



HAL
open science

An a contrario approach for the detection of patient-specific brain perfusion abnormalities with arterial spin labelling

Camille Maumet, Pierre Maurel, Jean-Christophe Ferré, Christian Barillot

► To cite this version:

Camille Maumet, Pierre Maurel, Jean-Christophe Ferré, Christian Barillot. An a contrario approach for the detection of patient-specific brain perfusion abnormalities with arterial spin labelling. *NeuroImage*, 2016, 134, 10.1016/j.neuroimage.2016.03.054 . inserm-01291748v2

HAL Id: inserm-01291748

<https://inserm.hal.science/inserm-01291748v2>

Submitted on 7 Apr 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

An *A Contrario* Approach for the Detection of Patient-Specific Brain Perfusion Abnormalities with Arterial Spin Labelling

Camille Maumet^{a,b,c,d,e,*}, Pierre Maurel^{a,b,c,d}, Jean-Christophe Ferré^{a,b,c,d,f}, Christian Barillot^{a,b,c,d}

^aUniversity of Rennes 1, Faculty of medicine, F-35043 Rennes, France

^bINSERM, U746, F-35042 Rennes, France

^cCNRS, IRISA, UMR 6074, F-35042 Rennes, France

^dInria, VISAGES project-team, F-35042 Rennes, France

^eWarwick Manufacturing Group, University of Warwick, CV4 7AL Coventry, United-Kingdom

^fCHU Rennes, Department of Neuroradiology, F-35033 Rennes, France

Abstract

In this paper, we introduce a new locally multivariate procedure to quantitatively extract voxel-wise patterns of abnormal perfusion in individual patients. This *a contrario* approach uses a multivariate metric from the computer vision community that is suitable to detect abnormalities even in the presence of closeby hypo- and hyper-perfusions. This method takes into account local information without applying Gaussian smoothing to the data. Furthermore, to improve on the standard *a contrario* approach, which assumes white noise, we introduce an updated *a contrario* approach that takes into account the spatial coherency of the noise in the probability estimation.

Validation is undertaken on a dataset of 25 patients diagnosed with brain tumors and 61 healthy volunteers. We show how the *a contrario* approach outperforms the massively univariate General Linear Model usually employed for this type of analysis.

Keywords: Arterial Spin Labelling, Brain tumour, Hypo-perfusion, Hyper-perfusion, *a contrario*, General Linear Model, searchlight approach.

*Corresponding author

Email address: contact@camillemaumet.com (Camille Maumet)

1. Introduction

Arterial Spin Labelling (ASL), a Magnetic Resonance Imaging (MRI) technique introduced in the early 1990's (Detre et al., 1992), provides a measure of the level of perfusion as a quantitative index: the cerebral blood flow (CBF). Contrary to standard perfusion imaging, including Positron Emission Tomography (PET) and Single Photon Emission Computed Tomography (SPECT) in nuclear medicine or Dynamic Susceptibility weighted Contrast (DSC) in MRI, ASL is completely non-invasive and does not require the injection of an exogenous contrast agent. The ASL sequence however suffers from a low SNR, which is still a serious obstacle for its use in clinical practice.

Since the introduction of ASL, and despite its low SNR, a large number of studies have demonstrated its usefulness in identifying patterns of abnormal perfusion at the group level, for example (Pinkham et al., 2011). While group statistical analyses are of tremendous importance in order to understand the general mechanisms underlying a pathology, there is today an increasing interest towards patient-specific analyses (Viviani et al., 2007; Klöppel et al., 2012; Scarpazza et al., 2013) that aim to draw conclusions at the level of a single individual. Inference at the patient level is highly desirable in order to perform a diagnosis or provide a personalised treatment. On top of this, some pathologies are intrinsically not well-suited to perform group voxel-wise analyses. It is for instance the case of brain tumours (Warmuth et al., 2003; Chawla et al., 2007; Weber et al., 2006; Sugahara et al., 2000), stroke (Huck et al., 2012; Ferré et al., 2012; Wheeler et al., 2013), or some types of epilepsy (Colliot et al., 2006; Pendse et al., 2010), that exhibit a different pattern of spatial abnormalities for each patient. In all of these pathologies, measuring the level of perfusion, as an indicator of the well-being of the tissues, has been demonstrated to be a useful information for diagnosis in clinical practice (see for example (Warmuth et al., 2003) for brain tumours, (Wheeler et al., 2013) in stroke and (Pendse et al., 2010; Boscolo Galazzo et al., 2015) for epilepsy).

Very few ASL studies focused on voxelwise quantitative perfusion abnormality detections at the individual level. The identification of individual patterns of hypo- and hyper-perfusions in patients with ASL mainly relies on visual analysis (Zaharchuk et al., 2012; Chen et al., 2012). In (Petr et al., 2013), a template-based analysis was presented in order to detect individual activation patterns in functional ASL data. Interestingly,

they also applied their method to a pathological case in order to detect a hypo-perfusion co-located with a dysplasia in an epileptic patient. The most widespread approach to compute statistical analyses in neuroimaging is the massively univariate General Linear Model (GLM). This versatile model can be used in patient-specific analyses to compare a group of healthy controls to a patient. In (Maumet et al., 2013), we showed that using a heteroscedastic GLM should be favored (over the homoscedastic approach) in ASL studies studying single patients. Recently this method was also applied in epileptic patients (Boscolo Galazzo et al., 2015).

In neuroimaging, Gaussian smoothing is typically applied on pre-processed data before computing the GLM to compensate for small misregistrations, to reduce the effect of potential outliers and to insure that the residuals follows a Gaussian random field (Poldrack et al., 2011). While smoothing has demonstrated its usefulness, its use can be problematic in some brain pathologies (such as brain tumors) where the co-localisation of hypo and hyper-perfusions is common. As an example, ring-enhanced lesions present a pattern of abnormal perfusion in which is a central (hypo-perfused) necrosis is surrounded by an (hyper-perfused) enhanced ring. One can easily picture how Gaussian smoothing in this context can be detrimental as it mixes hypo- and hyper-perfusions, possibly cancelling out the observable effects.

