = 1.80 × 1.80 × 5 mm^3, and interslice gap = 1.5 mm. For each slice, 50 images were recorded at intervals equal to the TR. After approximately eight time points, 0.2 mmol/kg of the high-concentration (1 M) gadolinium-based contrast agent gadobutrol (Bayer Schering Pharma AG, Berlin, Germany), was injected at a rate of 5 ml/s, immediately followed by a 20-ml bolus of saline (B. Braun Melsungen AG, Melsungen, Germany) also at 5 ml/s. Gadobutrol has been shown to exhibit very similar physicochemical properties to other low–molecular-weight gadolinium chelates that have been used for MR perfusion imaging (21).

Image Postprocessing

The images were postprocessed using a dedicated software package (nordicICE; NordicImagingLab, Bergen, Norway). Relative CBV (rCBV) maps were generated using established tracer kinetic models applied to the first-pass data (22,23) and corrected for possible extravascular contrast agent leakage (24). nCBV maps were created by voxel-wise division of the rCBV values with an unaffected slice-specific white matter rCBV value defined by two observers (5). The nCBV maps were displayed using a black-blue-green-yellow-red-white color map, and automatically coregistered with the conventional MR images using a normalized mutual information algorithm (25). Four of the five observers drew tumor regions of interest (ROIs) in 53 of the patients, whereas the fifth observer drew tumor ROIs in the remaining 33 patients. To conform to the histogram method, the five observers drew freehand tumor ROIs in each slice according to the combined underlay/overlay image information as described elsewhere (8,24,26). The observers were asked to avoid areas of necrosis, cysts, or nontumor macrovessels readily visible on the postcontrast T_1-weighted images and the CBV images. Although gliomas are known to infiltrate brain tissue beyond the radiographic margins (27), signal hyperintensity thought to represent tumor tissue and edema as seen on the T_2-weighted images was used to define the outermost tumor margins. Also, it has been shown that potentially imperfect tumor volume definition has little influence on the reproducibility of the glioma grading when using the histogram method (8).

The histogram grading method was performed as follows: histogram signatures were generated for each patient using nCBV values from the complete tumor ROIs. The range of the nCBV values along the *x*-axis was kept constant (values = 0–20) and the area under each histogram curve was normalized to 1. The histograms were divided into 108 bins, a number previously shown to give the maximal diagnostic accuracy (8).

Predictive Modeling by SVMs

The four sets of 53 histogram signatures from the four observers were used to create a dataset: the "training dataset". The remaining 33 histogram signatures from the fifth observer were used to create another dataset: "test dataset". To compare our results with previous studies on glioma grading, a preliminary υ-SVM model was derived

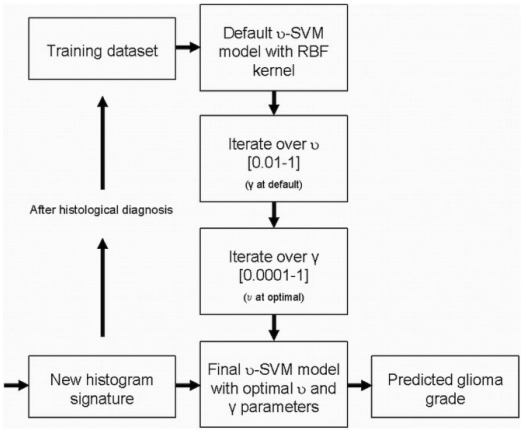


FIG. 3. Schematic flow-diagram of the υ-SVM model design used in our study. Based on the histogram signatures in the training dataset, the υ and γ parameters giving optimal diagnostic accuracy are obtained by iteration. Then, the model is applied on the new histogram signature for prediction of tumor grade. After a histological diagnosis is obtained, the new histogram signature is added to the training dataset, thereby continuously increasing the size of the training dataset.

with the only goal of optimally differentiating between high-grade and low-grade gliomas in the training dataset for each observer individually. Using histology as a reference, this υ-SVM model was derived by varying the values of υ and γ between 0.01 and 1 and between 0.00001 and 1, respectively.

