Diagnostic performance of texture analysis on MRI in grading cerebral gliomas

ABSTRACT

BACKGROUND AND PURPOSE
Grading of cerebral gliomas is important both in treatment decision and assessment of prognosis. The purpose of this study was to determine the diagnostic accuracy of grading cerebral gliomas by assessing the tumor heterogeneity using MRI texture analysis (MRTA).

MATERIAL AND METHODS
95 patients with gliomas were included, 27 low grade gliomas (LGG) all grade II and 68 high grade gliomas (HGG) (grade III = 34 and grade IV = 34). Preoperative MRI examinations were performed using a 3T scanner and MRTA was done on preoperative contrast-enhanced three-dimensional isotropic spoiled gradient echo images in a representative ROI. The MRTA was assessed using a commercially available research software program (TexRAD) that applies a filtration-histogram technique for characterizing tumor heterogeneity. Filtration step selectively filters and extracts texture features at different anatomical scales varying from 2mm(fine features) to 6mm(coarse features), the statistical parameter standard deviation (SD) was obtained. Receiver operating characteristics (ROC) was performed to assess sensitivity and specificity for differentiating between the different grades and calculating a threshold value to quantify the heterogeneity.

RESULTS
LGG and HGG was best discriminated using SD at fine texture scale, with a sensitivity and specificity of 93% and 81% (AUC 0.910, p<0.0001). The diagnostic ability for MRTA to differentiate between the different sub-groups (grade II - IV) was slightly lower but still significant.

CONCLUSIONS
Measuring heterogeneity in gliomas to discriminate HGG from LGG and between different histological sub-types on already obtained images using MRTA can be a useful tool to augment the diagnostic accuracy in grading cerebral gliomas and potentially hasten treatment decision.

Abbreviations
SD standard deviation
TA texture analysis
MRTA MR texture analysis
HGG high grade gliomas
LGG low grade gliomas
Introduction

Cerebral gliomas are the most common primary malignant brain tumour in adults and include astrocytoma, oligodendroglioma, mixed oligoastrocytoma, and ependymoma (1). For prognostic and treatment purposes they are further stratified into two groups, high grade gliomas (HGG), classified as WHO grade III-IV tumours, and low grade gliomas (LGG), classified as WHO grade I-II tumours (2). HGG have in most cases a very low life expectancy (3, 4), while LGG is associated with a longer life expectancy (5, 6). Treatment for HGG normally includes surgical resection followed by radiotherapy with or without chemotherapy, while treatment for LGG is usually surgical resection followed by close observation (7, 8). Hence, correct histological diagnosis is imperative for correct treatment. Brain biopsy represents the gold standard for histopathological diagnosis, which is based on nuclear pleomorphism, mitotic activity, cellularity, endothelial cell proliferation and presence of necrosis (9). This is increasingly challenged by new non-invasive advanced MRI techniques and research into additional sequences to improve radiological diagnostic accuracy. Image heterogeneity quantification could potentially lead to an objective and more accurate non-invasive radiological diagnosis which would impact patient management by allowing a more tailored and personalized management. Finding innovative ways to utilize pre-existing images has the potential to not only increase the accuracy of diagnoses but also optimally utilize scarce healthcare resources. Texture analysis (TA) assesses the distribution of gray-levels within an image to obtain texture features of intra-lesional heterogeneity (10, 11). Assessment of tumour heterogeneity, which is a well-recognized feature of malignancy, is important as it is related to poor prognosis (12). Imaging tumour heterogeneity via TA could be a non-invasive tool to assess prognosis, disease-severity and treatment-response
evaluation and prediction (10, 11). A considerable amount of work has been done on
CT texture analysis (CTTA) in the area of risk-stratification (13-16) compared to MRI.
Due to its multi-parametric approach, MRI is visually more heterogeneous than CT
and may be a powerful platform to quantify tumour heterogeneity with TA.
The purpose of this study was two-fold: firstly to exploit the use of TA on contrast-
enhanced T1-weighted MR imaging to differentiate between HGG and LGG, and
secondly to define the role of TA in the clinical decision-making process, in terms of
sensitivity, specificity, and accuracy.
Material and methods

Patients

The regional ethics committee was consulted and informed consent was waived as this is a single institution retrospective study on archived anonymised data on patients with cerebral gliomas. Preoperative MRI studies of 95 patients with histopathological evidence of cerebral glioma between December 2006 and September 2012 were evaluated retrospectively. Samples were obtained by tumour resection or biopsy, and histopathological diagnosis determined according to the World Health Organization (WHO) classification of Tumours of the Central Nervous System (2). A total of 27 patients (mean age, 44 years; age range, 21 to 70 years) had histologically verified grade II, 34 patients (mean age, 56 years; age range, 24 to 81 years) had histologically verified grade III and 34 patients (mean age, 58 years; age range, 25 to 87 years) had grade IV (glioblastoma multiforme). Detailed information is given in Table 1.