As opposed to the massively univariate GLM, a major trend is the use of machine learning methods (cf. (Klöppel et al., 2012; Poldrack et al., 2011) for a review or (Paloyelis et al., 2014) for an example in ASL). These multivariate approaches are able to combine the voxel-wise information and can therefore achieve a greater sensitivity. While there have been recent effort to extend the use of machine learning approaches to identify departure from a group (Mourão Miranda et al., 2011), the large majority of machine learning approaches found in the literature have been developed to tackle a two- (or more) class separation problems. In our context, where each individual presents a different pattern of abnormalities those approaches are ill-suited. Machine learning approaches are also more computationally intensive. Another alternative, lying in between the massively univariate and the fully multivariate methods, is the searchlight approach (Kriegeskorte et al., 2006). This locally multivariate procedure focuses on multivariate patterns found in a spherical neighbourhood around each voxel. At its inception, two multivariate

metrics were considered: a Euclidean distance and a Mahalanobis distance (taking into account the spatial autocorrelation of the noise) (Kriegeskorte et al., 2006; Kriegeskorte and Bandettini, 2007b). These locally multivariate metrics achieved a greater sensitivity than the massively univariate GLM in functional MRI studies (Kriegeskorte et al., 2006). However, those approaches do not provide a direct interpretation of the detections as hypo- or hyper-perfusions. To determine whether a detection at a given voxel corresponds to a hypo- or hyper-perfusion, those approaches usually rely on the sign observed in the voxel-wise map (Kriegeskorte and Bandettini, 2007a). Recently, this approach was used in combination with machine learning classifiers (for example, see (Uddin et al., 2011)) but training a classifier at each location can be very computationally intensive (Poldrack et al., 2011).

The *a contrario* approach is a recent statistical framework (Desolneux et al., 2003) which comes from the computer vision community. This approach proceeds by counting the number of elements observed in a local excursion set defined around each voxel. We therefore propose to extend the searchlight procedure to the *a contrario* multivariate metric. In *a contrario*, the locally multivariate probability estimation is typically performed by assuming a white noise (Desolneux et al., 2003; Aguerrebere et al., 2009). However, spatial autocorrelation of the residuals is a well-known phenomenon in MRI, see for example (Wang et al., 2003b) for ASL, and cannot be disregarded. To ensure the validity of the *a contrario* approach for MRI, we therefore updated the probability estimation to model noise spatial dependency.

In this paper, we propose a new locally multivariate approach, based on an *a contrario* metric, for the detection of patient-specific brain perfusion abnormalities with ASL. For validation purposes, the proposed *a contrario* method is compared to the GLM by means of Receiver-Operating-Characteristic (ROC) analyses. Quantitative validation is performed on a dataset of 25 patients diagnosed with brain tumours and 61 healthy volunteers.

We previously investigated the ability to detect hypo- and hyper-perfusions using an *a contrario* approach in two conference papers (Maumet et al., 2012a,b). On the methodological side, we present here a new variant of our model, which takes into account the spatial autocorrelation of the noise. Furthermore, a larger database is studied, a novel

quantitative validation is provided, and a comprehensive quantitative comparison with the GLM is performed.

In section 2, we first introduce the theory and in particular the principles of *a contrario* analyses. Then, we present the experiments and the datasets under study. The results are presented in section 3. Section 4 gives a discussion and section 5 concludes.

2. Material and Methods

2.1. Theory

This section presents the methods developed in order to detect areas of abnormal perfusion in a patient map by comparison to a group of control subjects. In 2.1.1, we review the standard massively univariate General Linear Model (GLM). In 2.1.2, we describe the proposed locally multivariate *a contrario* approach.

2.1.1. The massively univariate General Linear Model

The General Linear Model used to detect patterns of abnormal perfusion in a patient by comparison to a group of subjects can be defined as a hierarchical model with two levels: subject and group. As suggested in (Maumet et al., 2013), we use a heteroscedastic model considering two distinct sources of variance: inter-subject and intra-subject. The level of perfusion in the patient map at voxel v is therefore compared to the level of perfusion in the control group with the following t -statistic:

$$t_{(v)} = \frac{\hat{\beta}_{(n,v)} - \frac{1}{\sum_{s=1}^{n-1} w_{(s,v)}} \sum_{s=1}^{n-1} w_{(s,v)} \hat{\beta}_{(s,v)}}{\sqrt{\hat{\sigma}_{G(v)}^2 + \hat{\sigma}_{(n,v)}^2 + \frac{1}{\sum_{s=1}^{n-1} w_{(s,v)}}}} \quad (1)$$

with $w_{(s,v)}^{-1} = \hat{\sigma}_{G(v)}^2 + \hat{\sigma}_{(s,v)}^2$, where $\hat{\beta}_{(s,v)}$ is the estimated perfusion for subject s at voxel v , $\hat{\sigma}_{G(v)}^2$ is the estimated inter-subject variance, $\hat{\sigma}_{(s,v)}^2$ the estimated intra-subject variance (over the ASL repetitions) for subject s and index n denotes the patient of interest.

Under the null hypothesis H_0 that the perfusion value in the patient map is equal to the control group mean, the t -statistic $t_{(v)}$ follows a Student distribution T_{n-1} with $n - 1$ degrees of freedoms. The probability under the null hypothesis, $\pi_{(v)}^{GLM}$, of being

in presence of such a high level of perfusion at voxel v of the patient map is therefore estimated by:

$$\pi_{(v)}^{GLM} = \Pr(T_{n-1} \geq t_{(v)}). \quad (2)$$

Hypo-perfusions can respectively be studied with $\Pr(T_{n-1} \leq t_{(v)})$. Valid detections are then obtained by thresholding the probability map using a correction for multiple testing.

2.1.2. The *a contrario* approach: a locally multivariate procedure

The *a contrario* approach (Desolneux et al., 2003) is a locally multivariate procedure which uses the size of a local excursion set as statistic. As illustrated in fig. 1, the *a contrario* probability estimation is usually based on a two-step procedure. First, massively univariate voxel-wise probabilities are computed under a so-called “background model”. Second, locally multivariate probabilities are estimated assuming a given type of local noise (usually white).