The second υ-SVM model, used to predict glioma grades in the test dataset, was created as follows: The diagnostic accuracy of the model was derived by performing a 10-fold cross validation on the training dataset (repeated testing of the diagnostic accuracy of the model on 10 random selections based on the remaining 43 samples). First, the value of υ was varied between 0.01 and 1. Because the model is more sensitive to the υ parameter than the γ parameter of the RBF kernel (16–18), the γ value at this point was kept constant at default value (0.018 = 1/53; 53 patients in training dataset). To reduce observer specific variations, the final accuracy value for each value of υ was taken as the mean accuracy value across the four observers. The optimal υ value was assumed to be the υ value providing the maximal accuracy value of the cross validation. Using this optimal υ value, the value of γ was then varied between 0.00001 and 1. The optimal γ value was assumed to be the γ value providing the maximal accuracy value of the cross validation. A schematic flow-diagram of the υ-SVM model design is shown in Fig. 3. For each value of υ and γ, the υ-SVM model was applied on the test dataset to evaluate the accuracy of the model when predicting glioma grades in new patients. Also, to remove random spikes in the accuracy values from the cross validation, a low-pass filter with a filter coefficient equal to the reciprocal of the span (1/53) was applied to the curve of the accuracy values as a function of υ and γ, respectively.

In addition, an iterative algorithm was created to evaluate whether the υ-SVM model improved with the number of histogram signatures (i.e., data samples) included in the training dataset. The sample size in the training dataset was varied between five histogram signatures (randomly sampled) up to all 53 signatures, adding one signature at a time. To exclude selection bias, a new random sample of signatures was selected for each step and for each observer. At each step, the υ-SVM model design as described above was derived from the training dataset used on the full test dataset to predict glioma grades.

Statistical Analysis

The diagnostic accuracy (13) of the υ-SVM model on the training dataset was evaluated by assessing the ratio of gliomas in the training dataset that were correctly graded as either high-grade or low-grade by the υ-SVM model. The diagnostic accuracy of the υ-SVM model on the test dataset was evaluated by assessing true-positive rates (TPR) and true-negative rates (TNR). A glioma classified by both histopathology and the υ-SVM model as high-grade/low-grade was considered a true-positive/true-negative finding, respectively. Final TPR and TNR values for the different combinations of υ and γ were averaged across the four observers.

Based on the optimal υ and γ parameter values only, the effect of increasing the sample size (including more patients) of the training dataset was evaluated using linear regression. Here, mean TPR and TNR of the υ-SVM model on the test dataset across the four observers was derived at each sample size. Also, interobserver agreement values between the four observers were assessed at each sample size using Fleiss-Kappa (κ) statistics (28). Statistical analysis was performed using Matlab R2007a (The MathWorks, Natick, MA, USA), SPSS 13 (SPSS Inc., Chicago, IL, USA), and Minitab 15 (Minitab, State College, PA, USA).

RESULTS

Of the 53 patients in the training dataset, 29 received a histological diagnosis of a high-grade glioma (WHO grade III–IV), and 24 patients received a diagnosis of a low-grade glioma (WHO grade I–II). Of the 33 patients in the test dataset, 19 received a histological diagnosis of high-grade glioma, and 14 patients received a diagnosis of low-grade glioma. Of the total 38 low-grade gliomas, there were four grade-I pilocytic astrocytomas, one grade-II gemistocytic astrocytoma, 18 grade-II diffuse astrocytomas, five grade-II oligodendrogliomas, and 10 grade-II oligoastrocytomas. Of the total 48 high-grade gliomas, there were four grade-III anaplastic astrocytomas, three grade-III anaplastic oligodendrogliomas, four grade-III anaplastic oligoastrocytomas, and 37 grade-IV glioblastomas.

Figure 4 shows sample nCBV maps overlaid on axial T_2-weighted images for two patients diagnosed with a low-grade glioma and a high-grade glioma, respectively. The corresponding histogram signatures were derived from the complete distribution of nCBV values from the total tumor volume as defined as by the observers. Typically, the normalized nCBV distribution of a low-grade glioma has a

