Glioma Dynamics and Computational Models: A Review of Segmentation, Registration, and *In Silico* Growth Algorithms and their Clinical Applications

Elsa D. Angelini, Olivier Clatz, Emmanuel Mandonnet, Ender Konukoglu, Laurent Capelle and Hugues Duffau*

[1]*GET, Ecole Nationale Supérieure des Télécommunications, CNRS UMR 5141, Paris, France,* [2]*Asclepios Research Project - INRIA Sophia Antipolis, France,* [3]*Unité 678, Inserm/UPMC, 91 boulevard de l'Hôpital, 75634 Paris Cedex 13, France,* [4]*Department of Neurosurgery, Hôpital de la Salpêtrière, 47-83 Bd de l'hôpital, 75634 Cedex Paris Cedex 13, France,* [5]*Department of Neurosurgery, Hôpital Gui de Chauliac, CHU de Montpellier, 80 Avenue Augustin Fliche, 34295 Montpellier Cedex 5, France*

Abstract: Tracking gliomas dynamics on MRI has became more and more important for therapeutic management. Powerful computational tools have been recently developed in this context enabling *in silico* growth on a virtual brain that can be matched with real 3D segmented evolution through registration between atlases and patient brain MRI data. In this paper, we provide an extensive review of existing algorithms for the three computational tasks involved in patient-specific tumor modeling: image segmentation, image registration, and *in silico* growth modelling (with special emphasis on the proliferation-diffusion model). Accuracy and limits of the reviewed algorithms are systematically discussed. Finally applications of these methods for both clinical practice and fundamental research are also discussed.

INTRODUCTION

The advent of MRI scanning protocols has allowed accurate follow-up of tumor growth through volumetric measurements. Interpretation of the radiological evolution of the tumor appears of utmost importance for therapeutic management, especially for low-grade glioma. Indeed, patients are most of the time asymptomatic (except in the case of epilepsy) during the "low-grade" phase, and the tumoral evolution is only monitored by MRI follow-up, both prior and after treatment. However, such information about the tumor dynamics is usually not fully integrated with the therapeutic strategy, and the assessment of the tumor evolution is still limited to qualitative descriptions (recurrence, progression, regression, stability). Thus, it is expected that bio-mathematical models will help to quantify tumor dynamics, to simulate treatment effects, and finally to optimize therapeutic strategies.

Computational models of gliomas dynamics have been initiated more than ten years ago [1, 2]. The first studies aimed at modeling the effect of chemotherapy and surgical resection on the evolution of high-grade gliomas. If the mathematical framework introduced at that time is still in use, there have been considerable advances in its numerical resolution. In particular, digital brain templates, provided by MRI, enable to implement the biophysics equations onto accurate virtual anatomy. This in turn allows to refine the model, by introducing for example different cell motility in white and gray matter [3], and inside white matter, along and orthogonally to axonal fasciculus [4]. However, published studies have never seriously matched observed radiological evolution with virtual *in silico* dynamics. Such a comparison would require three different steps (as illustrated in Fig. **1**): segmentation of actual growth, registration on a virtual brain atlas, and identification of model parameters corresponding to optimal matching between actual and simulated evolution.

We thus propose in this paper to review existing algorithms for segmentation, registration and *in silico* growth, with special attention to the accuracy and reliability of the methods.

SEGMENTATION

An accurate determination of the actual tumor evolution requires full 3D segmentation on digital images. Manual segmentation by an expert is still considered as the reference method, but is a time consuming task with high inter and intra-observer variability.

Many automated or semi-automated approaches were developed over the past ten years showing great variability in results and performance in terms of reproducibility. Challenges in the segmentation of gliomas from MRI data are related to i) the infiltration of cells into the tissue, inducing unsharp borders with irregularities and discontinuities (a tumor is not necessary a single connected object), ii) the great variability in their contrast uptake (depending on their vascularisation) and iii) their appearance on standard MRI protocols.

MRI protocols used for brain imaging typically include Proton density (PD), T1-weighted (e.g. SPGR), T1-weighted enhanced (T1E) with contrast agent (usually Gadolinium), and T2-weighted (e.g. FLAIR) data. T1 data provides detailled anatomical views of the brain along with high signals on haemorrhages. T1E data shows strong signal on all vascularized structures (including tumors and haemorrhages), whereas usual FLAIR images (with a slice thickness around

*Address correspondence to this author at the Department of Neurosurgery, CHU Montpellier, Hôpital Gui de Chauliac, 80 Avenue Augustin Fliche, 34295 Montpellier Cedex 5, France; Tel: 33 4 67 33 02 90; Fax: 33 4 67 33 75 57; E-mail: h-duffau@chu-montpellier.fr
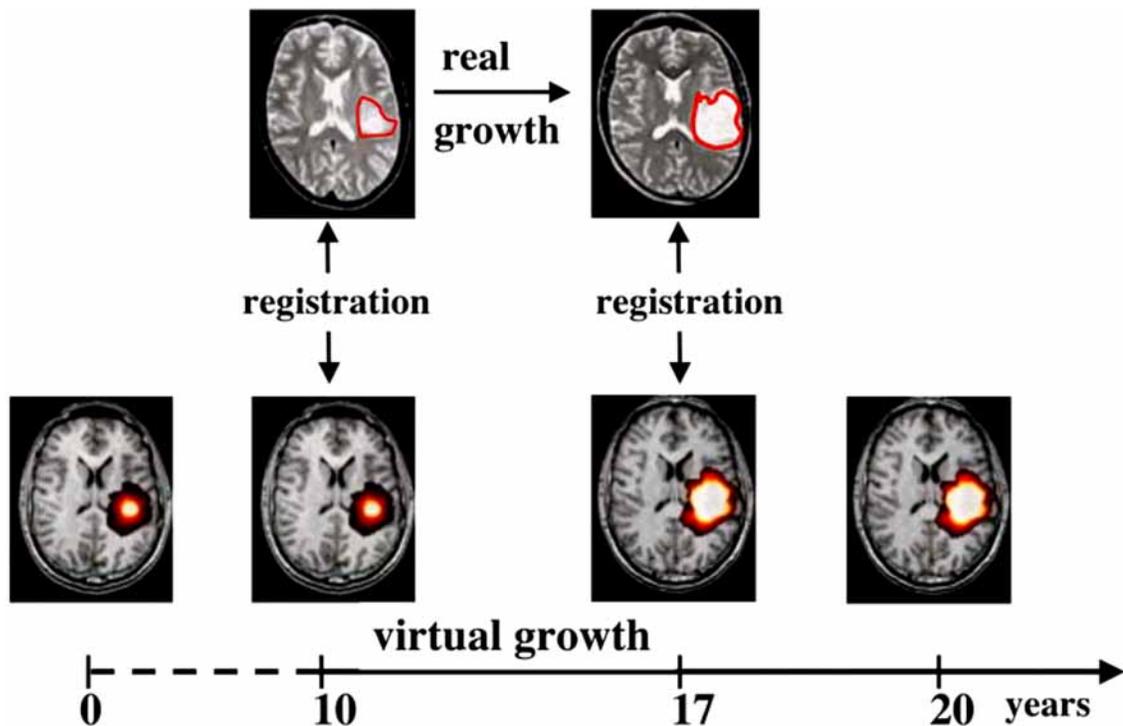
Fig. (1). Illustration of glioma growth evolution observed on temporal MRI studies.

5 mm) show less anatomical details, but high signal on tumors, infiltrations and edema.

Methods

Given an image, the segmentation task can be seen as the partition of the image into homogeneous objects, which correspond to a region-based segmentation approach, or as the detection of object contours within the image, corresponding to an edge-based segmentation approach. The majority of the MRI-based glioma segmentation methods that have been proposed in the literature are region-based. More recent methods, based on deformable models, also included edge-based information. In the case of MRI segmentation, several factors introduce a large amount of uncertainty in the segmentation process, including partial volume effects, integration of multi-protocol image data and observer variability. In this context, a large set of segmentation methods was designed in a statistical framework, providing a classification of the image data into different tissue types, while only few were designed with a deterministic approach.

Deterministic Approaches

In 1996, Gibbs *et al.* [5] introduced a morphological edge detection technique combined with simple region growing to segment enhancing tumors on T1 MRI data. Based on an initial sample of the enhanced tumor signal and the surrounding tissues, provided manually, an initial segmentation was performed combining pixel thresholding, fitting to an edge map of the image data and morphological opening and closing, inspired by the work proposed by Kennedy *et al.* [6]. The tumor area was defined based on pixel values in the range of 4 standard deviations around the mean value, constrained by the edge map.

In 2004 Letteboer *et al.* [7], proposed an interactive segmentation method for three types of tumors: full-enhancing, ring-enhancing and non-enhancing. After manual tracing of an initial slice, a series of morphological filtering operations (based on the watershed algorithm) was applied to partition the MRI volume data into homogeneous areas. A multiscale framework (i.e. analysis of the image data at different spatial resolutions) was employed to correlate segmented regions across different scales. The overall segmentation process was guided via an interactive user interface.

In 2005, Droske *et al.* [8], proposed to use a deformable model, implemented with a level set formulation, to partition the MRI data into regions with similar image properties, based on prior intensity-based pixel likelihoods for tumoral tissues. The deformable model optimization was performed on a spatially-adaptive grid, only refined in inhomogeneous regions. Homogeneity measures included gray value intervals, defined from a user input, and image gradient values. Some manual supervision of the deformable model was required, so that incremental segmented areas were proposed to the user who controlled the final segmentation results. More specifically, heterogeneous tumors, involving necrosis for example, required successive segmentations by addition or removal of intermediate results.