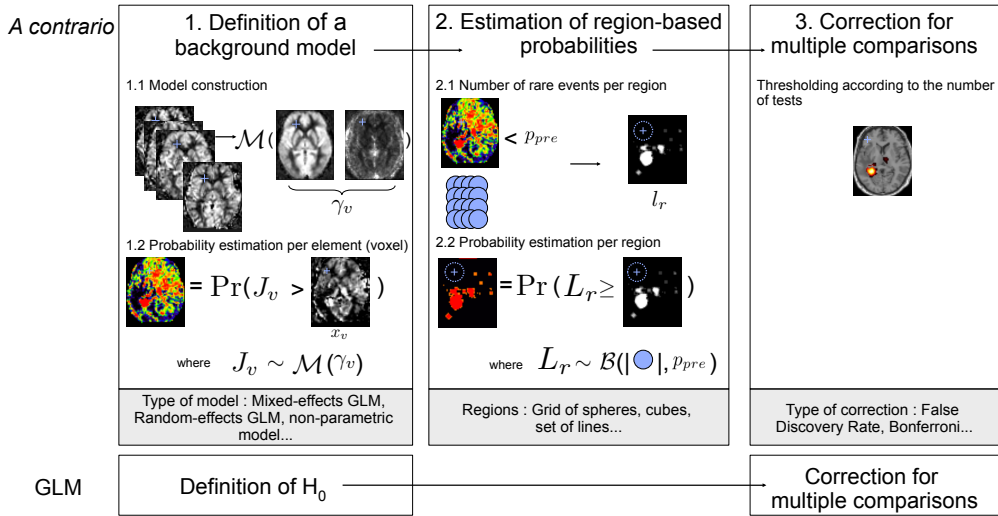


Figure 1: Overview of the *a contrario* approach: the *a contrario* statistic is computed in two steps: 1. Voxel-wise probabilities are estimated under a selected background model, 2. Region-based probabilities are calculated, those typically correspond to the voxel count of a local excursion set in pre-defined regions of interest. Finally, 3. Inference is performed using an appropriate correction for multiple comparisons. The GLM makes inference directly on the voxel-wise probabilities.

In the first step, a model of the background is defined at the voxel level. This model can represent the background noise or be learned from a control population. Let $J_{(v)}$

be the random variable representing the value observed at the voxel v of the image of interest. Under the background model \mathcal{M} , depending on the set of parameters γ , we have:

$$J_{(v)} \sim \mathcal{M}(\gamma_{(v)}). \quad (3)$$

Given $x_{(v)}$, the value observed at voxel v , each voxel of the image can be associated with a (univariate) probability under the background model:

$$\Pr(J_{(v)} \geq x_{(v)}), \text{ where } J_{(v)} \sim \mathcal{M}(\gamma_{(v)}) \quad (4)$$

Rousseau et al. first noticed that the term “background model” commonly employed in the *a contrario* literature could be well considered as a null hypothesis in standard hypothesis testing (Rousseau et al., 2007). So that the standard massively univariate GLM (based on unsmoothed data) can in fact be used to produce the input voxel-wise probability map of the *a contrario* analysis leading to:

$$\Pr(J_{(v)} \geq x_{(v)}) = \pi_{(v)}^{GLM}. \quad (5)$$

In the second step, in order to estimate locally multivariate probabilities, we then need:

1. A partition of the image into sub-regions;
2. A model to estimate the locally multivariate probability in a given region of interest from the univariate voxel-based probabilities.

Partition of the image into sub-regions. Given an *a priori* shape of the structures to be detected the regions can be selected accordingly. For instance, in the first *a contrario* application, which aimed at detecting alignments, segments of variable sizes were chosen as regions of analysis (Desolneux et al., 2003). In our case, the shape of the expected detections is unknown and, as it has been proposed in *a contrario* (Rousseau et al., 2007; Aguerrebere et al., 2009) and with the searchlight procedure (Kriegeskorte et al., 2006), the sub-regions are defined as spheres centred at each voxel.

From univariate voxel-based probabilities to a locally multivariate statistic. The computation of a locally multivariate statistic involves the notion of “rare events”. By definition, a rare event occurs at voxel v if the univariate probability (under the null hypothesis) to

observe such value, or a more extreme, is smaller than a pre-defined threshold p_{PRE} . The initial voxel-wise probability map is thus thresholded to produce a binary map :

$$K_{(v)} = \begin{cases} 1, & \text{if } \pi_{(v)}^{GLM} \leq p_{PRE}, \\ 0, & \text{otherwise.} \end{cases} \quad \text{with } K_{(v)} \sim \text{Bern}(p_{PRE}) \quad (6)$$

where $\text{Bern}(p_{PRE})$ is a Bernoulli distribution with probability p_{PRE} .

Then, the random variable describing the number of rare events in a region r , $L_{(r)}$, is determined by:

$$L_{(r)} = \sum_{v \in r} K_{(v)} \quad (7)$$

which is equal to the cardinal of the excursion set (parametrised by p_{PRE}) observed in each region.

The T-statistic used in the massively univariate GLM, is therefore replaced by a locally multivariate statistic in the *a contrario* approach: the number of rare events $L_{(r)}$. In order to be able to make inference using the *a contrario* approach, we must determine the distribution of the statistic $L_{(r)}$ under the null hypothesis.

The original *a contrario* approach assumes spatial independence of the residuals (Desolneux et al., 2003; Aguerrebere et al., 2009), i.e. the probability under the background model (or equivalently under the null hypothesis) to observe a rare event at a given voxel is independent of the probability to observe a rare event at any other voxel in the image, and in particular in its neighbourhood. However, the presence of spatial autocorrelation of the noise has been well described in the MRI literature, see for example (Chumbley and Friston, 2009) for fMRI or (Wang et al., 2003b) for ASL, and cannot be disregarded. We draw the attention of the reader that, under the *a contrario* approach, as for any hypothesis testing procedure, assumptions are made on the distribution of the error but not on the signal itself. Hence, the *a contrario* approaches makes no assumption on the spatial coherency of the (normal or abnormal) perfusion signal. In line with recent developments in the *a contrario* approach (Myaskovskey et al., 2012), we propose to rely on a multivariate Gaussian distribution in order to estimate the probability of observing a given number of rare events per sphere, as developed hereafter.