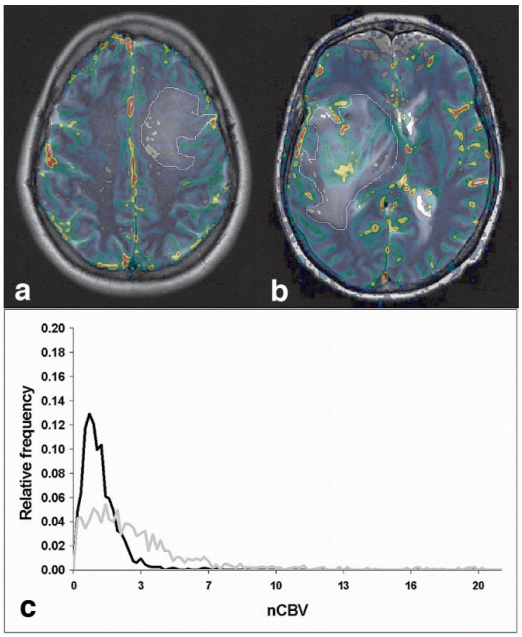


FIG. 4. Coregistered nCBV maps overlaid on axial T_2-weighted fast spin-echo images (repetition time [TR]= 4000 ms/echo time [TE] = 104 ms) of a patient with a low-grade (WHO grade-II) diffuse astrocytoma (**a**) and a patient with a high-grade (WHO grade-IV) glioblastoma (**b**). A tumor area (white ROIs) was defined by the observers in each tumor slice. **c:** Final histogram signatures of the complete distribution of nCBV values from the total tumor volumes. The black line represents the histogram signature of low-grade glioma in (a), whereas the gray line represents the histogram signature of the high-grade in (b).

high peak and narrow distribution compared to the lower peak and broader distribution of a high-grade glioma.

The υ-SVM Model

The entire procedure of generating an optimized υ-SVM model based on a training dataset with 53 patients took approximately 1 min (standard desktop computer with Pentium 4 CPU, 3.19 GHz, and 3 GB RAM). Using $\upsilon = 0.07-0.09$ and γ at the default value (1/53), the υ-SVM model differentiated between high-grade and low-grade gliomas in the training dataset with an accuracy of 100% (53/53) for all four observers.

For the predictive υ-SVM model, the maximal filtered accuracy value (83.96%) of the cross-validation was found at $\upsilon = 0.31$ when the γ parameter was kept at the default value. Figure 5 shows the mean diagnostic accuracy values across the four observers for the υ-SVM model as a function of the υ parameter, keeping γ at default value. The boundary conditions of the υ-SVM model did not allow a υ value above 0.9. The TPR and TNR of the υ-SVM model on the test dataset were 0.64 and 0.76 when using $\upsilon = 0.31$. Keeping υ at 0.31, the maximal filtered accuracy value (83.96%) of the cross-validation was found at $\gamma = 0.02$. Here, the TPR and TNR of the υ-SVM model on the test dataset remained unchanged, at 0.64 and 0.76, respectively.

However, for the υ iteration, using the highest possible υ value (0.56) providing accuracy values of the cross-validation within 99% of the maximal filtered accuracy value resulted in TPR = 0.75 and TNR = 0.81 when using the υ-SVM model on the test dataset. Further, for $\upsilon = 0.56$, using the lowest possible γ value (0.0065) providing accuracy values of the cross-validation within 99% of the maximal filtered accuracy value increased the TPR to 0.76 and the TNR to 0.82 when using the υ-SVM model on the test dataset. Figure 6 shows the mean diagnostic accuracy values across the four observers for the υ-SVM model as a function of γ, keeping υ at 0.56.

Effect of Increasing the Training Dataset Sample Size

Figure 7 shows the Fleiss-Kappa agreement values between the four observers as a function of patients included in the training dataset. A significant increase ($R^2 = 0.46$, $P < 0.001$) in the agreement values was observed from five patients ($\kappa = 0.139$) up to 53 patients ($\kappa = 0.766$). Using optimal υ and γ values only, Fig. 8 shows the effect of increasing the number of patients in the training dataset on the mean TPR and TNR values when using the υ-SVM model on the test dataset. Whereas TPR increased significantly with sample size ($R^2 = 0.311$, $P < 0.001$), no significant sample size dependence was observed for TNR ($R^2 = 0.004$, $P = 0.692$).