Statistical Approaches

In 1995, Vaiddynathan *et al.* [9], compared two supervised multispectral classification methods: k nearest neighbour (kNN) and spectral fuzzy C-means (FCM). For these two classification approaches, nine tissue classes were considered (background, CSF, WM, GM, fat, muscle, tumor, edema, necrosis). The authors also tested an interactive seed-growing segmentation approach on T1E MRI data. The seed-

growing algorithm only segmented tumor tissue based on a sample pixel population manually selected by the user.

In 1998, Clark *et al.* [10] introduced a knowledge-based (KB) automated segmentation method for glioblastomas on multispectral data combining T1E, PD and T2 weighted data. A training phase was performed on 17 slices from seven patients, extracting tumor size and enhancement level characteristics. Slices were first characterized as normal or abnormal via a fuzzy C-means (FCM) classification and the analysis of the clustering result through an expert system. Two examples of knowledge used in the predecessor system were: 1) in a normal slice, CSF belongs to the cluster center with the highest value in the intracranial region; 2) in image space, all normal tissues are roughly symmetrical along the vertical axis.

After a brain mask was computed, initial tumor segmentation, generated from vectorial histogram thresholding in the T1, PD and T2 images, was post-processed with a KB approach to eliminate non-tumor pixels.

Tumor heuristics used in the KB system were the following: "1) Gadolinium-enhanced tumor pixels occupy the higher-end of the T1 spectrum; 2) Gadolinium-enhanced tumor pixels occupy the higher-end of the PD spectrum, though not with the degree of separation found in T1 space; 3) Gadolinium-enhanced tumor pixels are generally found in the "middle" of the T2 spectrum, making segmentation based on T2 values difficult; 4) Slices with greater enhancement have better separation between tumor and non-tumor pixels, while less enhancement results in more overlap between tissue types". It is important to note that their notion of tumor pixel included edema and necrosis. A final processing stage was performed, based on histogram analysis of the tumor pixels and heuristics on the "density" of intensity features of non-tumor tissues. Indeed, based on the observation that tumors can show different levels of enhancement and very complex shapes, the final KB approach was focused on characterizing non-tumoral tissues.

In 2001, Kaus *et al.* [11] presented a complete validation of an automated segmentation method on T1E data from twenty patients with meningiomas and low-grade gliomas. The segmentation method, called an adaptive template-moderated classification, and described in [12, 13] was based on an iterative process. It alternated between a kNN classification of voxels into five hierarchical tissue types (background, skin-fat-bone, brain, ventricles, tumor) and a nonlinear registration of the data with an anatomical atlas (manually segmented MRI data of a single subject) to align the data with the template. The kNN classification used features from data intensity values and anatomical priors on the tissue location from the atlas. This method performed extraction of the five tissues in a pre-determined hierarchical order. Tissue mean values were learned on the patient's data via manual selection of three or four points for each tissue. To handle the presence of the tumor in the registration process, voxels assigned to the tumor class were masked with brain labels prior to registration with the atlas. This method obviously relied on a strong homogeneity assumption of the tumor's appearance on MRI data, which was reinforced by the use of anisotropic diffusion filtering.

In 2001 Moonis *et al.* [14] proposed a segmentation framework based on fuzzy connectedness (FC) which optimally clustered voxels into classes of high connectivity (analogous to a similarity measure). The method was applied to T1, T1E and T2 data, and initialized with an MRI data standardisation of the gray levels based on non-linear transformation of the histograms [15].

In 2005, Liu *et al.* [16], from the same group, used a similar approach based on a volume of interest on co-registered T1 and T2 data, to process only slices containing the tumor. A set of points inside the tumor were selected to initialize the statistics used in the FC. The threshold level applied to the FC maps to define the final segmentation result was determined empirically on five datasets and then fixed once for all. Segmentation was performed separately on the T2, T1E and subtracted (T1-T1E) data sets in 3D. Manual corrections of the segmentation results were performed by experts.

In 2001, Fletcher-Heath *et al.* [17], proposed a combination of unsupervised classification with FCM and knowledge-based (KB) image processing for segmentation of non-enhancing tumors. The FCM was run on spectral data (T1, T2, PD). As the authors pointed out, FCM tended to define clusters with similar sizes, which required an initial classification in ten classes. A KB system was then designed to re-cluster the segmentation results into seven classes based on a training phase. Difficulties principally arose in the separation of CSF and tumor signals.

In 2004, Mazzara *et al.* [18], compared the kNN approach from [9] and the KG-based approach from [10] for growth tumor volume (GTV) measurements on eleven patients with high and low-grade gliomas. As used in oncology radiation therapy, GTV corresponded to the area enclosing several contiguous clusters of enhancing pixels (i.e. including non-enhancing pixels within the area). The study showed severe limitations of the KG-system (which was not trained with the dataset to segment) in handling particular cases such as non-enhancing tumor margins or the presence of non-enhancing cystic necrotic tissues at the center of the tumor. On the other hand, the kNN segmentation method, trained with sample data from MRI slices to segment, lead to robust segmentation results on all patients. In 2006, Beyer *et al.* [19], from the same group, presented a similar and more recent comparative study, extracting GTV with the same two segmentation methods and evaluating the results in terms of predictive dose measurement for therapy planning.

In 2004, Zou *et al.* [20], proposed a continuous probabilistic segmentation framework, based on mixture modeling for two classes: tumor and non-tumor tissues. After initialization of the segmentation with the semi-automated method from Kaus *et al.* [11], the segmentation process involved estimation of the distribution parameters and probability values thresholding. Three metrics were proposed and evaluated to optimize the threshold selection: Receiver operating curve (ROC), which weights the sensitivity versus the specificity of the segmentation result, a Dice similarity coefficient, which is also a function of sensitivity and specificity and mutual information that directly compares the segmentation result to a ground truth.

In 2004, Prastawa *et al.* [21] proposed a segmentation framework based on outlier detection on T2 data. The abnormal tumor region was detected via registration on a normal brain atlas. Statistical clustering of the abnormal voxels, followed by a deformable model, were then used to isolate the tumor and the edema.

A summary of the reviewed papers and the specificities of the clinical aspects of the evaluation setup are provided in Table **1**.

Gold Standard Definition

Evaluation of tumor segmentation accuracy typically requires the definition of a ground truth segmentation. In this section we call Ω the image data domain (i.e. the entire set of pixels to segment), V the tumor volume (i.e. set of pixels identified as belonging to the tumor) provided by the evaluated segmentation method, by V_t the tumor volume identified by the ground truth method and by $|\;|$ the cardinality (i.e. number of voxels included) of a volume. The complementary set of V (i.e. set of pixels identified as not belonging to the tumor) is written as $\bar{V}$.

Static as well as statistical ground truth volumes have been defined and used in the reviewed literature to compare tumor segmentations. We review their definitions and constructions in this section.

Deterministic Ground Truth

The most widely used approach to define a ground-truth tumor object, is via manual tracing of the contours by one or several experts in neuroradiology. There are obvious limitations to such approach:

- Enhancement appearance of tumor on T1E data is very variable and depends on the degree of vascularization as well as permeability of the vessels to Gadolinium.

- Manual tracing bears some variability, which has been evaluated in several studies as detailed below in this section.

In the majority of papers, a standard manual segmentation of the tumor (and eventually of the brain) was defined from the segmentations of one or more independent human observers.

For example, in Kaus *et al.* [11], pixels were assigned to the ground truth, if at least three of four observers agreed with this assignment.
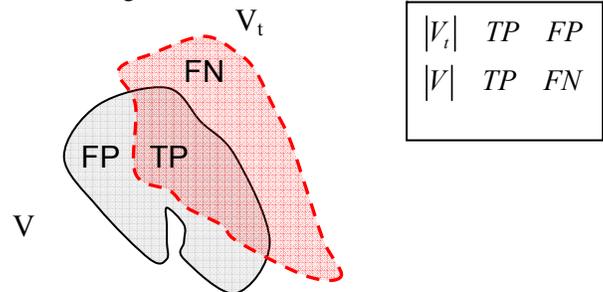
In Fletcher-Heath *et al.* [17], in Vaiddynathan *et al.* [9] in Droske *et al.* [8] and in Prastawa *et al.* [21], ground truth was defined from manual tracing from one expert.

To evaluate manual tracing variability, in Kaus *et al.* [11], one observer manually segmented four times the same 2D MRI slice, over one week.

In Liu *et al.* [16], two sets of manual tracings were used, based on 2 operators (a neuroradiologist and a trained expert in medical imaging assisted by a semi-automated method), and to evaluate segmentation precision, the two operators repeated the segmentation three times on ten T2, ten T1E and five T1E subtracted studies. To evaluate segmentation accuracy, five T2 cases were manually traced three times.

In Stadlbauer *et al.* [22], the authors compared manual delineation of hyperintense signal on T2 data to high ratio of Cho/NAA on MR spectroscopy imaging (MRSI). The ground truth for this study consisted of biopsy samples of

Table 1.　Summary of Reviewed Papers and Clinical Setup

	Data	LGG	Glioma HG	Year
Gibbs	T1E		10	1996
Letteboer		20		2004
Droske	T1E	?		2005
Liu	FLAIR, T1, T1E		10	2005
Vaidyanathan	T1, PD, T2		4	1995
Fletcher-Heath	T1, PD, T2		6	2001
Clark	T1, PD, T2 (all with Gd)		7	1998
Kaus	SPGR-Enh	14		2001
Moonis	FLAIR		19	2001
Mazzara	T1E, FLAIR (CT)	3	8	2004
Zou	T1E (SPGR)	3	3	2004
Prastawa	T2	1	1	2004

brain tissues in the tumor region detected on MRSI but not on T2 data.