Let's first notice, that the probability of observing $l_{(r)}$ or more rare events can be

calculated as a sum of probabilities:

$$\Pr(L_{(r)} \geq l_{(r)}) = \sum_{i=l_{(r)}}^e \Pr(L_{(r)} = i) \quad (8)$$

Then, given $C_{(r)}^{(i)}$ the set of combinations of i elements among r , we have:

$$\Pr(L_{(r)} = i) = \sum_{S \in C_{(r)}^{(i)}} \Pr(K_{(v)} = 1, \forall v \in S \text{ and } K_{(v)} = 0, \forall v \notin S) \quad (9)$$

The problem now lies in the calculation of the joint probability to observe a given combination of “rare”/“not rare” voxels in the region of interest. Also, by definition of the term “rare event” we have:

$$\begin{aligned} & \Pr (\quad K_{(v)} = 1, \forall v \in S \quad \text{and} \quad K_{(v)} = 0, \forall v \notin S \quad) \\ &= \Pr (\quad \pi_{(v)} \leq p_{PRE}, \forall v \in S \quad \text{and} \quad \pi_{(v)} > p_{PRE}, \forall v \notin S \quad) \quad (10) \\ &= \Pr (\quad T_{n-1(v)} \geq \Phi_{T_{n-1}}^{-1}(p_{PRE}), \forall v \in S \quad \text{and} \quad T_{n-1(v)} < \Phi_{T_{n-1}}^{-1}(p_{PRE}), \forall v \notin S \quad), \end{aligned}$$

where $\Phi_{T_{\nu}}^{-1}$ is the inverse cumulative function of a Student distribution with ν degrees of freedom, outlining that the joint probability can be estimated from a Student’s t random field. As previously proposed in the neuroimaging literature regarding the use of random field theory for multiple testing correction (Kiebel et al., 1999; Zhang et al., 2009), we focus on the Gaussianised t random field:

$$z_{(v)} = \Phi^{-1}(\Phi_{T_{n-1}}(t_{(v)}))$$

where Φ^{-1} is the inverse cumulative function of a standard Normal distribution, and on the corresponding multivariate Gaussian random variable $Z_{(v)}$.

Then we approximate:

$$\begin{aligned} & \Pr (\quad T_{(v)} \geq \Phi_{T_{n-1}}^{-1}(p_{PRE}), \forall v \in S \quad \text{and} \quad T_{(v)} < \Phi_{T_{n-1}}^{-1}(p_{PRE}), \forall v \notin S \quad) \\ & \approx \Pr (\quad Z_{(v)} \geq \Phi^{-1}(p_{PRE}), \forall v \in S \quad \text{and} \quad Z_{(v)} < \Phi^{-1}(p_{PRE}), \forall v \notin S \quad). \end{aligned} \quad (11)$$

As a Gaussian random field, $Z_{(v)}$ is also representable by a multivariate Gaussian distribution, we have:

$$\begin{bmatrix} Z_{(1)} \\ \vdots \\ Z_{(e)} \end{bmatrix} \sim \mathcal{N}_{\epsilon} \left(\begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \sigma_{(1,2)}^2 & \cdots & \sigma_{(1,e)}^2 \\ \sigma_{(1,2)}^2 & & & \\ \vdots & \ddots & & \\ \sigma_{(1,e)}^2 & \cdots & & 1 \end{bmatrix} \right). \quad (12)$$

where \mathcal{N}_e is a multivariate Gaussian distribution of dimension e and $\sigma_{(i,j)}^2 = \text{cov}(Z_{(i)}, Z_{(j)})$. The problem now lies in the estimation of the covariance between each pair of voxel $\sigma_{(i,j)}^2, i \neq j$. To this aim, we use the great amount of work developed in the context of random field theory. In particular, assuming that $Z_{(v)}$ can be described as a Gaussian random field with a stationary homogeneous Gaussian correlation function, (Worsley et al., 1996) proposed an estimation of the spatial correlation. This estimate corresponds to the full-width-at-half-maximum of a Gaussian kernel that, applied to white noise, would lead to the same amount of autocorrelation. We computed this estimate using SPM8 (Statistical Parametric Mapping 8) on a subset of our subjects and found a full-width-at-half-maximum approximately equal to 1.5 voxels (in each direction). This value is in line with previous estimates obtained in ASL data (Wang et al., 2003a). Given this estimate of the autocorrelation, $\sigma_{(i,j)}^2$ is easily computed (cf. for example (Rasmussen and Williams, 2006) p.84).

The locally multivariate region-based probability of the *a contrario* approach, $\pi_{(r)}^*$, can therefore be calculated as the sum of joint probabilities in a multivariate Gaussian distribution:

$$\pi_{(r)}^* = \Pr(L_{(r)} \geq l_{(r)}) = \sum_{i=l_{(r)}}^e \Pr(L_{(r)} = i)$$

where $\Pr(L_{(r)} = i) = \sum_{S \in C_{(r)}^{(i)}} \Pr(Z_{(v)} \geq \Phi^{-1}(p_{PRE}), \forall v \in S \text{ and } Z_{(v)} < \Phi^{-1}(p_{PRE}), \forall v \notin S)$.

(13)

In practice, the probabilities $(\Pr(L_{(r)} = i))_{0 \leq i \leq e}$ are computed once at the beginning of the analysis. In the context of this paper, the probability associated to each region r is assigned to its centre voxel v to produce a voxel-wise probability map:

$$\pi_{(v)}^{*CONT} = \pi_{(r)}^*. \tag{14}$$

The standard *a contrario* approach assuming independence of the residuals can be seen as a special case where $\sigma_{(i,j)} = 0 \quad \forall (i,j) \in \{1 \dots e\}^2, i \neq j$.

2.2. Experiments

2.2.1. Simulated data analysis

Simulations were used to study the detections under different level of signal-to-noise

ratio (SNR) and for signal of different sizes. Two concentric spheres of opposite signal were generated to mimic ring-enhancing lesions (typically observed in brain lesions). The inner sphere represented a hypo-perfused necrosed area and the outer sphere the ring-enhancement. A Gaussian noise was added to the signal image so that the SNR was set to 0.5, 1 and 2 for both hypo- and hyper-perfusions. Each simulation was performed on a 3D image of size 30 voxels in each direction. The radius of the inner necrosis was set to 2, 4 or 6 voxels with the outer ring having a thickness of 1 voxels. For each setting, 100 repetitions were computed. Fig. 2 provides simulated signal for each necrosis size and examples of simulated images.