DISCUSSION

In this study we propose a threshold-independent approach for prospective glioma grading using the υ-SVM

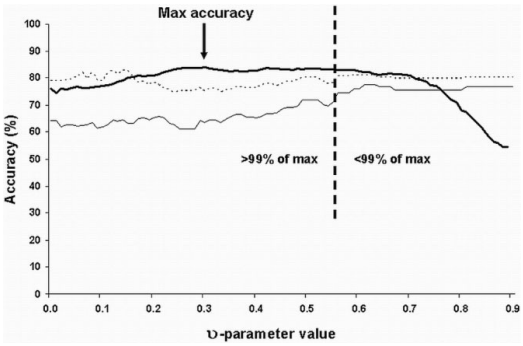


FIG. 5. Mean diagnostic accuracy values across the four observers for the υ-SVM model as a function of changing the υ parameter from 0 to 1 in steps of 0.01, keeping the γ parameter at default value of 1/53). The boundary conditions of the υ-SVM model did not allow a υ value above 0.9. Based on the resulting low-pass-filtered accuracy values (percent of gliomas graded correctly) in the training dataset (bold solid line), the maximal accuracy value (83.96%) was found at $\upsilon = 0.31$. Using the highest υ value (0.56) providing accuracy values within 99% (>83.12%) of the maximal filtered accuracy value resulted in higher TPR (thin solid line) and TNR (thin dotted line) when using the υ-SVM model on the test dataset. The values of TPR and TNR are presented on a scale from 0% to 100%.

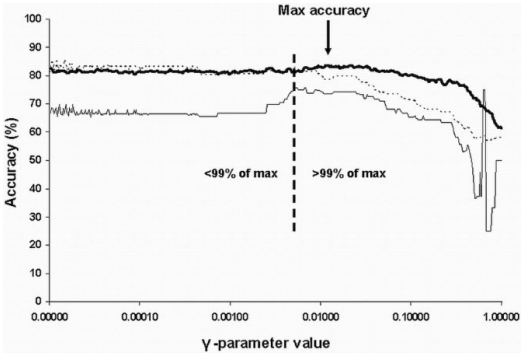


FIG. 6. Mean diagnostic accuracy values across the four observers for the υ-SVM model as a function of changing the γ parameter from 0.00001 ($1/10^5$) to 1 in steps of $1/10^n$ (n=5: −0.1:1), keeping the υ parameter at 0.56. Based on the resulting low-pass-filtered accuracy values (percent of gliomas graded correctly) in the training dataset (bold solid line), the maximal accuracy value (83.49%) was found at γ = 0.0112. Using the lowest γ value (0.0065) providing accuracy values within 99% (>82.66%) of the maximal filtered accuracy value resulted in higher TPR (thin solid line) and TNR (thin dotted line) when using the υ-SVM model on the test dataset. The values of TPR and TNR are presented on a scale from 0% to 100%.

model on nCBV histograms derived from MR perfusion data. Using the entire nCBV tumor distribution rather than a single threshold, a more robust base for prediction is created. One potential feature of the υ-SVM model is that multiple channels of information can be included in the training dataset. In addition to MR perfusion imaging, a standard MRI protocol in patients suspected of a glioma typically consist of conventional MRI (29), MR diffusion

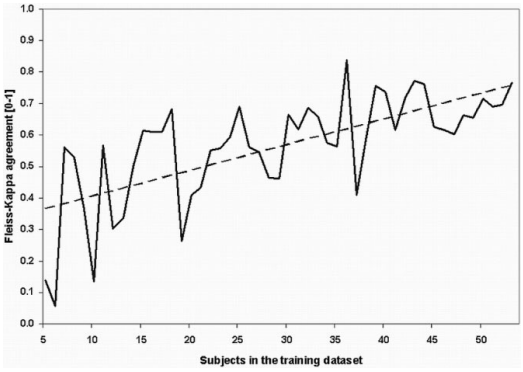


FIG. 7. Fleiss-Kappa agreement values between the four observers as a function of patients included in the training dataset (at optimal values of υ and γ). For each observer, a random sample of five histogram signatures up to all 53 signatures was included in the training dataset, adding one signature at the time. As illustrated by the trend line (dotted line), a significant increase (R^2 = 0.463, $P <$ 0.001) in the agreement values was observed as more patients were included in the training dataset.