In their early work, Gibbs *et al.* [5] evaluated their segmentation method on a simple phantom made of a water-filled glass head containing a balloon of known volume (varying from 10 to 16 cm^3 with 2cm^3 increment) of CuSO4, which was screened with the same T1E protocol as used for patients. They also compared their method to a thresholding-based segmentation method provided by a commercial software (ISG Technologies Inc), requiring 2D manual initializations. Efficiency of the commercial method and the proposed method was evaluated by repeating the initialization three times, from a single user.

Statistical Ground Truth

In Mazzara *et al.* [18], a statistical ground truth was defined from a set of manual tracings as the probability that a pixel is properly considered as part of the tumor. These probabilities were used as weights in the accuracy measurement, so that the true-positive accuracy measurement corresponded to the ratio of the total sum of weights contained in the segmented area versus the total sum of weights generated from nine manual tracings. A similar approach was used to measure false-positive values.

In Zou *et al.* [20], the authors used the "simultaneous truth and performance level estimation" (STAPLE) method introduced by the same group in [23], to compute a probabilistic estimate of the true segmentation, given multiple manual tracings, and provided a measure of the performance level represented by each segmentation.

In Letteboer *et al.* [7], the authors considered that manual tracing did not provide ground truth *per se* and compared their segmentation results to such tracings via a Bland and Altman statistical analysis. Considering that differences in volume measurements depend on the tumor size, which is correlated with the fact that the majority of the segmentation errors occur on the surface of the tumor, they proposed a normalization of the volume difference values D evaluated in the test:

$$D = \frac{\left(|V| - |V_t|\right)}{\left(\left(|V_D| - |V_t|\right) + \left(|V_t| - |V_E|\right)\right)/2}, \qquad (1)$$

where V_D is the segmented volume dilated by one pixel and V_E is the segmented volume eroded by one pixel. This normalized difference value directly correlates the D value with the number of different voxels.

Clinical Validation Methods

Given a ground truth representation of the tumor, several error measurements can be used to evaluate the accuracy of the segmentation method, including the notions of specificity, sensitivity, repeatability and efficiency, which are discussed in this section.

Error Measurements for a Given Segmentation

In Kaus *et al.* [11], validation was performed with measurements of segmentation accuracy (SA) defined as:

$$SA = \frac{\left(TP + TN\right)}{|\Omega|} \qquad (2)$$

using the true positive (TP) volume overlap measured as:

$$TP = \left|V \cap V_t\right| \qquad (3)$$

and the true negative (TN) volume overlap measured as :

$$TN = \left|\bar{V} \cap \bar{V}_t\right| \qquad (4)$$

It is important to note that this measure is very favorable to small objects such as circumscribed tumors.

In Clark *et al.* [10] and in Fletcher-Heath *et al.* [17], the authors used two measurements:

percent match (PM) ratio defined as:

$$PM = \frac{TP}{|V_t|} \qquad (5)$$

This measure corresponds to a TP volume fraction (TPVF).

correspondence ratio (CR) defined as:

$$CR = \frac{TP - FP/2}{|V_t|} \qquad (6)$$

An additional comparison was performed, evaluating the segmentation method for accuracy and precision of **longitudinal tumor evolution**. In this study they evaluated four patients with at least two scans and failed in measuring a growing tumor in one case, where a significant amount of fluid at the initial time lead to an overestimation of the initial tumor size.

Analogous measures were used by Letteboer *et al.* [7] as:

$$PM = \frac{TP}{\left(\left(|V_t| + |V|\right)/2\right)}, \qquad (7)$$

and

$$CR = \frac{\left|V - V_t\right|}{\left(\left(|V_t| + |V|\right)/2\right)} \qquad (8)$$

In Liu *et al.* [16] ,ten segmentations of patients with glioblastomas were evaluated with the following. Measurments:

1. **Precision**: reproducibility of the segmentation varying all parameters for the scanning protocol (i.e. using repetitive scans of individual patients) and for the segmentation

algorithm. This measure is analogous to the concept of **reproducibility** or **repeatability**.

2. **Accuracy**: comparison of segmented regions with respect to manual tracing. Two volume fraction (VF) measures were used:

`FPVF:

$$\frac{\left|V \cap \bar{V}_t\right|}{\left|V_t\right|} \tag{9}$$

and FNVF:

$$\frac{\left|V_t \cap \bar{V}\right|}{\left|V_t\right|} \tag{10}$$

Efficiency: computation and manual intervention time.

Similar measurements were also evaluated in **Letteboer** *et al.* [7] on twenty patients.

In Gibbs *et al.* [5], efficiency of the commercial method and the proposed method was evaluated by repeating the initialization three times, from a single user.

In Zou *et al.* [20], the three metrics evaluated to optimize the segmentation method were also used to evaluate the segmentation results: area under the receiver operating curve (AUC) which weights the sensitivity versus the specificity of the segmentation result, a Dice similarity coefficient (DSC), which is also a function of sensitivity and specificity and mutual information (MI) that directly compares the segmentation result to a ground truth.

In Prastawa *et al.* [21], the authors used three error metrics, including surface comparison of the tumor's outline. Theses metrics were the *PM* overlap measure, the Hausdorff distance and average surface distance.

Inter and Intra-Observer Variability

Inter and intra-observer variability is typically measured with the **coefficient of variation** (CV) of volume measurement defined as:

$$CV\% = \frac{\sigma_{V_t}}{\mu_{V_t}} \times 100 \tag{11}$$

where μ_{V_t} is the mean value and σ_{V_t} is the standard deviation of volume measurement made on V_t.

In Moonis *et al.* [14] the authors used the CV measure to evaluate variability of the user tracings and inputs to the segmentation method. In Vaiddynathan *et al.* [9], the authors also used the CV measures to evaluate reproduci-bility of tumor segmentation with respect to parameter setting of the different segmentation methods.

In Mazzara *et al.* [18], **intra-observer variability** of manual tracing was assessed with three tracings from one expert, and measured via the ratio of average disagreement:

$$\frac{\sum_{i=1}^{N}\left|V_t^i\right| - \left|\cap V_t^i\right|}{\sum_{i=1}^{N}\left|V_t^i\right| / N} \tag{12}$$

Based on this disagreement ratio, **inter-observer variability** was evaluated among three experts, comparing one expert to the two other ones.

In Letteboer *et al.* [7], three operators segmented the tumors twice. Observer variability was assessed through a Bland and Altman statistical analysis, measuring deviations from an average volume value. From this analysis, variability can be measured as the Bias ± 1.96 SD, where SD is the standard deviation of the volume measurements.

In Prastawa *et al.* [21], area overlap and surface distances were used to evaluate inter-observer variability.

Results

In Gibbs *et al.* [5], correlation of volume measurements based on phantom data was perfect, with less then 5% errors. Patient data segmentation required 2-3 minutes of user intervention and around 10 minutes of computational time. Segmentation with the commercial software required about 30 minutes of user time. Both methods provided between 9% and 13% precision in volume measurements (for tumor volumes in the range of 2 cm^3-80 cm^3), with a mean difference in observations of 0.1 ± 4.5 cm^3. Mean difference between two observers is 0.8 ± 1.8 cm^3. In conclusion, despite its simplicity and limitation to well enhanced tumor, not close to the skull, this early method provided fast and accurate global tumor volume measures.

In Letteboer *et al.* [7], average intra and inter-observer volume differences were 3.2% and 9.7%, while average manual PM values were 93.5% and 90%. Variability with Bland and Altman analysis was 0.04 ± 1.79 for manual tracing and -0.01 ± 0.76 for the segmentation method. They concluded that the watershed method was more reproducible than manual tracing. A negative bias showed under-segmentation from the watershed and they obtained best PM similarity measures for enhancing tumors and worst for non-enhancing ones.

In Droske *et al.* [8], computation time, including manual initialization required 3 minutes. They concluded that accuracy was high for homogeneous enhanced tumors and lower for heterogeneous tumors, without quantitative numbers.

In Vaidyanathan *et al.* [9], precision of manual tracing and three segmentation methods was evaluated in terms of user input and showed very large variability. Manual tracing average variability (inter-intra) was 6%-17%. Segmentation methods average variability was around 5%-8% for trained classification methods and 17%-6% for a seed growing method. These results demonstrated the better precision of trained-based segmentation approaches but also the weakness of these methods in terms of sensitivity to user inputs. They concluded that reproducibility was a major factor affecting tumor volume measurements (with differences

above 20 ml in their study), with a critical need for less-supervised methods.

In Fletcher-Heath *et al.* [17], they obtained a PM between 53% and 90% and a CR between 0.37 and 0.87 over six cases. For the two cases used for training the KB system, average PM was 72% and average CR was 0.67. For the four tested cases, average PM was 75% and average CR was 0.56.

In Clark *et al.* [10], overall results showed an overestimation of the tumor volume with automated segmentation. Regarding longitudinal studies, the methods also showed one case where tumor shrinkage was falsely measured, potentially due to high haemorrhage signal in the baseline MRI scan. They also reported 5% inter-observer variability in tumor volume manual measurement. For the proposed KB segmentation method, repeatability was fully guaranteed. PM was typically above 0.9 (range=0.69:1, average= 0.93), and CR was more variable (range=0.43:0.85, average=0.66). They also showed that less than 15% of the FP were really FP, not spatially connected to any ground-truth tumor pixel. Regarding a kNN- based method, they reported poorer PM measures (range=0.22:0.99, average= 0.71) and CR measures (range=-2.21:0.64, average =0.19).

In Kaus *et al.* [11], evaluation of the proposed segmentation method reported extremely high SA (over 99%), with very low segmentation variability. Intra and inter -observer variability (CV) for manual and automated segmentation was below 5%-15% and 4%-7%. This again suggested the gain in precision when using an automated (classification-based) segmentation method. Computation time for a three-dimensional data set was 75 minutes.