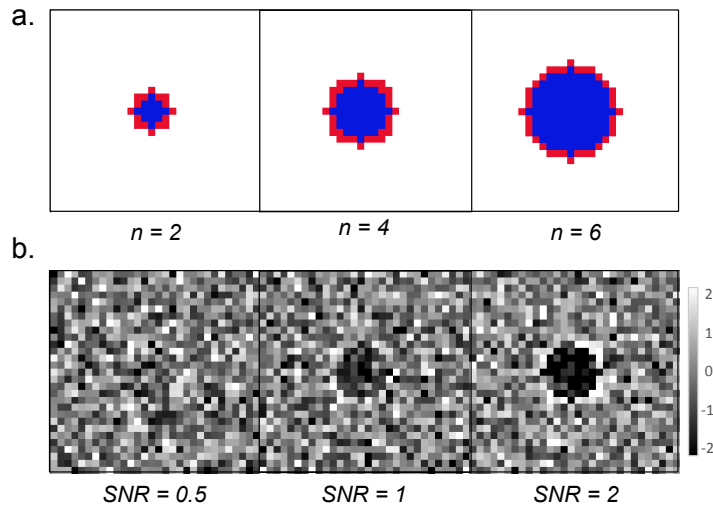


Figure 2: Simulated lesions. a) Ground truth hypo-perfusion (blue) and hyper-perfusion (red) for a necrosis radius of $n \in \{2, 4, 6\}$ voxels. b) Examples of simulated images for $SNR \in \{0.5, 1, 2\}$ and $n = 4$ voxels.

2.2.2. Real data analysis

25 patients diagnosed with brain tumours and 61 healthy volunteers were involved in this study. One control subject and four patients were excluded because of strong borderzone sign (Zaharchuk et al., 2009). The final dataset therefore included 21 patients (13 males, 8 females, age: 55.2 ± 14.1 years) and 60 healthy volunteers (28 males, 32 females, age: 29.4 ± 7.6 years).

We studied seven smoothing kernels for the GLM (with full-width-at-half-maximum $w \in \{0, 2, 4, 6, 8, 10, 12\}$ mm³) and nine parameter sets for the *a contrario* approach (the pre-defined p-value $p_{PRE} \in \{0.01, 0.005, 0.001\}$ and the sphere radius $rd \in \{1, 2, 3\}$ voxel(s)). Area under the curve were estimated for false positive rates ranging from 0 to 10%, as this is the area of interest in detection analysis (Skudlarski et al., 1999).

2.2.3. Experiment 1: Group ROC curves

In order to combine the data of all the patients included in this analysis, we computed an average ROC curve (across subjects). Given N_{hyper} , the number of subjects presenting at least one hyper-perfusion, the group true positive rate was then estimated by:

$$TPR_{G(p_{unc})} = \frac{1}{N_{hyper}} \sum_{s=1}^{N_{hyper}} TPR_{(s,p_{unc})} \quad (15)$$

On the other hand, the group false positive rate was estimated across all subjects by:

$$FPR_{G(p_{unc})} = \frac{1}{N} \sum_{s=1}^N FPR_{(s,p_{unc})} \quad (16)$$

2.2.4. Experiment 2: Individual ROC curves

Best parameters. For each subject, the parameter set (pair (p_{PRE}, rd) in *a contrario* or w in GLM) leading to the highest area under the ROC was identified and used to plot the corresponding “best” ROC curve.

Best parameter set across all subjects. As statistical analyses are usually performed using a single set of parameters across all subjects (and not adjusting for the best parameter set independently for each subject as performed in the previous paragraph), we also compared the *a contrario* and GLM methods across all subjects for each parameter set.

2.2.5. Ground truth

Based on clinical knowledge, as in (Maumet et al., 2013), we used a semi-automatic procedure that took advantage of the complementary anatomical (T1w-Gd, T2w FLAIR) and perfusion (DSC) information to get an estimation of the ground truth. The next subsections will briefly describe this procedure.

Due to its low SNR, ASL is not very well suited to measure low level of perfusion (Wintermark et al., 2005). That is why we focused on hyper-perfusions for sensitivity estimation. Hypo-perfusions were nevertheless retained for specificity calculation. Fig. 3 displays an example of ground truth for a representative subject.

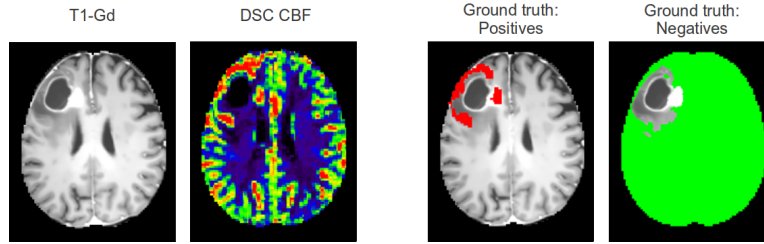


Figure 3: Ground truth for a representative subject. From left to right: T1-Gd, DSC CBF, ground-truth hyper-perfusions (i.e. ground truth positives) in red and ground-truth area of normal perfusion (i.e. ground truth negatives) in green overlaid on the T1-Gd.

Hyper-perfusions. In order to get an estimation of the positives (hyper-perfusions), we implemented a method inspired by the hotspot technique (Noguchi et al., 2008) commonly used in clinical practice. First, the affected tissue (tumour and oedema) was segmented using a semi-automated method based on the T2w and T1w-Gd images and visually inspected by an expert neuro-radiologist. Second, we compared the segmented regions to their contralateral counterparts in the DSC CBF map. Voxels overtaking the upper decile were identified as potential hyper-perfusions. Each potential perfusion abnormality was then visually inspected by an expert neuro-radiologist and manually corrected if needed. Special care was taken in order to avoid inclusion of hyper-perfusions related to the presence of arteries. Out of the 17 patients who underwent a DSC sequence, 9 presented hyper-perfusions.

Areas of normal perfusion. According to clinical knowledge, in the absence of metastasis, the perfusion abnormalities should be confined to the affected tissue (tumour and oedema) identifiable on T1w-Gd and T2w. Therefore, the non-affected tissue (brain voxels not segmented as part of the tumour or oedema) was considered as an area of normal perfusion (ground truth negatives). Here, all the 21 patients were included, including patient for who the DSC images were not available.