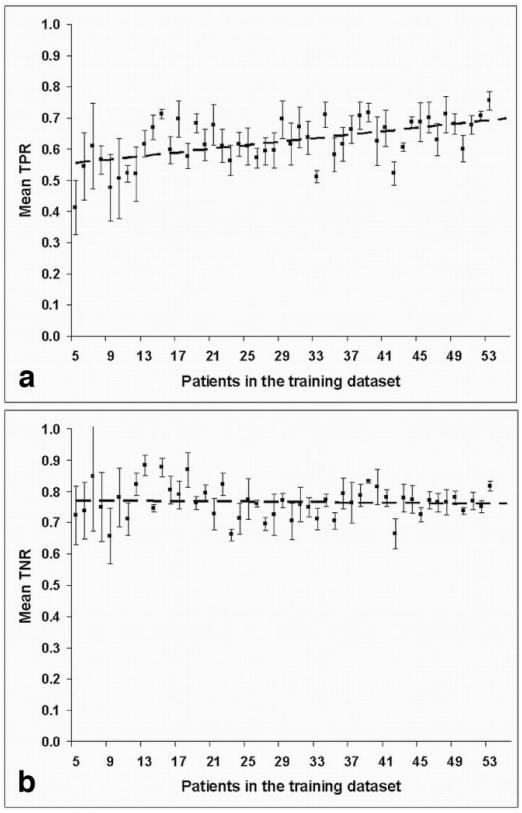


FIG. 8. For an increasing number of patients included in the training dataset (5–53), the scatter plots show mean TPR (**a**) and mean TNR (**b**) across the four observers when using the υ-SVM model on the test dataset. For each step, the υ-SVM model was based on optimal υ and γ values only. The error bars indicate standard errors of the mean values. As illustrated by the trend lines (dotted lines), TPR increased significantly with sample size (R^2 = 0.311, $P <$ 0.001), whereas TNR (R^2 = 0.004, P = 0.692) did not.

imaging (30), and MR spectroscopy (7). Also, including parameters from alternative dynamic imaging techniques such as permeability analysis (K^{trans}) (31), vessel size imaging (32), and vascular-space-occupancy imaging (33) might improve the diagnostic accuracy obtained with the υ-SVM model. A more comprehensive model including multiple information channels is currently being developed and evaluated. Furthermore, it has been shown that SVM models have very good generalization ability and are well-founded in theory (34,35). In addition, besides handling large dataset and large feature vectors efficiently, this kind of classification approach also yields superior performance in many real-world applications when compared to classical approaches like nearest neighbor classifiers (kNN) (35,36). All these properties are attractive in a classification task such as glioma grading since sufficient flexibility and robustness against data variability can be provided.

The υ-SVM model could differentiate between the 53 high-grade and low-grade gliomas in the training dataset with a perfect accuracy of 100% (53/53) for all observers. When the histological diagnosis is known, our results suggest that a SVM-based model might be a better method for retrospective grading of gliomas than current grading methods (1–4,6–8,12). The clinical utility of a diagnostic test depends, however, on its ability to correctly diagnose new cases based on historical data. As shown by our results, this is a challenging task. Although the "retrospective" accuracy was not achieved when using this model on new patient data, TPR and TNR values above 0.76 were obtained with a modest training set sample size. To our knowledge, there are no previous studies that have investigated the prospective diagnostic accuracy of glioma grading based on DSC imaging. It is therefore difficult to assess the quality of the proposed method compared to other potential prospective grading methods. Clearly, the diagnostic accuracy should be as high as possible to have clinical utility, and the accuracy of about 0.8 obtained with the proposed υ-SVM model may not be sufficient for extensive clinical use. However, the TPR was shown to significantly increase with increasing sample size, suggesting that the diagnostic accuracy will reach an acceptable level with a sufficient increase in the sample size.

When iterating over the υ parameter, the maximal filtered accuracy value of the cross-validation was found at $\upsilon = 0.31$. However, using $\upsilon = 0.31$ did not give optimal TPR and TNR values for the test dataset. The diagnostic accuracy for the test dataset could be increased by increasing υ beyond the optimal value for the training dataset, with minimal influence (<1% reduction) on the accuracy for the training dataset. The same result was observed when iterating over the γ parameter. Using the lowest γ value providing accuracy values in the training dataset within 99% of the maximum accuracy value resulted in a more generalized model (higher TPR and TNR in the test dataset) without losing more than 1% accuracy in the training dataset. This suggests that any grading model, or threshold value, is inherently overadapted to the data from which it was derived. Further, at least for the relatively low sample size used in our study, the model was rather insensitive to changes in the γ-value. Hence, it is therefore recommended to focus on optimizing the υ parameter.