MR Imaging

All patients were examined with the same imaging acquisition protocol on a 3 T whole-body MRI system (Signa HDx, GE Medical Systems, Milwaukee, Wisconsin). The MRI protocol consisted of a T1 inversion recovery (T1 IR) sequence with the following parameters: TR / TE / TI 2500 ms / 9.6 ms / 920 ms, FOV 240x240 mm, matrix size 384x224, slice thickness 5 mm, slice gap 1.5 mm. The T2-weighted sequence was performed with TR/TE 6000 ms / 95 ms, FOV 240x240 mm, matrix size 480x480,
slice thickness 5 mm, slice gap 1 mm. The fluid-attenuated inversion-recovery (FLAIR) sequence was performed with TR/TE/TI 9500 ms/120 ms/2250 ms, FOV 240x240 mm, matrix size 384x224, slice thickness 5 mm, slice gap 1.5 mm. After contrast medium gadopentate dimeglumine (Magnevist, Bayer Schering Pharma, Berlin, Germany), a three-dimensional isotropic spoiled gradient echo (SPGR) sequence, using the following parameters: TR/TE 7.8 ms/3 ms, FOV 256x256 mm, matrix size 256x256 mm, slice thickness 1 mm was acquired.

Texture Analysis (TA)

Post-contrast SPGR sequence was retrieved from the brain tumour preoperative protocol. The contrast-enhanced SPGR image with the largest axial cross-section of the tumour was selected for TA and the corresponding slice in DICOM (Digital Imaging and Communications in Medicine) format was retrieved. The TA was performed by a clinical researcher with 4 years experience with TA on radiographic imaging. The ROI was placed around the area of the tumour with the largest diameter of solid tumor, including both contrast enhanced tissue (or the solid portion of the tumour in cases of non-enhancing tumours) and necrosis, but avoiding large cysts, haemorrhages and oedema. The ROIs were drawn by the same researcher with guidance of a board-certified neuroradiologist, without knowledge of the final histological tumour diagnosis. The window level was manually set by the researcher to best display the tumour. This has no impact on the texture analysis results (i.e changing the window does not change the outcome of the generated texture parameters) just the visual perception. The MR texture analysis (MRTA) results were compared to histopathological diagnosis.
The MRTA methodology used for this study follows a previously published study on CTTA in gliomas (16), which employed a filtration-histogram technique. It extracts and enhances texture features at different sizes within the largest cross-section of the glioma on contrast enhanced SPGR image followed by histogram quantification. MRTA was carried out using a commercially available research software called TexRAD (TexRAD – www.texrad.com, part of Feedback Plc, Cambridge, UK) (13). The initial filtration step employs a Laplacian of Gaussian (LoG) band-pass filtration, which extracts and highlights image features at different sizes corresponding to spatial scale filter (SSF) ranging from 2-6 mm in width (radius) where SSF 2 mm = fine, SSF 3-5 mm= medium and SSF 6 mm = coarse texture features. SSF values lower than 2 were not considered, as they tend to represent image noise. Quantification of histogram (with and without filtration) was based on the parameter standard deviation (SD), which represents the width of the histogram or degree of variation from the mean pixel value (equation shown below).

\[
SD = \left( \frac{1}{(n-1)} \sum_{(x,y)\in R} [a(x,y) - \bar{a}]^2 \right)^{\frac{1}{2}},
\]

Were R is the ROI within the image a(x, y), \( n \) is the total number of pixels in R and \( \bar{a} \) is the mean value within R.

A recent study by Miles et al (14) demonstrated that SD increases approximately in proportion to the square root of the number of objects (a component of image heterogeneity) highlighted by the filter and their mean intensity difference compared to background tissue (i.e. dark and bright objects are both positive). In other words a higher SD-value, implies increased tumor heterogeneity. (Figure 1)
Statistical Analysis

The parameter SD was used to measure texture analysis with and without filtration which was correlated against grade II-IV using non-parametric Spearman rank correlation. To individually discriminate LGG from HGG, grade II from grade III, and grade III from grade IV, a two-tailed Mann Whitney test was applied. The areas under the curve (AUC) for SD to distinguish defined tumour grades were calculated by receiver operating characteristic (ROC) curve analysis. The optimal thresholds and their determined sensitivity and specificity were reported. Positive predictive value (PPV) and negative predictive value (NPV) was also calculated from the ROC analysis. All statistical significance was set to 5% and was performed on SPSS (22.0.0.1) and MedCalc (MedCalc software, Mariakerke, Belgium).
Results

Correlation analysis

The correlation between the defined tumour grades and HGG vs LGG for SD was significant and is displayed in Table 2. Unfiltered SD with rs = 0.685 (p<0.0001) showed the best correlation between all the grades and SD filter value 2 displayed rs = 0.640 (p<0.0001) for HGG vs LGG correlation.