Computation of the ROC curves. For a given subject and a given method, a point of the ROC curve is obtained as described hereafter. First, the probability map is thresholded with an uncorrected p-value p_{unc} to outline the detections. Then, the true positive rate, $TPR_{(s,p_{unc})}$, and false positive rates, $FPR_{(s,p_{unc})}$, are computed by comparing the detections to the ground truth defined in section 2.2.5. This operation is iterated with uncorrected p-values ranging from 0 to 1 in order to draw the complete ROC curve. Here, it is worth noting that, since a cumulative probability distribution is, by definition, a monotonically increasing function, the true positive rates and false positive rates can equivalently be calculated by thresholding the statistic map instead of the probability map. The *a contrario* ROC curve can therefore be computed by thresholding the map of rare event counts and is thus independent of the method selected to compute the region-based probabilities. For ease of calculation, the ROC curves were estimated using the standard *a contrario* approach.

2.3. Data acquisition and pre-processing

2.3.1. Data acquisition

Data acquisition was performed on a 3T Siemens Verio MR scanner with a 32-channel head-coil. Patients were scanned in the context of clinical practice. The imaging protocol included a 3D T1-weighted anatomical sequence (T1w) (TR: 1900 ms, TE: 2.27 ms, FOV: 256 mm \times 256 mm \times 176 mm, flip angle: 9°, resolution: 1 mm \times 1 mm \times 1 mm), a PICORE Q2TIPS sequence (Wong et al., 1998) with crusher gradients (TR: 3000 ms, TE: 18 ms, FOV: 192 mm \times 192 mm, flip angle: 90°, resolution: 3 mm \times 3 mm, slice thickness: 7 mm, inter-slice gap: 0.7 mm, TI: 1700 ms, TI_{wd} : 700 ms, 60 repetitions (35 subjects) or 30 repetitions (25 subjects), mSENSE parallel imaging with accelerating factor of 2). In addition to these sequences, the patients also underwent a 3D T1w post gadolinium (T1w-Gd) sequence (TR: 1900 ms, TE: 2.27 ms, flip angle: 9°, FOV: 250 mm \times 250 mm \times 176 mm, resolution: 1 mm \times 1 mm \times 1 mm) and a 2D T2w FLAIR sequence (TR: 9000 ms, TE: 90 ms, FOV: 220 mm \times 199.4 mm, flip angle: 150°, resolution: 0.69 mm \times 0.69 mm, slice thickness: 4 mm). Out of the 21 patients, 17 subjects also underwent a DSC sequence (GRE EPI, TR: 1500 ms, TE: 30 ms, FOV: 230 mm \times 230 mm, flip angle: 90°, in plane resolution: 1.8 mm \times 1.8 mm, slice thickness: 4 mm, inter-slice gap: 1.2 mm).

2.3.2. Pre-processing

Arterial Spin Labelling. Image pre-processing was performed using SPM8 (Statistical Parametric Mapping 8, Wellcome Department of Imaging Neuroscience, University College, London) in Matlab R2012a (Mathworks, Natick, MA). The anatomical image of each subject was segmented using the unified segmentation (Ashburner and Friston, 2005). A subject-specific anatomical brain mask was created, excluding voxels with less than 50% of brain tissue in subsequent statistical analyses. A six-parameter rigid-body registration of the ASL volumes was carried out in order to reduce undesired effects due to subject motion. Pair-wise subtraction of the control and labelled scans was then computed. Rigid coregistration onto the whole brain anatomical map was then performed based on normalised mutual information. A standard kinetic model (Buxton et al., 1998) was then applied in order to get ASL CBF, according to the following:

$$f = 6000 \times \frac{\lambda \Delta M}{2 M_0 \alpha T I_{wd} \exp^{-(T I + i d x_{sl} * T I_{sl}) / T I_b}} \quad (17)$$

where f is the CBF map in $\text{mL} \cdot 100 \text{g}^{-1} \cdot \text{min}^{-1}$, M_0 the acquired M_0 map, λ the blood/tissue water partition coefficient, α measures the labelling efficiency, ΔM is the perfusion-weighted map, $T I = 1.7$ s the inversion time (Ferré et al., 2012), $i d x_{sl}$ the slice index (0 for the first slice), $T I_{sl} = 0.045$ s the readout time for one slice, $T I_{wd} = 0.7$ s the temporal width of the bolus, $T I_b$ the T1 of blood. According to (Wang et al., 2011), we assumed: $\lambda = 0.9 \text{ mL} \cdot \text{g}^{-1}$, $T I_b = 1.5$ s and $\alpha = 0.95$. The 6000 factor allows the conversion from $\text{mL} \cdot \text{g}^{-1} \cdot \text{s}^{-1}$ to $\text{mL} \cdot 100 \text{g}^{-1} \cdot \text{min}^{-1}$ which is the standard unit for CBF. We bring the attention of the reader to the fact that, contrary to what is usually done in ASL pre-processing, ΔM represents the set of perfusion-weighted maps (one volume per repetition) instead of a single perfusion-weighted map obtained by averaging across the repetitions. This is necessary in order to allow for the measurement of the within-subject variance.

Spatial normalization parameters estimated during the segmentation step were then applied to the T1 and ASL CBF map in order to normalize the subjects into the ICBM-452 T1 template space (Mazziotta et al., 2001). This registration algorithm was selected since it gives good results even in the presence of large anatomical lesions (Crinion et al., 2007). Normalisation in intensity was then applied to each ASL CBF map in order to

reduce the inter-subject variability (Aslan and Lu, 2010). The normalisation parameter was calculated as the average CBF in grey matter similarly to (Petr et al., 2013).

Dynamic Susceptibility weighted Contrast imaging. The DSC images were processed using MR manufacturer software by manually choosing an arterial input function to calculate: CBF, cerebral blood volume and mean transit time maps. The method is based on a deconvolution algorithm as described in (Ostergaard et al., 1996). Similarly to ASL, DSC CBF maps were co-registered on anatomical maps and spatially normalized.

3. Results

3.1. Simulations

Figure 4 plots the area under the ROC curve (AUC) obtained on the detections of hyper-perfusions and hyper-perfusions for the GLM and the *a contrario* approach on simulated lesions.