In Moonis *et al.* [14], on a set of ten patients, average CV for two different users was 0.27-0.21 on FLAIR data and 0.37-0.27 on T1E data. Inter-observer CV values were 0.38 and 0.29 for FLAIR and T1E data, with non-significant differences. The study reported that manual editing of the segmentation results lead to smaller tumor volumes, with a median change of volume of -17%. No significant difference was found in CV values with or without manual editing.

In Liu *et al.* [16] ,ten patients with glioblastomas were evaluated. Manual correction by experts was performed on all segmentation outputs evaluated. Results showed:

-   **Precision**: for repeated segmentations by two experts, the average CV measure for each expert and inter-observer CV were: (0.33-0.26, 0.43) on FLAIR data. (0.33-0.3, 0.36) on T1E data and (0.37-0.33, 0.41) on subtracted T1E. Regarding these three experiments, intra- and inter-observer percentage overlap was above 98%. Mean volume variation was 1.2% over two repeated FLAIR scans on five patients.

-   **Accuracy**: On five FLAIR cases, average FPVF was 4.45% and average FNVF 4.28%. Volume estimation accuracy was above 95%.

-   **Efficiency**: computational time was 16 minutes (7 minutes for registration, and 8 minutes for operator supervision).

Manual tracing average inter and intra variability was measured below 2% on the three protocols.

In Mazzara *et al.* [18], average intra-observer variability was very large (20%±16%). Average inter-observer variability was also high (28%±12%). They observed that they obtained better average reproducibility in preoperative cases (15% - 24%) than in postoperative cases (27%-32%). In terms of accuracy measurements based on a statistical ground truth, manual tracing provided 85%±7%, KNN provided 56%±6% and KG method provided 52%±7%, while FP values were 8%±11% for the KNN method, 8%±8% for the KG method and 17%±11%, showing a tendency of the segmentation method to underestimate tumor volume. The authors concluded that the segmentation accuracy was within the manual tracing variability range. It is important to note that the accuracy of the automated methods could not have exceeded the one of the manual tracing with the proposed accuracy measurements with a statistical ground truth. Average computational time for the KNN approach was 30 minutes, including training time. For the KG approach, preparing MRI scans for segmentation required 90 minutes of operator time, and segmentation time required 30 minutes. In Beyer *et al.* [19], the same group showed that expert physician reference volumes were irradiated within the same level of conformity when using radiation plans generating from automated segmented contours.

In Zou *et al.* [20], high AUC and DSC measures were obtained on 2D slices, from nine cases, but they observed that the recommended optimal threshold for the probability values seemed to be case- and task metric-dependent, largely depending on the clinical goal of the segmentation task. For example, AUC was suited for overall accuracy while MI was better suited for longitudinal studies of the tumor's evolution.

In Prastawa *et al.* [21], inter-observer variability was very variable, between 59% and 89% of area overlap. Average surface distances remained below 2mm while the Hausdorff distance reached 13mm for one case. Surface overlaps with the automated segmentation method were comparable, between 68% and 80%, the Hausdorff distance reached 18mm and the average distance 4mm. Computa-tional time was around 90 minutes per volume.

Summary

Based on the literature review of tumor segmentation from MRI data it is extremely difficult to conclude on a best available method. Obviously, automation and minimal user supervision as in [9], are desirable. Obviously, atlas-based methods are limited to tumors without any mass effect. To cover the range of sophistication in segmentation methods, simple seed growing showed very poor reproducibility, while integration of multispectral MRI data, from several protocols, seems critical to mimic visual interpretation of the tumor borders from neurologists.

Recent methods have included a great amount of interactive manual supervision of the segmentation process, reflecting multiple observations of high variability from manual or automated tracing of tumors. Moreover, MRI data provides images with varying tumor appearance due to the heterogeneity of the tumor physiology as well as important variations in MRI scanners in terms of image quality.

Related to this factor, a recent paper from Angelini *et al.* [24] proposed a longitudinal method for quantification of low-grade glioma evolution based on non-linear image normalization and direct difference comparison, avoiding the need for specific tumor segmentation. As an alternative, notions of fuzzy segmentation decisions are used in recent works from Khotanlou *et al.* and Dou *et al.* [25, 26].

Future Improvements

- Performing **multispectral image segmentation:** we can believe that classification or deformable models incorporating multispectral data will show superior performance within the next few years.

- **Role of MRI spectroscopy imaging (MRSI):** in a recent study, Stadlbauer *et al.* [22] showed on ten patients with low-grade (I&II) gliomas that abnormal tumor areas on ^{1}H-MRSI exceeded by an average of 24% that area delineated via supervised seed-growing segmentation of T2 hyperintense signal. The study also showed that computation of metabolic ratios on MRSI enabled automated segmentation of tumoral areas. On the other hand, spatial resolution of spectroscopic data is still limited, with a voxel volume close to 1 mL (compared to 0.005 mL in MRI).

- **Different evaluation setups:** It is essential to evaluate brain tumor segmentation methods in a clinical setting. In a study from Pallud *et al.* [27] biopsy samples isolated tumoral cells beyond imaging abnormalities observed on T2 and FLAIR MRI data in 17 patients with low-grade gliomas. In the context of radiation therapy, Beyer *et al.* [19] have recently shown the superiority of knowledge-based segmentation methods in automatically determining brain gross tumor volume (GTV), with higher reproducibility than manual tracing.

REGISTRATION

Introduction

Since numerical simulations are usually computed in a different space from the one encompassing the patient's anatomy, modeling results have to be warped back to match the specific patient geometry. This deformation of the reference space to the patient space is achieved through a Registration algorithm. Registration is a primordial step to compare the actual segmented evolution with virtual simulated growth.

Image registration can be defined as the process of aligning a target image to a source image. It consists in determining the geometrical transformation that maps corresponding structures in the target and the source image. Explicitly the registration algorithm aims at deforming the source images so that similar structures appear at the same location in the images.

Historically, the primary interest of clinicians was the fusion of multimodal images of the same subject. This registration -usually based on rigid transformations- allows the comparison of different images from a single subject and the visualization of multiple modalities at the same spatial location (see Fig. **2** for an example with MR T1 and MR T2). It could also be used to finely analyze the progression over time of an evolving process (such as lesion growing) [28].

Once mono-subject rigid registration algorithms were considered mature, computer scientists started developing non rigid registration methods for multi-subjects registration. Those algorithms aim at finding the deformation that maps different subjects in the same space. Such methods are usually used in combination with atlases. An atlas is an anatomical image of a subject coupled with a second image. This associated image represents segmented structures of interest, local diffusion properties or label probabilities. The registration algorithm is used to compute the deformation field from the anatomical image of the atlas to the image of the subject. This deformation field is then applied to the property image, so that this information can be mapped to the patient's anatomy.

Methods

Registration Algorithm for Healthy Subject

The majority of registration algorithms share the same three components. First the type of transformation, which controls the geometric flexibility of the algorithm, needs to be defined. The transformation $T()$ defines the mathematical formulation that relates a point X in the first (floating) image and a point X' in the second (fixed) image: $X' = T(X)$.
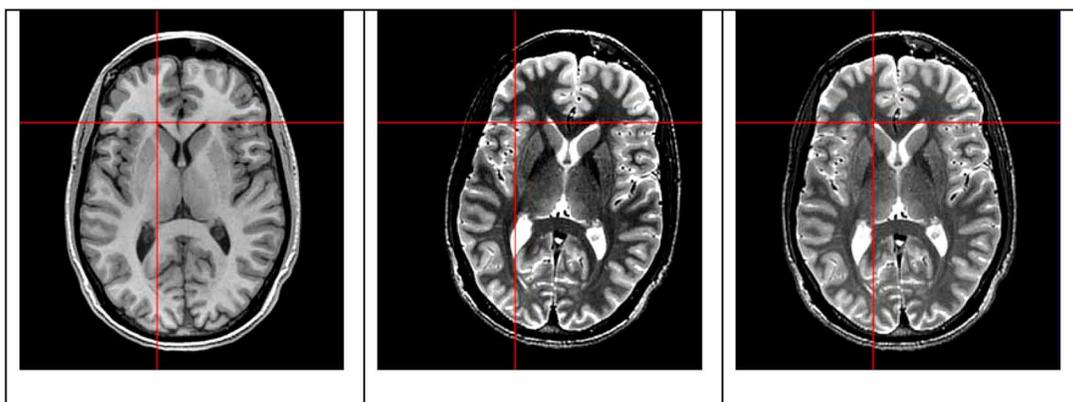
Fig. (2). Rigid registration of MR T1 (left) and T2 (middle) images of the same patient. The registered T2 image (right) now have homologous structures displayed at the same location in the image.

Usually, the transformation is assumed to be rigid between images of the same subject. This transformation is parameterized by 6 degrees of freedom: $X' = RX + D$, 3 for translation (vector *D*) and 3 for rotation (matrix *R*). Different subjects are registered using a non-rigid transformation. Various transformation models have been proposed in the literature. These models were chosen for their mathematical simplicity, and none of them relies on *a priori* brain variability. However, they allow for multi-resolution registration, meaning that the user could define the level of complexity of the transformation through a specified number of degrees of freedom. Most common transformation models are splines [29], cosine basis [30], tetrahedral mesh [31, 32], multi-affine [33-35] and free (one vector per voxel) [36-38].