Differences in MRTA values in glioma grading

Table 3 summarizes all pair wise comparisons among the tumour grades. SD significantly differentiate between HGG and LGG, between grade II and grade III and grade III and grade IV gliomas for all the filters.

Diagnostic performance of MRTA

Tables 4 summarizes the ROC analysis results for SD, used to distinguish HGG from LGG (a), grade II from grade III gliomas (b) and grade III from grade IV (c). In the ROC curve analysis, the AUC for SD at SSF 2 best differentiated HGG from LGG (AUC = 0.910, p <0.0001, sensitivity = 80.9% specificity = 92.6% for a threshold above 1127, Table 4a, and ROC curve displayed in Figure 2). This filter revealed a PPV of 96.5% and NPV value of 65.8%. Likewise SD at SSF 2 was the best parameter for differentiating grade II from grade III (AUC = 0.837, p <0.0001, sensitivity = 70.6% specificity = 85.2%, Table 4b). Furthermore, SD without filtration
best differentiated grade III from grade IV (AUC = 0.728, \( p < 0.001 \), sensitivity = 97\% 
specificity = 44.1, Table 4c).

A scatterplot of SD values at SSF 2 against glioma grade (Grade II, III, IV) for each 
individual case is shown in Figure 3.
Discussion

Gliomas are heterogeneous lesions and histopathology is the gold standard for diagnosis, but has its limitations due to the inherently invasive procedure, sampling errors and variability in interpretation, especially in association with biopsy (9, 17). This is one of the reasons why physiologic MRI techniques are playing an increasing role in grading gliomas and guiding neurosurgical management of brain tumours (18). Recent studies have demonstrated the value of metabolic and physiologic MR imaging techniques, such as DWI (19), DTI (20, 21), MRS (22, 23), DSC MRI (17, 24, 25), and DCE MRI in the assessment of grading gliomas (26, 27). In addition, assessing the heterogeneity within a tumour by imaging might be of clinical benefit, particularly with the increased emphasis on personalized medicine. Hence, tumour heterogeneity is a clinically relevant parameter for imaging that may be quantifiable and could augment standard reporting methods (10). This study demonstrates the potential for MRTA to distinguish between LGG and HGG by quantifying heterogeneity, without additional imaging. TA is an easy post-processing step that can be performed on existing DICOM format images.

The preliminary assessment of the diagnostic performance for MRTA yielded a sensitivity of 81% and specificity of 93% for differentiating HGG from LGG using the primary histogram parameter SD post filtration (fine-texture scale) at a threshold of 1127. Patients with a SD value above 1127 indicate a high probability of HGG. For the same filter and parameter discussed above there is a significant correlation with tumour grade (rs = 0.640, p < 0.0001).

These results can in part be explained by the correlation between tumour heterogeneity and tumour grade (4). The heterogeneity of gliomas is most likely
related to intra-tumoural spatial variation, angiogenesis, extravascular/extracellular matrix, and areas of necrosis (10).

It is worth noting that other authors have selected different texture analysis parameters, such as entropy derived from ADC maps (28), or combining imaging features from several sequences and using support vector machines (29). A recent paper by Ryu et al (28) using TA on ADC maps found an AUC with entropy of 0.839. Entropy has also been related to the heterogeneity of tumour using CTTA (13, 18-20).

The method we have used for quantifying heterogeneity on MRI differs from other studies in two ways. Firstly we have used image filtration, where image features of 6 pixels or higher were highlighted (LoG filter is similar to a non-orthogonal wavelet – providing the flexibility to tune the filter value to extract and separately evaluate fine, medium and coarse texture features). This filtration step is important to remove image heterogeneity that is due to noise and highlight biological important heterogeneity. This proved to be critical in our study as the most significant correlations with grade were obtained from filtered texture parameters, in contrast to texture quantification of conventional images without filtration. Secondly, first order histogram statistical parameters which are directionally independent, such as standard-deviation (degree of variation from the mean pixel value, SD) were used. Previous MR texture studies in gliomas have employed second-order statistical parameters which are direction-oriented, i.e. statistical methods focusing on direction and length of pixels such as grey-level run length or co-occurrence matrices. These are in general more complex and more difficult to reproduce. Using both image filtration and an easily comprehensible directionally independent parameter strengthens this study compared to that mentioned above.
Quantifying heterogeneity to distinguish the different sub-grades was significant. Differentiating between grade II vs. grade III was more significant than between grade III and IV. However, the discrimination between the two high grades was quite pronounced in a previous study with TA on CT with contrast, sensitivity 91% and specificity 90% for SD (16). In addition, Ryn et al (28), using the ADC textural analysis parameters for differentiation between grade III and IV gliomas, found a sensitivity, specificity and accuracy of 81.8 %, 90 % and 84.4 % respectively.