With the GLM, we observe an opposite effects of smoothing on the detection of hypo-perfusions versus hyper-perfusions. While smoothing is beneficial to detect hyper-perfusions when the necrosis radius is greater than 2 voxels, for hyper-perfusions, on the other hand, smoothing only helps with a radius of 2. This illustrates the fact that with a Gaussian smoothing the signal can be cancelled by neighboring voxels if they hold signal of opposite sign.

With the *a contrario* approach, a good level of detection can be reached at the same time for hypo- and hyper-perfusions. As the SNR increase, the *a contrario* detections get closer to the ground truth with any set of parameters studied.

3.2. Real data

In 3.2.1, the ROC curves corresponding to each individual patient are computed and compared. Then, in 3.2.2, the average ROC curve over the group is computed for each method. Finally, a qualitative comparison is presented in 3.2.3.

3.2.1. Individual ROC curves

Best parameters. In fig. 5, it is clearly visible that the proposed *a contrario* approach outperforms the GLM for all subjects. A paired two-sample *t*-test comparing the *a contrario*

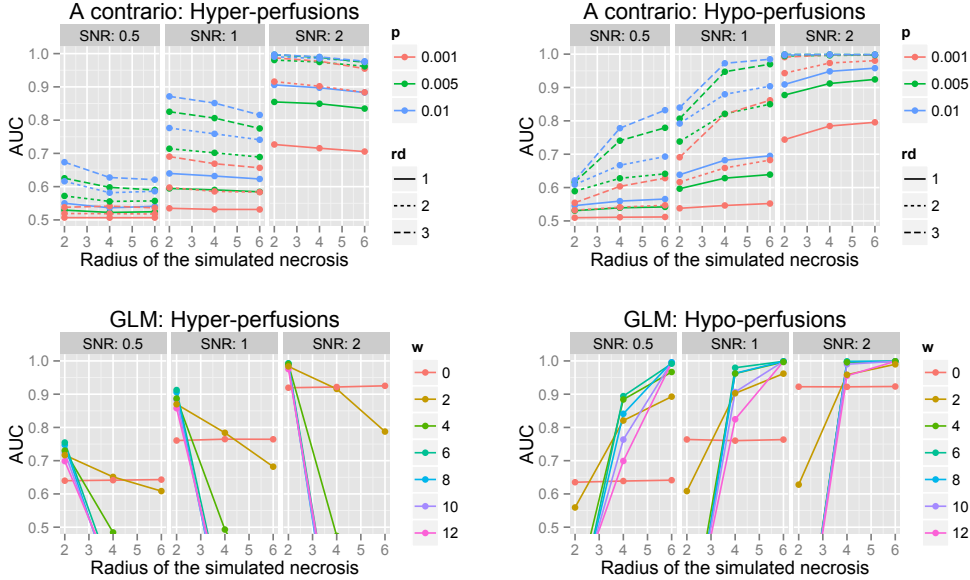


Figure 4: Area under the ROC curves (AUC) for the proposed *a contrario* approach (first row) and the massively univariate GLM (second row) in the detection of hyper-perfusions (first column) and hypo-perfusions (second column) on simulated data (necrosis radius $n \in \{2, 4, 6\}$ voxels and $SNR \in \{0.5, 1, 2\}$).

and the GLM with the best parameter set for each subject outlines a significant difference ($p = 0.008$). Furthermore the best parameter set found for each subject was less variable for the *a contrario* approach (the best set was ($rd = 3; p_{PRE} = 0.001$) for 6 subjects, ($rd = 2; p_{PRE} = 0.001$) for 2 subjects and ($rd = 2; p_{PRE} = 0.005$) for the remaining subject) than for the GLM (the best set was $w = 4 \text{ mm}^3$ for 2 subjects, $w = 6 \text{ mm}^3$ for 1 subject, $w = 8 \text{ mm}^3$ for 2 subjects and $w = 12 \text{ mm}^3$ for 4 subjects).

Best parameter set across all subjects. Table 1 presents the area under the ROC curves (average SNR over the 9 patients) for the *a contrario* approach and the GLM. The best area under the curve is obtained for the *a contrario* approach with $p_{PRE} = 0.001$ and $rd = 3$ and for the GLM with $w = 6 \text{ mm}^3$. A paired two-sample *t*-test between the best configurations outlines the superiority of the *a contrario* approach ($p = 0.004$). This is in line with the higher variability of the best parameter set for the GLM outlined in the previous section.

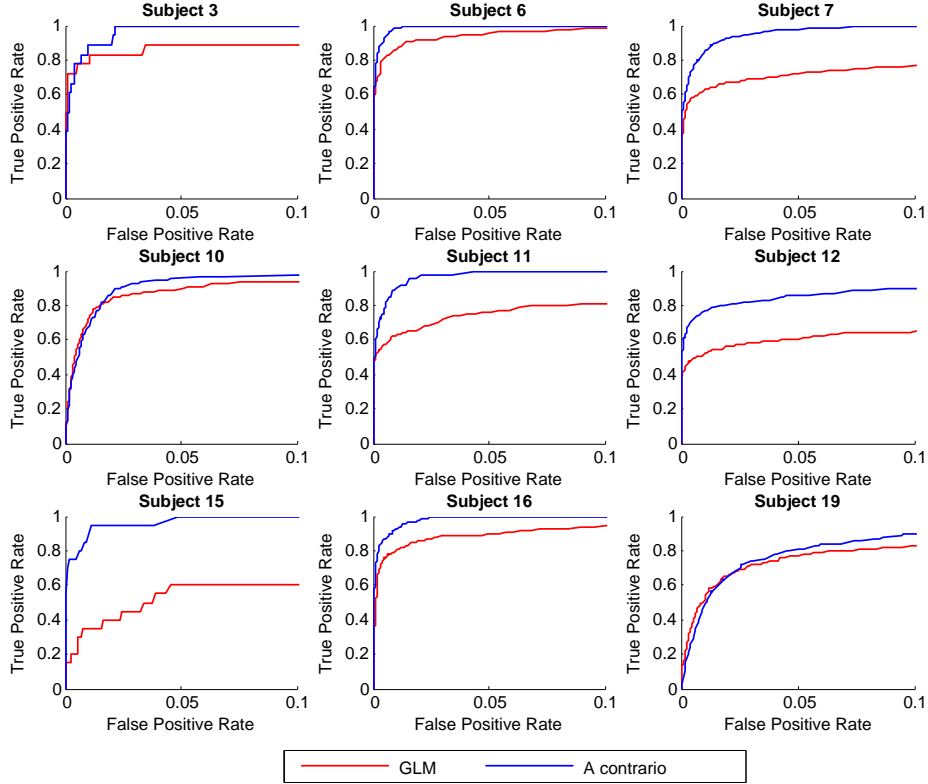


Figure 5: Best ROC curves for false positive rates ranging between 0% and 10%, for the massively univariate GLM (red) and the proposed *a contrario* approach (blue) in nine patients. Overall, the *a contrario* approach outperforms the GLM.