When increasing υ toward the upper bound of the admissible vales (0.9) (16), the accuracy of the υ-SVM model on the training data decreased, probably because the model becomes too generalized, resulting in poor discriminating power. Interestingly, this had little effect on the prediction of glioma grades in the test dataset, as TPR and TNR values were relatively unchanged for $\upsilon > 0.7$. As confirmed by theory, this is because the υ parameter altering the support vectors of the training dataset may not result in a similar variation in the corresponding test dataset (18,37). However, as confirmed by our data, this does not imply that the optimal υ value is found at the highest υ value ($\upsilon = 0.9$).

The υ-SVM model on the test dataset provided higher TNR than TPR. In our study, the final υ-SVM model was based on υ and γ values providing the highest accuracy (combination of TPR and TNR) on the training dataset. However, depending on the desired output, a model providing a high TPR at the cost of a reduced TNR might be suitable. In glioma grading, it can be argued that a high TPR is more important than a high TNR as the consequence of treating a high-grade glioma as a low-grade glioma is more severe than treating a low-grade glioma as a high-grade glioma (7,38).

The TPR of the υ-SVM model on the test dataset was found to significantly increase with increasing sample size, whereas TNR showed no significant sample-size dependence. This result suggests a smaller interpatient variation in the histogram signatures from low-grade gliomas compared to the signatures from high-grade gliomas. This is confirmed by studies showing that the distributions of nCBV values in high-grade gliomas are more heterogeneous than the distributions in low-grade gliomas (8,12). Thus, by including more samples for training, the variability within the histogram signatures of the high-grade gliomas allows a better optimization of the margin and the separating hyperplane, thereby resulting in a higher generalization and robustness of the υ-SVM model.

In addition, the interobserver agreement between the four observers increased significantly with increasing sample size. Hence, as a larger sample size becomes available, the υ-SVM model should become less influenced by operator-induced variations in the input data. A further aim with increasing sample size would be to obtain sufficient discriminating power between classes to differentiate between WHO grades I, II, III, or IV instead of just low-grade or high-grade tumors.

A potential limitation in our model was that a standard RBF kernel function was used. Although the chosen kernel has previously been shown to fit histogram data (19), it is likely that the model could be further optimized by a more careful selection of the kernel function (16). Also, all glioma grades (I–IV) were included in our study. Even though glioma grading might be difficult if patients with grade-I pilocytic astrocytomas and grade-II oligodendroglial tumors are included (39,40), it would represent a major limitation to the predictive model if these tumors had to be characterized by other diagnostic techniques. Another limitation to the study is that the developed υ-SVM model has so far not been tested on data from other institutions, and therefore, although the current histogram grading method shows high interobserver agreement (8), the stability of the model for interinstitutional data has not been shown. To address this limitation, a multicenter evaluation of the υ-SVM model is currently being initiated by our group. Potential confounding variability in the tumor ROI definitions between observers and institutions will be addressed in this evaluation.

In summary, we have shown that a υ-SVM-based predictive model can be applied to tumor histogram signatures obtained from nCBV maps derived from MR-based DSC imaging. The method enables prospective determination of glioma grade with a significant increase in TPR and interobserver agreement with increasing size of the training dataset. The υ-SVM model as presented in this study can readily be implemented in a clinical setting, providing an important step toward a fully-automated MR-based glioma grading procedure.