Concluding from this, portrayed texture from necrosis is probably more applicable to extract from CT and ADC textural analysis than MRTA based on contrast-enhanced SPGR, as this is the major factor that distinguishes grade III from grade IV. Thus, using TA as a multimodal approach (combining CT TA and MRTA) or as a multiparametric approach (combining ADC and contrast-enhanced SPGR) to brain tumour analysis can provide a more complete tumour heterogeneity profile than any one modality alone for the diagnosis and grading of gliomas.

This is a novel technique to quantify these features in the different grades of gliomas. Previous studies using different acquisition techniques have been published. Many of these techniques are time consuming and can be difficult to reproduce. Our technique is user-friendly and uses an existing MRI sequence which is already part of the routine preoperative protocol. This is a single institution study and a large-scale clinical validation is needed. Biopsy will remain the gold standard, however, as brain biopsy is invasive and associated with sampling error, this study could be employed as an adjunct in confident decision-making, as part of the imaging contribution to the multidisciplinary approach to diagnosing patients with gliomas. Further emphasis should be placed on optimizing the potential of pre-existing radiological images. This should include exploring differences between the glioma sub-grades, as the
management of patients with gliomas differs not only between high grade and low grade tumours, but also between sub-grades. Recently the question has been raised of how non-histological data such as molecular information can be incorporated into the next WHO classification of CNS tumours (30). In agreement with Louis et al (30), we believe that the inclusion of non-tissue based information, i.e. clinical radiological information, can improve the clinical diagnostic work-up and biopsy guidance. Meaning the radiologist, in addition to reporting a preliminary diagnosis, can assist the neurosurgeons in getting a representative sample from the tumours area of highest heterogeneity, indicating the highest grade of the tumour and hence ensuring a representative tissue sample. This can make neuroimaging a more powerful noninvasive technique and have a positive impact on the care of patients with brain tumours.

It should be noted that the use of the same acquisition protocol in all our examinations strengthens our study. It is well known that texture features are affected by the MR imaging acquisition, especially the gain-factor and spatial resolution (29). However, a few limitations were encountered: firstly, as a pilot study using the single largest axial cross-section MRI slice of the tumour, might not correspond to the specific site where the diagnostic biopsy was retrieved, poses as weighted limitations of this study; however, acquiring stereotactic biopsies before tumour resection is logistically difficult and not normally performed in patients with newly diagnosed gliomas. In addition, a volumetric TA approach could show a more optimal representation of the tumour heterogeneity. Secondly, although the population comprises of consecutively selected patients meeting the inclusion criteria, there is a larger proportion of HGG compared to LGG, which has the
potential to skew the results. We attempted to mitigate this by using the first 34 of the total 142 patients with GBM. Thirdly, this was a retrospective study and therefore a prospective study using the thresholds highlighted in this study is needed to validate these results.

**Conclusions**

Quantifying glioma heterogeneity by using filtration and first order histogram statistical parameters such as SD determined by texture analysis highlights the different heterogeneous appearance of LGG and HGG as well as between glioma subgroups on contrast-enhanced SPGR images. This additional diagnostically useful information provided by texture analysis can augment the radiologists accuracy when grading gliomas.
Table 1. Showing inclusion and exclusion of patients.

Table 2. Spearman Rank Correlation coefficients with P values, the best rs values and their p value is in bold.

Table 3. Mann Whitney p values for SD SSF 2 to differentiate between HGG and LGG and sub-grades.

Table 4. ROC analysis of SD differentiating A) HGG vs LGG, B) grade II vs grade III and C) grade III vs grade IV. The best AUC values are highlighted in bold.

Figure 1. a) Contrast-enhanced SPGR image shows an intraaxial enhancing necrotic tumour (histopathologically verified as glioblastoma multiforme). The corresponding images selectively display, b) fine, c) medium and d) coarse textures.

Figure 2. Graph shows receiver operating characteristic (ROC) curve of SD SSF 2 for differentiation of HGG from LGG.

Figure 3. Scatterplot of SD SSF 2 for differentiating LGG from HGG, showing all individual cases in glial tumour grades II, III and IV.
REFERENCES