The second component of a registration algorithm is a similarity metric between the two images *I* and *I'*. This metric will depend on the assumptions made on image intensities distribution for similar structures in the two images [39, 40]. For instance, if similar structures are considered to have homologous intensities (as for mono-modal registration), the appropriate metric will be the sum of squared intensities: $\sum \left(I(X) - I(X') \right)^2$. If the relationship between the two intensities is supposed linear (as for mono modal registration, but from different MR scanners), the correlation coefficient should be used:

$$\frac{\sum (I(X) - \bar{I})(I'(X') - \bar{I'})}{\sum I(X) I'(X) - \overline{\bar{I'}\bar{I}}}, \text{ where } \bar{I} \text{ defines the average}$$

intensity of image *I*. For more complex and probabilistic relationship, mutual information is the natural choice [41, 42]. Because of its adaptability to multi-modal images, mutual information is now the default similarity measure used in most registration algorithms.

Usually, the transformation model does not explicitly impose the smoothness of the deformation. This smoothness is then ensured through a regularization component in the global registration formulation. Popular smoothing energies rely on second order derivative of the displacement field like thin plates splines [29] or Laplacian [37] and continuum mechanics based energy [31]. It has recently been proposed to take into account inter-subject variability to constrain the regularization energy [35, 43].

These components are integrated into the registration algorithm through an optimization process based on an energy formulation. The algorithm tries to estimate the parameters of the transformation that minimizes an energy composed of the similarity criteria and the regularization energy.

Registration Algorithm in the Presence of Tumor

The presence of a tumor in images represents a challenge for registration algorithms: the assumption of an intensity relationship between homologous structures does not hold, as the presence of a tumor in the image changes the MR signal in invaded areas. The initial approach to tackle this problem consisted in discarding the tumor region from the

similarity criteria [31, 44]. The displacement of the tumor region is then guided on its boundary by the surrounding healthy structures, and the regularization criterion ensures the continuity of the displacement inside the tumoral region. The registration algorithm could be considered as passive in the tumor region, in the sense that it follows the motion of its environment.

An alternative approach consists in defining a model of the tumor-induced displacement. Initially, a statistical method was proposed, trained on a dataset of possible displacements [45]. These methods make the assumption that similar tumors have similar deformation patterns; so that the deformation induced by a new tumor can be deduced with a linear combination of other displacement induced by tumors at the same location. However, these statistical methods need a large number of patient images and the corresponding true displacements. To overcome this problem, it was proposed to train the statistical model on tumor growth simulations [46, 47]. Numerical simulation could then be used to train the model for any tumor and at any location in the brain. In this case, the registration algorithm could be considered as active, in the sense that it tries to find an appropriate model to fit to observed growth in the image. Nevertheless, using such registration methods for tumor growth simulation in an atlas image is highly controversial.

Clinical Validation Methods

Validation of non-rigid registration algorithms is a research topic on its own. Because it is a relatively recent research domain, emphasize has mostly been put on the development of new algorithms. In addition, validation methods highly depend on the application of the registration algorithm.

When the displacement observed between the two images to be registered is due to a mechanical deformation (for example: a tumor induced mass effect), comparison of manually identified landmarks seems to be the standard method of evaluation. Points on similar structures are manually identified in the moving image (P_m) and the reference image (P_r). The computed displacement of the moving point $T(P_m)$ is then compared to the manually evaluated displacement $D = (P_m - P_r)$ and the error is defined as the sum of the squared differences: $\sum (T(P_m) - D)^2$. This method is usually considered as the standard method for mono-subject registration validation. It is however subject to inter-expert variability and human error in identifying corresponding points.

In the context of multi-subject registration, the objective is different: the deformation now represents the variability between the two individuals and is more difficult to evaluate. In addition, distinguishing between variability-induced deformation and tumor-induced deformation is challenging for pathological cases. However, the registration algorithm is usually used in combination with an atlas to import the associated property image in the patient space. It then makes more sense to validate the quality of this final matching. For example, the registration of a segmented atlas will be evaluated on the quality of the segmentation on the new image [48].

Recently, new methods for validation based on atlas building have been proposed [49, 50]. In [50], multiple registration algorithms are used to build an average diffusion tensor image. The quality of the registration is then evaluated on the variability (variance) around the mean image: the better the registration, the closer every deformed diffusion tensor images should be to the mean. In [49], an atlas of resection probability is build. The performance of the registration algorithm is then evaluated using the predictive power of the atlas on new cases.

Results

In [31], the registration error is evaluated using manually identified landmarks. Results are obtained on mono-subject registration of patients in the course of tumor resection. Images are acquired with an open MR system. The average registration error measured on 6 cases and 54 landmarks is bellow 1mm, and the maximum error bellow 3mm. It is mentioned that an accuracy decrease is observed in the regions very close to the tumor. The algorithm is not distributed.

In [46], the impact of the active tumor model is evaluated on atlas-to-subject image registration of real and simulated tumors. The benefit of the active model over the usual passive registration algorithm is evaluated using landmark errors: the average accuracy improvement is 58% for real tumors, and 39% for simulated tumors. The average error is bellow 4mm in both cases. Part of the algorithm is distributed [51], but it does not include the active tumor model.

A first step towards the validation of registration algorithms for healthy subject has been proposed in [50]. The quantitative study of Sanchez *et al.* shows that best registration algorithms available for DTI mapping (where splines and demons are compared) do not have statistical difference in their respective measured errors.

In [49], the software LesionMask [52] (available for download) is used to build a tumor resection probability atlas. Retrospective evaluation of the predictive power of the atlas for the preoperative classification of subtotal versus partial tumor resection shows a correct prediction in 82% of cases. Such method could be used in the very near future to assess the relative performance of registration algorithms in the presence of tumors.

Summary

To summarize, different registration methods have been proposed in the literature to tackle the problem raised by the presence of tumors in the images. There is however today no consensus on a satisfactory registration method able to map a patient image with a tumor on an (healthy) atlas. Indeed, validation of registration algorithm seems to be very application dependant. Most promising validation methods seem to be based on atlas building [49, 50].

The mathematical formulation of the problem seems to be now better understood. It is then expected that recent efforts in the open source community [53], implementing most popular algorithms and making them available, will improve the research efforts in comparing and validating

these algorithms. To the best of our knowledge, LesionMask [52] is the only non-rigid registration algorithm available that could be used in the presence of tumors. More recent tools based on ITK have been proposed for the registration of healthy subjects [54, 55]. It is probable that those algorithms will soon be adapted to account for the tumor deformation in the images.

IN SILICO GLIOMA GROWTH

The mathematical description of tumor growth can be formulated at different spatial scales: either one tries to simulate the growth at the cellular level (cellular automata), or one defines on a macroscopic scale how the tumor density will evolve (with partial differential equations). In both cases, the models attempt to predict the mathematical law of the tumor growth.

The cellular automata approach has been proposed for different tumors, including high grade gliomas.[56-58]. The rules of division and invasion are the key elements of this approach. Since these models describe the tumor growth at the microscopic scale, their prediction could be also validated in clinical practice by microscopic histological analysis of tumor samples (spatial correlation in graph cells [59]) or *in vitro*, with dynamic study of glial cells migrations [60]. On the contrary, the partial differential equations approach does not tell anything about the spatial ordering of the tumor cells: it simulates a coarse-grained cell density.

There is of course a link between these two spatial scales of description: partial differential equations can be solved by stochastic methods, mimicking the cellular phenomena (Gaussian random walk for example). Fractals models [61-63] also propose to make a link between these two scales. By determining the fractal dimension of the tumor boundaries, a scaling analysis gives critical exponents from which dynamics law can be inferred. This approach seems to be in good agreement with both *in vivo* and *in vitro* data of bulky tumors, but its relevance for infiltrative tumors like glioma is not established.

We will now focus on the macroscopic model of tumor growth, which seems to be the most appropriate for comparison with clinical MRI data, and is the most widely used. Of interest is a phenomenological approach very recently proposed [64]: in this case, machine learning algorithms are used to determine, based on observed growth patterns, a 3D probabilistic classification of diffusion patterns. This promising approach, yet preliminary, will help in the future to refine the deterministic model of proliferation-diffusion that we will now present.

Methods

Proliferation-Diffusion Model for Tumor Cell Density

This modeling framework is based on an equation governing (coarse-grained) tumor cell density (denoted c, expressed in cells/mm^3) increase by time unit [4]. The generic form of the equation is the following:

$$\frac{\partial c}{\partial t} = \rho c + \nabla.(D\nabla c), \qquad (13)$$

stating that, at a given spatial location, new tumor cells appear either by division (proliferation), or by moving from

a close location (diffusion). The term diffusion is used here as « passive diffusion » (random walk of cells, also called brownian or fickian process). It does not model cell migration (i.e. active motion of cells), nor invasion (i.e. destruction of the extra-cellular matrix). The two parameters introduced in the equation are the proliferation coefficient (ρ), and the diffusion coefficient (*D*). Several mathematical variants have been proposed for spatial variations of ρ and *D*.

Exponential proliferation (with ρ constant), stating that cellular division obeys a cycle, of doubling time (ln(2)/ρ), is the most widely used. This basic model can be refined, by decreasing the proliferation parameter ρ in areas of high cellular density (e.g. logistic growth, Gompertz growth) [4].

Initially, diffusion was considered as homogeneous in the whole brain. A first improvement consisted in introducing a diffusion parameter higher in white matter than in gray matter [3]. Indeed, standard virtual MRI atlases of the brain are available from averaging of healthy volunteers. In such segmented atlases, each voxel is labelled in reference to its category (gray matter, white matter, or subarachnoidal space i.e. brain boundaries). This allows to match the numerical implementation of the equations to the specific geometry of the brain, with distinct diffusion coefficient values in each compartment.