REFERENCES

1. Law M, Oh S, Babb JS, Wang E, Inglese M, Zagzag D, Knopp EA, Johnson G. Low-grade gliomas: dynamic susceptibility-weighted contrast-enhanced perfusion MR imaging—prediction of patient clinical response. Radiology 2006;238:658–667.
2. Covarrubias DJ, Rosen BR, Lev MH. Dynamic magnetic resonance perfusion imaging of brain tumors. Oncologist 2004;9:528–537.
3. Aronen HJ, Gazit IE, Louis DN, Buchbinder BR, Pardo FS, Weisskoff RM, Harsh GR, Cosgrove GR, Halpern EF, Hochberg FH. Cerebral blood volume maps of gliomas: comparison with tumor grade and histologic findings. Radiology 1994;191:41–51.
4. Knopp EA, Cha S, Johnson G, Mazumdar A, Golfinos JG, Zagzag D, Miller DC, Kelly PJ, Kricheff II. Glial neoplasms: dynamic contrast-enhanced T_2*-weighted MR imaging. Radiology 1999;211:791–798.
5. Wetzel SG, Cha S, Johnson G, Lee P, Law M, Kasow DL, Pierce SD, Xue X. Relative cerebral blood volume measurements in intracranial mass lesions: interobserver and intraobserver reproducibility study. Radiology 2002;224:797–803.
6. Hakyemez B, Erdogan C, Ercan I, Ergin N, Uysal S, Atahan S. High-grade and low-grade gliomas: differentiation by using perfusion MR imaging. Clin Radiol 2005;60:493–502.
7. Law M, Yang S, Wang H, Babb JS, Johnson G, Cha S, Knopp EA, Zagzag D. Glioma grading: sensitivity, specificity, and predictive values of perfusion MR imaging and proton MR spectroscopic imaging compared with conventional MR imaging. AJNR Am J Neuroradiol 2003;24:1989–1998.
8. Emblem KE, Nedregaard B, Nome T, Due-Tonnessen P, Hald JK, Scheie D, Casar Borota O, Cvancarova M, Bjornerud A. Glioma grading using histogram analysis of blood volume heterogeneity from MR-derived cerebral blood volume maps. Radiology 2008;247:808–817.
9. Swanson KR, Bridge C, Murray JD, Alvord EC Jr. Virtual and real brain tumors: using mathematical modeling to quantify glioma growth and invasion. J Neurol Sci 2003;216:1–10.
10. Collins VP, Loeffler JS, Tivey H. Observations on growth rates of human tumors. Am J Roentgenol Radium Ther Nucl Med 1956;76:988–1000.
11. Hogea C, Davatzikos C, Biros G. An image-driven parameter estimation problem for a reaction-diffusion glioma growth model with mass effects. J Math Biol 2008;56:793–825.
12. Law M, Young R, Babb J, Pollack E, Johnson G. Histogram analysis versus region of interest analysis of dynamic susceptibility contrast perfusion MR imaging data in the grading of cerebral gliomas. AJNR Am J Neuroradiol 2007;28:761–766.
13. Fryback DG, Thornbury JR. The efficacy of diagnostic imaging. Med Decis Making 1991;11:88–94.
14. Boser BE, Guyon IM, Vapnik VN. A training algorithm for optimal margin classifiers. In: Proceedings of the Fifth Annual ACM Conference on Computational Learning Theory (COLT 1992), July 27–29, 1992, Pittsburgh, PA, USA. New York: Association for Computing Machinery (ACM); 1992. p:144–152.
15. Vapnik VN. Universal learning technology: support vector machines. NEC J Adv Tech 2005;2:137–144.
16. Chang CC, Lin CJ. Training nu-support vector classifiers: theory and algorithms. Neural Comput 2001;13:2119–2147.
17. Fan RE, Chen PH, Lin CJ. Working set selection using second order information for training SVM. J Mach Learn Res 2005;6:1889–1918.
18. Scholkopf B, Smola AJ, Williamson RC, Bartlett PL. New support vector algorithms. Neural Comput 2000;12:1207–1245.
19. Chapelle O, Haffner P, Vapnik VN. Support vector machines for histogram-based image classification. IEEE Trans Neural Netw 1999;10:1055–1064.
20. Louis DN, Ohgaki H, Wiestler OD, Cavenee WK. Astrocytic tumours and Oligodendroglial tumours. WHO classification of tumours of the central nervous system. 4th edition. Lyon: International Agency for Research on Cancer; 2007. p 13–67.