Moreover, it is a well known fact that glioma cells migrate more easily along the direction of white matter tracts [65-67]. This observation has been included in the model by assuming an anisotropic cell diffusion tensor, deduced from water diffusion tensor [4]. The virtual brain used for numerical simulations now takes into account the 3D architecture of white matter : it is a segmented anatomical image, co-registered with its diffusion tensor image (DTI). Unfortunately, there is no standard atlas for DTI, since the average of diffusion tensor images from a population of healthy volunteers is still an area of medical imaging research.

Finally, biomechanical properties of the brain should also be modeled. This has been already done for high-grade gliomas [68]. Even if it is more frequently present for high grade gliomas, some grade II glioma exhibit mass effect, with deformation of surrounding brain parenchyma, and the method should also be extended to low-grade gliomas.

Clinical Validation Methods

The modeling equations can be numerically solved, on virtual brain atlases. Results are typically represented by serial color maps, where the color reflects the tumor cell density.

The clinical validation of these simulations – an issue of utmost importance – raises two separate problems: (1) which variable of the model can be experimentally measured ? (2) are the values of the parameters *D* and ρ in good agreement with their biological estimations ?

Three-dimensional maps of tumor cell density have never been built from histological specimens, except *in vitro* [69]. Thus, it is not possible to compare directly the results of some simulations with experimental data. The only way is to estimate cell density from MRI. Unfortunately, the link between cell density and MRI signal abnormalities remains largely unknown. The common hypothesis considers a threshold of visibility, estimated only for CT data [70], around 8000 cells/cm^3. For MRI, T2 (FLAIR) sequences are probably more sensitive, but this issue has never been investigated. Both experimental and theoretical studies are thus needed to determine tumor cell density within an heterogeneous tumor [69] and to correlate these results with MRI signal.

The model also gives an estimation of cell density outside the tumor region visible on MRI (with a cell density bellow the threshold value). The possibility to visualize the immerged part of the iceberg is called virtual imaging [71]. In the model, the extent of this non-visible tumor is linked to the ratio *D/ρ*. At the present time, only very few studies have tried to quantify the concentration of isolated tumor cells in these microscopically infiltrative areas. A recent study has shown, using a histo-radiological correlation from serial stereotactic biopsies of untreated grade II glioma, that isolated tumor cells are present far beyond the margins of MRI signal abnormalities [27, 72]. A spatial quantification of the tumor cell density distribution would be more informative for the purpose of model validation, even if it would still be restricted to a limited number of biopsy samplings. More global methods are currently under investigation, like optical imaging at the end of the surgical resection [73-75], or spectroscopy by magnetic resonance (sMR) during pre-operative surgical planning [76].

Thus, it is necessary to assume that the tumor is visible on MRI only if the cell density reaches a threshold value (usually 8000 cells/mm^3). This allows to outline the visible spatial extent of the tumor on the simulations, and to compare with observed tumoral tissues on MRI data. A first simplified approach consists in the measurement of the linear dimensions of the tumor. It can indeed be shown that, under some basic assumptions, an approximate solution of the equation is given by a constant velocity of tumor diameter expansion. The visible tumor front edge is therefore seen as a wave, propagating at a constant speed [77]. Since the velocity is equal to $4\sqrt{D\rho}$, determining the slope of the diametric growth curve thus gives an estimation of the product *Dρ*. In a next step, serial maps provided by numerical simulations can be visually compared with observed tumoral tissues on MRI data. Several simulations can be launched with different sets of parameters (*D, ρ,* anisotropic ratio, starting point) and the one leading to the best fit is selected. Actually, accepted values of *D* and ρ for gliomas of different grades have been estimated by this rough method. A more rigorous validation, solving the inverse problem as explained later, is a complex task, still under development.

The second problem has never been seriously investigated. Once parameter values of the couple (*D, ρ*) have been determined by fitting with simulations, one would indeed test if the values are in good agreement with their experimental measures. Unfortunately, the experimental estimation of ρ seems impossible. Even if counting the

density of cycling cells with biomarkers (MIB-1, Ki67, MCM2) is conceivable, the cell cycle period remains unknown, and consequently, the exact value of ρ cannot be measured. On the contrary, new cellular imaging techniques (using SPIO [78] or quantum dots [79]) could be a mean to track the diffusion of single glioma cells. *In vitro* dynamic study of microscopic slices is an alternative way to measure diffusion properties of glioma cells [60]. These methods would allow a direct estimation of D and are promising tools in the understanding of the tumor spreading behavior.

Results

In silico growth of glioma was initially developed for HGG [1, 2, 70]. These first attempts were essentially conceptual, since simulations were performed in 2D, with boundaries including the ventricles, but without white/gray matter differentiation. Nevertheless, this allowed to propose values for D and ρ, for high-grade gliomas as well as low-grade gliomas [2]. With the advent of MRI, Swanson *et al.* have paved the way to simulate tumor growth in a 3D virtual brain, modeling anatomical structures [3]. They introduced 3D segmentation of ventricles, sub-arachnoid spaces, white and gray matter. This allowed to achieve simulations closer to reality, with the additional possibility to visualize the invasive part of the tumor, unseen on MRI data. However, as already stated before, there is no validation of this virtual imaging setup. Clinical validation of simulations for HGG has been achieved only statistically with survival curves of actual and virtual cases after different extents of resections quite similar [77].

Actually, HGG are highly heterogeneous tumors: they eventually include necrosis, contrast enhancement areas, edema, infiltration, and significant mass effect. These diverse biological phenomena increase the complexity related to modeling HGG. The case of LGG could be more favourable for modeling purpose. Indeed, during the "low-grade" phase, biological behaviour seems to be relatively constant. Moreover, since therapeutic abstention was the rule, patients were clinically and radiologically monitored over several years, and serial MRI data sets before any treatment were available for comparison with model predictions.

In a first study, it has been shown, on a series of 27 patients, that LGG exhibited, as predicted by the model, a constant velocity of their diametric expansion, with an average of 4 mm/year [80]. This result confirmed the values of D and ρ initially proposed [2]: typical values are $D = 7.5 \times 10^{-3} \text{ mm}^2 \text{ day}^{-1}$ and $\rho = 0.0012 \text{ day}^{-1}$.

In a second step, the shapes of simulated tumors were compared, for the first time, with the ones clinically observed on serial MRI data. Anisotropic growth along white matter fibers was introduced to obtain a better fit [4] : a new parameter was introduced to quantify this anisotropic diffusion of glioma cells. This parameter could be very useful to identify, from micro-arrays analysis, new proteins involved in the migration along axons.

Finally, an indirect way to validate a model is to test if the dynamic parameters are of prognosis value. This has been recently proved on a series of 143 patients: patients

with diametric expansion velocity greater than 8 mm/year had a prognosis similar to those harbouring a HGG, and this despite an histological diagnosis of LGG [81].

Future Improvements

The inverse problem of finding the optimal model parameters from observations is the key ingredient in adapting the general tumor growth models to specific patient cases. This adaptation is necessary to utilize theoretical models in clinical practice. One can think of two different and coupled inverse problems in the case of tumor growth: solving for reaction-diffusion parameters and solving for the deformation field caused by the tumor. Naturally, both problems are based on clinically available observations. Considering the scale of the growth models explained previously, medical images form the most important class of observations.

In the case of identifying the mass effect of the tumor, estimating the induced deformation field, using medical images, has received some attention from the medical image analysis community [45]. This inverse problem becomes especially important in the application of atlas-based segmentation. In using the comparison between healthy anatomy and the pathology-bearing brain, these approaches face the problem of distinguishing deformation caused by the tumor from inter-subject variability.

On the other hand, the inverse problem on the reaction-diffusion model explained in the previous section has received almost no attention yet. This problem consists of identifying several parameters from time sequences of medical images. One can list these parameters as:

- D_w, D_g : Diffusion tensors in the white and in the gray matter. These tensors can consist of different numbers of parameters based on their construction. The state of the art for this construction uses two different parameters: one for the diffusion along the fibers, and one for the diffusion across the fibers and in the gray matter.

- ρ: The proliferation rate. Although this parameter is microscopic and cannot be determined experimentally, one can find an average value that would force the model to fit the observation.

- X_0: The initial point. One can retrieve the point where the tumor has started by finding the other parameters and running the model backwards.

- t_0: The initial time. Assuming the tumor growth is homogeneous in time, one can find the onset time of the tumor.

Although we can list these parameters, it does not mean that we can identify them for any specific case. Number of observations strictly limits the number of parameters we can extract confidently. In other words, more parameters can be identified if one has a larger time sequence of images with regular follow-ups. As a first attempt to solve this inverse problem, Konukoglu *et al.* proposed a method to solve for speed of tumor invasion in white and in gray matter separately using two images taken at different time instances [82]. Their method was based on the asymptotic properties

of the reaction-diffusion model formulating the evolution of the enhanced signal regions in MR images.

Perspectives

The dream of the physicians would be to click a point and run a software tool allowing to determine the individual dynamics parameters of the tumor for each patient from two (or more) successive MRI data sets. To pass from dream to reality, the process must include the three steps previously described: determination of real tumor growth by (semi-) automated segmentation, registration with the virtual brain, and resolution of the inverse problem.

This formidable achievement would open the door to tailor therapies to individual tumor dynamics [83]:

- The product of parameter values $D\rho$ will lead to a dynamical classification of gliomas. For example, given the worst prognosis associated with a LGG growing faster than 8 mm/year [81], it should be recommended to treat the patient as if his tumor were of higher grade.

- The value of the parameters ratio D/ρ, which controls the extent of non-visible tumor has also important therapeutic consequences. For example, it is expected that the benefit of gross total resection for tumors with high values of D/ρ is limited. In this case, a lot of isolated tumor cells are left even after a radiologically complete resection. On the contrary, these tumors would greatly benefit of a supra-radical resection.