21. Laurent S, Elst LV, Muller RN. Comparative study of the physicochemical properties of six clinical low molecular weight gadolinium contrast agents. Contrast Media Mol Imaging 2006;1:128–137.
22. Rosen BR, Belliveau JW, Vevea JM, Brady TJ. Perfusion imaging with NMR contrast agents. Magn Reson Med 1990;14:249–265.
23. Ostergaard L, Weisskoff RM, Chesler DA, Gyldensted C, Rosen BR. High resolution measurement of cerebral blood flow using intravascular tracer bolus passages. Part I: Mathematical approach and statistical analysis. Magn Reson Med 1996;36:715–725.
24. Boxerman JL, Schmainda KM, Weisskoff RM. Relative cerebral blood volume maps corrected for contrast agent extravasation significantly correlate with glioma tumor grade, whereas uncorrected maps do not. AJNR Am J Neuroradiol 2006;27:859–867.
25. Maes F, Collignon A, Vandermeulen D, Marchal G, Suetens P. Multimodality image registration by maximization of mutual information. IEEE Trans Med Imaging 1997;16:187–198.
26. Schmainda KM, Rand SD, Joseph AM, Lund R, Ward BD, Pathak AP, Ulmer JL, Badruddoja MA, Krouwer HG. Characterization of a first-pass gradient-echo spin-echo method to predict brain tumor grade and angiogenesis. AJNR Am J Neuroradiol 2004;25:1524–1532.
27. Grier JT, Batchelor T. Low-grade gliomas in adults. Oncologist 2006;11:681–693.
28. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159–174.
29. Talos IF, Zou KH, Ohno-Machado L, Bhagwat JG, Kikinis R, Black PM, Jolesz FA. Supratentorial low-grade glioma resectability: statistical predictive analysis based on anatomic MR features and tumor characteristics. Radiology 2006;239:506–513.
30. Lam WW, Poon WS, Metreweli C. Diffusion MR imaging in glioma: does it have any role in the pre-operation determination of grading of glioma? Clin Radiol 2002;57:219–225.
31. Patankar TF, Haroon HA, Mills SJ, Baleriaux D, Buckley DL, Parker GJ, Jackson A. Is volume transfer coefficient (K(trans)) related to histologic grade in human gliomas? AJNR Am J Neuroradiol 2005;26:2455–2465.
32. Kiselev VG, Strecker R, Ziyeh S, Speck O, Hennig J. Vessel size imaging in humans. Magn Reson Med 2005;53:553–563.
33. Lu H, Law M, Johnson G, Ge Y, van Zijl PC, Helpern JA. Novel approach to the measurement of absolute cerebral blood volume using vascular-space-occupancy magnetic resonance imaging. Magn Reson Med 2005;54:1403–1411.
34. Anguita D, Boni A, Ridella S. Evaluating the generalization ability of support vector machines through bootstrap. Neural Process Lett 2000;11:51–58.
35. Hearst MA. Support vector machines. IEEE Intel Syst 1998;13:18–28.
36. Lee JW, Lee JB, Park M, Song SH. An extensive comparison of recent classification tools applied to microarray data. Comput Stat Data Anal 2005;48:869–885.
37. Chalimourda A, Schoelkopf B, Smola AJ. Experimentally optimal nu in support vector regression for different noise models and parameter settings. Neural Netw 2004;17:127–141.
38. Claus EB, Black PM. Survival rates and patterns of care for patients diagnosed with supratentorial low-grade gliomas: data from the SEER program, 1973-2001. Cancer 2006;106:1358–1363.
39. Lev MH, Ozsunar Y, Henson JW, Rasheed AA, Barest GD, Harsh GR, Fitzek MM, Chiocca EA, Rabinov JD, Csavoy AN, Rosen BR, Hochberg FH, Schaefer PW, Gonzalez RG. Glial tumor grading and outcome prediction using dynamic spin-echo MR susceptibility mapping compared with conventional contrast-enhanced MR: confounding effect of elevated rCBV of oligodendrogliomas [corrected]. AJNR Am J Neuroradiol 2004;25:214–221.
40. Cha S, Tihan T, Crawford F, Fischbein NJ, Chang S, Bollen A, Nelson SJ, Prados M, Berger MS, Dillon WP. Differentiation of low-grade oligodendrogliomas from low-grade astrocytomas by using quantitative blood-volume measurements derived from dynamic susceptibility contrast-enhanced MR imaging. AJNR Am J Neuroradiol 2005;26:266–273.