- For radiotherapy, the tumor margins are at present time defined 2 cm beyond the MRI signal abnormalities borders. The simulations could propose a more precise contour for isodose regions, adapted to each patient [82].

- Chemotherapy responses could be evaluated by quantifying how much ρ is decreased by the cytotoxic drug. Each patient can serve as its own control, and the therapeutic response can be defined by the difference of dynamic parameters (after treatment – before treatment). For example, a patient with a stable lesion under chemotherapy should be considered as a responder if his initial growth rate was about 4 mm/year [84]. By the way, the model could explain why chemotherapy could be efficient even if the visible tumor remains unchanged, corroborating the well known feeling that some responses are only clinical and not radiological.

Finally, such dynamic parameters would also be of utmost interest for fundamental research. There is growing evidence that such tumors are complex systems, implying their analyze at nanoscopic, microscopic and macroscopic scales. The dynamic parameters previously described characterize the tumor at the macroscopic scale and they will allow correlations with pathological examinations at the microscopic scale and molecular profiles at the nanoscopic scale, as a first step toward a multi-scale approach of grade II gliomas.

At the present stage, the clinical and mathematical analysis of glioma dynamics is still in its infancy; but given the advances in both clinical imaging and computer modelling, it should play a prominent role in the future neuro-oncological practice.

ABBREVIATIONS

AUC	=	Area under the receiver operating curve
CR	=	Correspondence ratio
CSF	=	Cerebrospinal fluid
CV	=	Coefficient of variation
DSC	=	Dice similarity coefficient
DTI	=	Diffusion tensor imaging
FC	=	Fuzzy connectedness
FN	=	False negative
FNVF	=	False negative volume fraction
FP	=	False positive
FPVF	=	False positive volume fraction
FCM	=	Fuzzy C-means
GM	=	Gray matter
HGG	=	High grade glioma (i.e. WHO grade III & IV glioma)
KB	=	Knowledge-based
kNN	=	k Nearest-neighbour
LGG	=	Low grade glioma (i.e. WHO grade II glioma)
MI	=	Mutual information
MRI	=	Magnetic resonance imaging
PD	=	Proton density MRI data
MRSI	=	MRI spectroscopy imaging
PM	=	Percent match
SA	=	Segmentation accuracy
SD	=	Standard deviation
T1	=	T1 MRI data
T1E	=	T1 MRI data enhanced (with the use of a contrast agent)
T2	=	T2 MRI data
TP	=	True positive
WM	=	White matter
GTV	=	Growth tumor volume

REFERENCES

[1] Tracqui P, Cruywagen GC, Woodward DE, *et al*. A mathematical model of glioma growth: the effect of chemotherapy on spatio-temporal growth. Cell Prolif 1995; 28(1): 17-31.

[2] Woodward DE, Cook J, Tracqui P, *et al*. A mathematical model of glioma growth: the effect of extent of surgical resection. Cell Prolif 1996; 29(6): 269-88.

[3] Swanson KR, Alvord EC, Jr, Murray JD. A quantitative model for differential motility of gliomas in grey and white matter. Cell Prolif 2000; 33(5): 317-29.

[4] Jbabdi S, Mandonnet E, Duffau H, *et al*. Simulation of anisotropic growth of low-grade gliomas using diffusion tensor imaging. Magn Reson Med 2005; 54(3): 616-24.

[5]   Gibbs P, Buckley DL, Blackband SJ, *et al.* Tumour volume determination from MR images by morphological segmentation. Phys Med Biol 1996; 41(11): 2437-46.

[6]   Kennedy DN, Filipek PA, Caviness V. Anatomic segmentation and volumetric calculations in nuclear magnetic resonance imaging. IEEE Trans Med Imag 1989; 8: 1-7.

[7]   Letteboer MMJ, Olsen OF, Dam EB, *et al.* Segmentation of tumors in magnetic resonance brain images using an interactive multiscale watershed algorithm. Acad Radiol 2004; 11: 1125-1138.

[8]   Droske M, Meyer B, Rumpf M, *et al.* An adaptive level set method for interactive segmentation of intracranial tumors. Neurol Res 2005; 27(4): 363-70.

[9]   Vaidyanathan M, Clarke LP, Velthuizen R.P, *et al.* Comparison of supervised MRI segmentation methods for tumor volume determination during therapy. Magn Reson Imaging 1995; 13(5): 719-28.

[10]   Clark M, Hall LO, Goldgof, DB, *et al.* Automatic tumor segmen-tation using knowledge-based techniques. IEEE Trans Med Imag 1998; 17(2): 187-201.

[11]   Kaus MR, Warfield SK, Nabavi A, *et al.* Automated segmentation of MR images of brain tumors. Radiology 2001; 218(2): 586-91.

[12]   Kaus M.R, Warfield SK, Jolesz FA, *et al.* Segmentation of meningiomas and low grade gliomas in MRI. in International Conference on Medical Image Computing and Computer Assisted Intervention. 1999. Cambridge, UK.

[13]   Warfield SK, Kaus MR, Jolesz FA, *et al.* Adaptive template moderated spatially varying statistical classification. in Interna-tional Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI). 1998. Cambridge, MA, USA.

[14]   Moonis G, Liu J, Udupa JK, *et al.* Estimation of tumor volume with fuzzy-connectedness segmentation of MR images. AJNR Am J Neuroradiol 2002; 23(3): 356-63.

[15]   Nyul LG, Udupa JK. On standardizing the MR image intensity scale. Magn Reson Med 1999; 42: 1072-1081.

[16]   Liu J, Udupa JK, Odhner D, *et al.* A system for brain tumor volume estimation via MR imaging and fuzzy connectedness. Comput Med Imaging Graph 2005; 29(1): 21-34.

[17]   Fletcher-Heath LM, Hall LO, Goldgof DB, *et al.* Automatic segmentation of non-enhancing brain tumors in magnetic resonance images. Artif Intell Med 2001; 21(1-3): 43-63.

[18]   Mazzara GP, Velthuizen RP, Pearlman JL, *et al.* Brain tumor target volume determination for radiation treatment planning through automated MRI segmentation. Int J Radiat Oncol Biol Phys 2004; 59(1): 300-12.

[19]   Beyer GP, Velthuizen RP, Murtagh FR, *et al.* Technical aspects and evaluation methodology for the application of two automated brain MRI tumor segmentation methods in radiation therapy planning. Magn Reson Imaging 2006; 24(9): 1167-78.

[20]   Zou KH, Wells WM, 3rd, Kikinis R, *et al.* Three validation metrics for automated probabilistic image segmentation of brain tumours. Stat Med 2004; 23(8): 1259-82.

[21]   Prastawa M, Bullitt E, Ho S, *et al.* A brain tumor segmentation framework based on outlier detection. Med Image Anal 2004; 8: 275-283.

[22]   Stadlbauer A, Moser E, Gruber S, *et al.* Improved delineation of brain tumors: an automated method for segmentation based on pathologic changes of 1H-MRSI metabolites in gliomas. Neuroimage 2004; 23(2): 454-61.

[23]   Warfield SK, Zou KH, Wells WM. Simultaneous truth and performance level estimation (STAPLE): an algorithm for the validation of image segmentation. IEEE Trans Med Imag 2004; 23(7): 903 -921.

[24]   Angelini E, Atif J, Delon J, *et al.* Detection of glioma evolution on longitudinal MRI studies. in International Symposium on Biomedical Imaging (ISBI). 2007. Arlington USA.

[25]   Khotanlou H, Colliot O, Atif J, *et al.* 3D brain tumor segmentation in MRI using fuzzy classification, symmetry analysis and spatially constrained deformable models. Fuzzy Sets and Systems 2007(in press).

[26]   Dou W, Ruan S, Chen Y, *et al.* A framework of fuzzy information fusion for segmentation of brain tumor tissues on MR images. Image and Vision Computing 2007; 25: 164-172.

[27]   Pallud J, Devaux B, Nataf F, *et al.* [Spatial delimitation of low grade oligodendrogliomas]. Neurochirurgie 2005; 51(3-4 Pt 2): 254-9.

[28]   Rey D, Subsol G, Delingette H, *et al.* Automatic detection and segmentation of evolving processes in 3D medical images: Application to multiple sclerosis. Med Image Anal 2002; 6(2): 163-79.

[29]   Rueckert D, Sonoda LI, Hayes C, *et al.* Nonrigid registration using free-form deformations: application to breast MR images. IEEE Trans Med Imaging 1999; 18(8): 712-21.

[30]   Kiebel SJ, Ashburner J, Poline JB, *et al.* MRI and PET coregistration--a cross validation of statistical parametric mapping and automated image registration. Neuroimage 1997; 5(4 Pt 1): 271-9.

[31]   Clatz O, Delingette H, Talos IF, *et al.* Robust nonrigid registration to capture brain shift from intraoperative MRI. IEEE Trans Med Imaging 2005; 24(11): 1417-27.

[32]   Ferrant M, Nabavi A, Macq B, *et al.* Serial registration of intraoperative MR images of the brain. Med Image Anal 2002; 6(4): 337-59.

[33]   Pitiot A, Bardinet E, Thompson PM, *et al.* Piecewise affine registration of biological images for volume reconstruction. Med Image Anal 2006; 10(3): 465-83.

[34]   Arsigny V, Pennec X, Ayache N. Polyrigid and polyaffine transformations: a novel geometrical tool to deal with non-rigid deformations - application to the registration of histological slices. Med Image Anal 2005; 9(6): 507-23.

[35]   Commowick O, Stefanescu R, Fillard P, *et al.* Incorporating statistical measures of anatomical variability in atlas-to-subject registration for conformal brain radiotherapy. Med Image Comput Comput Assist Interv Int Conf Med Image Comput Comput Assist Interv 2005; 8(Pt 2): 927-34.

[36]   Stefanescu R, Pennec X, Ayache N. Grid powered nonlinear image registration with locally adaptive regularization. Med Image Anal 2004; 8(3): 325-42.

[37]   Thirion JP. Image matching as a diffusion process: an analogy with Maxwell's demons. Med Image Anal 1998; 2(3): 243-60.

[38]   Vercauteren T, Pennec X, Malis E, *et al.* Insight into efficient image registration techniques and the Demons algorithm. In Infor Proc Med Imag (IPMI'07). 2007.

[39]   Roche A, Malandain G, Ayache N, *et al.* Towards a better comprehension of similarity measures used in medical image registration., in Int. Conf. on Medical Image Computing and Computer-Assisted Intervention (MICCAI). 1999. p. 555-566.

[40]   Zollei L, Fisher JW, Wells WM. A unified statistical and information theoretic framework for multi-modal image registration. Inf Process Med Imaging 2003; 18: 366-77.

[41]   Maes F, Collignon A, Vandermeulen D, *et al.* Multimodality image registration by maximization of mutual information. IEEE Trans Med Imaging 1997; 16(2): 187-98.

[42]   Wells WM, 3rd, Viola P, Atsumi H, *et al.* Multi-modal volume registration by maximization of mutual information. Med Image Anal 1996; 1(1): 35-51.

[43]   Pennec X, Stefanescu R, Arsigny V, *et al.* Riemannian elasticity: a statistical regularization framework for non-linear registration. Med Image Comput Comput Assist Interv Int Conf Med Image Comput Comput Assist Interv 2005; 8(Pt 2): 943-50.

[44]   Stefanescu R, Commowick O, Malandain G, *et al.* Non-rigid atlas to subject registration with pathologies for conformal brain radiotherapy. In Med Image Comput Comput Assist Interv Int Conf Med Image Comput Comput Assist Interv. 2005; 8(Pt 2): 927-34.

[45]   Davatzikos C, Shen D, Mohamed A, *et al.* A framework for predictive modeling of anatomical deformations. IEEE Trans Med Imaging 2001; 20(8): 836-43.

[46]   Mohamed A, Zacharaki EI, Shen, D, *et al.* Deformable registration of brain tumor images via a statistical model of tumor-induced deformation. Med Image Anal 2006; 10(5): 752-63.

[47]   Prastawa M, Bullitt E, Gerig G. Synthetic ground truth for validation of brain tumor MRI segmentation. Med Image Comput Comput Assist Interv Int Conf Med Image Comput Comput Assist Interv 2005; 8(Pt 1): 26-33.

[48]   Bondiau PY, Malandain G, Chanalet S, *et al.* Atlas-based automatic segmentation of MR images: validation study on the brainstem in radiotherapy context. Int J Radiat Oncol Biol Phys 2005; 61(1): 289-98.

[49]   Mandonnet E, Jbabdi S, Taillandier L, *et al.* Preoperative estimation of residual volume for WHO grade II glioma resected with intraoperative functional mapping. Neuro Oncol 2007; 9(1): 63-9.

[50] Sanchez Castro F, Clatz O, Dauguet J, *et al*. Evaluation of brain image nonrigid registration algorithms based on Log-Euclidean MR-DTI consistency measures. in IEEE International Symposium on Biomedical Imaging (ISBI). 2007.

[51] Section for Biomedical Image Analysis, Upenn.

[52] LesionMask - MRIcro - Chris Rorden.

[53] Insight Segmentation and Registration Toolkit.

[54] MedINRIA - Asclepios Research Project Team.

[55] Slicer3 - NAMIC.

[56] Kansal AR, Torquato S, Chiocca EA, *et al*. Emergence of a subpopulation in a computational model of tumor growth. J Theor Biol 2000; 207(3): 431-41.

[57] Kansal AR., Torquato S, Harsh GI, *et al*. Simulated brain tumor growth dynamics using a three-dimensional cellular automaton. J Theor Biol 2000; 203(4): 367-82.

[58] Kansal AR, Torquato S, Harsh IG, *et al*. Cellular automaton of idealized brain tumor growth dynamics. Biosystems 2000; 55(1-3): 119-27.

[59] Gunduz C, Yener B, Gultekin SH. The cell graphs of cancer. Bioinformatics 2004; 20 Suppl 1: I145-I151.

[60] Palfi S, Swanson KR., De Bouard S, *et al*. Correlation of *in vitro* infiltration with glioma histological type in organotypic brain slices. Br J Cancer 2004; 91(4): 745-52.

[61] Bru A, Albertos S, Luis Subiza, J, *et al*. The universal dynamics of tumor growth. Biophys J 2003; 85(5): 2948-61.

[62] Guiot C, Degiorgis PG, Delsanto PP, *et al*. Does tumor growth follow a "universal law"? J Theor Biol 2003; 225(2): 147-51.

[63] Guiot C, Delsanto PP, Carpinteri A, *et al*. The dynamic evolution of the power exponent in a universal growth model of tumors. J Theor Biol 2006; 240(3): 459-63.

[64] Morris M, Greiner R, Sander J, *et al*. Learning a classification-based glioma growth model using MRI data. J Comput 2006; 1(7): 21-31.

[65] Giese A, Bjerkvig R, Berens ME, *et al*. Cost of migration: invasion of malignant gliomas and implications for treatment. J Clin Oncol 2003; 21(8): 1624-36.

[66] Mandonnet E, Capelle L, Duffau H. Extension of paralimbic low grade gliomas: toward an anatomical classification based on white matter invasion patterns. J Neurooncol 2006; 78(2): 179-85.

[67] Scherer H. The forms of growth in gliomas and their practical significance. Brain 1940; 63: 1-35.

[68] Clatz O, Sermesant M, Bondiau PY, *et al*. Realistic simulation of the 3-D growth of brain tumors in MR images coupling diffusion with biomechanical deformation. IEEE Trans Med Imaging 2005; 24(10): 1334-46.

[69] Mojsilovic A, Rogowitz B, Gomes J, *et al*. Analysis and 3D reconstruction of heterogeneity in malignant brain tumors: an interdisciplinary case study using a novel computational visualization approach. Anal Quant Cytol Histol 2002; 24(3): 125-33.

[70] Burgess PK, Kulesa PM, Murray JD, *et al*. The interaction of growth rates and diffusion coefficients in a three-dimensional mathematical model of gliomas. J Neuropathol Exp Neurol 1997; 56(6): 704-13.

[71] Swanson KR, Alvord EC, Jr.,Murray JD. Virtual brain tumours (gliomas) enhance the reality of medical imaging and highlight inadequacies of current therapy. Br J Cancer 2002; 86(1): 14-8.

[72] Watanabe M, Tanaka R, Takeda N. Magnetic resonance imaging and histopathology of cerebral gliomas. Neuroradiology 1992; 34(6): 463-9.

[73] Kircher MF, Mahmood U, King RS, *et al*. A multimodal nanoparticle for preoperative magnetic resonance imaging and intraoperative optical brain tumor delineation. Cancer Res 2003; 63(23): 8122-5.

[74] Leppert J, Krajewski J, Kantelhardt SR, *et al*. Multiphoton excitation of autofluorescence for microscopy of glioma tissue. Neurosurgery 2006; 58(4): 759-67; discussion 759-67.

[75] Toms SA, Lin WC, Weil RJ, *et al*. Intraoperative optical spectroscopy identifies infiltrating glioma margins with high sensitivity. Neurosurgery 2005; 57(4 Suppl): 382-91; discussion 382-91.

[76] Ganslandt O, Stadlbauer A, Fahlbusch R, *et al*. Proton magnetic resonance spectroscopic imaging integrated into image-guided surgery: correlation to standard magnetic resonance imaging and tumor cell density. Neurosurgery 2005; 56(2 Suppl): 291-8; discussion 291-8.

[77] Harpold HL, Alvord EC, Jr.,Swanson KR. The evolution of mathematical modeling of glioma proliferation and invasion. J Neuropathol Exp Neurol 2007; 66(1): 1-9.

[78] Dunning MD, Kettunen MI, Ffrench Constant C, *et al*. Magnetic resonance imaging of functional Schwann cell transplants labelled with magnetic microspheres. Neuroimage 2006; 31(1): 172-80.

[79] Gao X, Yang L, Petros JA, *et al*. *In vivo* molecular and cellular imaging with quantum dots. Curr Opin Biotechnol 2005; 16(1): 63-72.

[80] Mandonnet E, Delattre JY, Tanguy ML, *et al*. Continuous growth of mean tumor diameter in a subset of grade II gliomas. Ann Neurol 2003; 53(4): 524-8.

[81] Pallud J, Mandonnet E, Duffau H, *et al*. Prognostic value of initial magnetic resonance imaging growth rates for World Health Organization grade II gliomas. Ann Neurol 2006; 60(3): 380-3.

[82] Konukoglu E, Clatz O, Bondiau PY, *et al*. Extrapolating tumor invasion margins for physiologically determined radiotherapy regions. International Conference on Medical Image Computing and Computerized Assisted Intervention (MICCAI) 2006; 1: 338-46.

[83] Mandonnet E, Pallud, J, Clatz, O, *et al*. Computational modelling of WHO grade II gliomas: Principles and applications to surgical strategy. Neurosurg Rev 2007; Submitted.

[84] Ricard, D, Kaloshi, G, Amiel-Benouaich, A, *et al*. Dynamic history of low-grade gliomas before and after temozolomide treatment. Ann Neurol 2006; In Press.