3.2.2. Group ROC curves

Fig. 6 plots the resulting average ROC curve for each parameter set for the *a contrario* approach and the GLM. The corresponding area under the curves are summarized in table 2. The results are very close to the one obtained on the set of 9 subjects in the previous section, outlining the fact the false positive rate at a given uncorrected p-value is not varying much across subjects. As before, the *a contrario* approach outperforms the GLM.

Table 1: Area under the ROC curve, averaged over the 9 subjects presenting hyper-perfusions, for false positive rates ranging between 0% and 10%, with the a contrario approach (with a sphere radius $rd \in \{1, 2, 3\}$ voxel(s) and a pre-defined p-value $p \in \{0.01, 0.005, 0.001\}$) and the GLM (smoothed with a Gaussian kernel of full-width-at-half-maximum $w \in \{0, 2, 4, 6, 8, 10, 12\}$ mm³). The proposed a contrario approach outperforms all tested configurations of the GLM.

<i>a contrario</i>	$rd = 1$			$rd = 2$			$rd = 3$			
	$p = 0.01$	$p = 0.005$	$p = 0.001$	$p = 0.01$	$p = 0.005$	$p = 0.001$	$p = 0.01$	$p = 0.005$	$p = 0.001$	
	0.68	0.73	0.78	0.76	0.81	0.87	0.74	0.83	0.91	
GLM	$w = 0$	$w = 2$	$w = 4$	$w = 6$	$w = 8$	$w = 10$	$w = 12$			
	0.63	0.68	0.71	0.72	0.72	0.69	0.64			

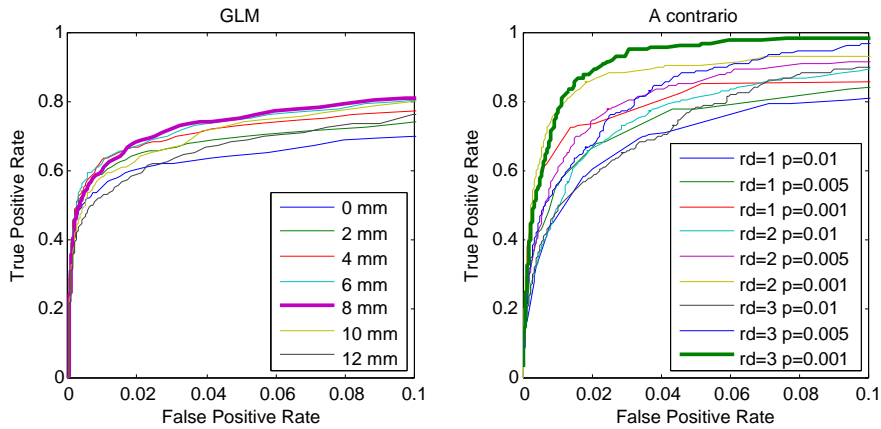


Figure 6: Average ROC curve (based on 21 subjects), for false positive rates ranging between 0% and 10%, with the GLM (smoothed with a Gaussian kernel of full-width-at-half-maximum $w \in \{0, 2, 4, 6, 8, 10, 12\}$ mm³) and the a contrario approach (with a sphere radius $rd \in \{1, 2, 3\}$ voxel(s) and a pre-defined p-value $p_{PRE} \in \{0.01, 0.005, 0.001\}$)

3.2.3. Qualitative comparison

In order to illustrate the advantage of the *a contrario* approach over the GLM, we chose 2 representative subjects and compared the methods at fixed false positive rate and true positive rate. To this aim, we selected the uncorrected p-values that would lead to a pre-specified false positive rate (or true positive rate) from the ROC analysis. We

Table 2: Area under the average ROC curve (based on 21 subjects), for false positive rates ranging between 0% and 10%, with the *a contrario* approach (with a sphere radius $rd \in \{1, 2, 3\}$ voxel(s) and a pre-defined p-value $p \in \{0.01, 0.005, 0.001\}$) and the GLM (smoothed with a Gaussian kernel of full-width-at-half-maximum $w \in \{0, 2, 4, 6, 8, 10, 12\}$ mm³). The proposed *a contrario* approach outperforms the GLM.

<i>a contrario</i>	$rd = 1$			$rd = 2$			$rd = 3$		
	$p = 0.01$	$p = 0.005$	$p = 0.001$	$p = 0.01$	$p = 0.005$	$p = 0.001$	$p = 0.01$	$p = 0.005$	$p = 0.001$
	0.68	0.73	0.79	0.75	0.80	0.87	0.72	0.81	0.91
GLM	$w = 0$	$w = 2$	$w = 4$	$w = 6$	$w = 8$	$w = 10$	$w = 12$		
	0.63	0.67	0.70	0.72	0.73	0.70	0.65		

worked with the best set of parameters identified for each subject (from section 3.2.1). Fig. 7 presents the detections obtained with both methods on the 2 selected patients. The first subject, which presents a small hyper-perfusion, is studied at a true positive rate of 50% and the second at a false positive rate of 5%.

Patient 15 suffers from a gliosarcoma in the left hemisphere near the parahippocampal region. The lesion displays a small hyper-perfusion in its dorsal part, small hypoperfusions are seen in the surrounding oedema. At a true positive rate of 50%, the GLM displays a larger number of false positives than the *a contrario* approach.

Patient 7 presents a meningioma of both occipital lobes characterized by a large hyper-perfusion. At a false positive rate of 5%, most of the hyper-perfusion was properly detected by both methods. The *a contrario* approach is however more sensitive as it detects a larger proportion of the hyper-perfusion.

These 2 cases illustrate the reduced sensitivity of the GLM by comparison to the *a contrario* approach at the same specificity level.

4. Discussion

Quantitative assessment of the detections and comparison between detection methods are challenging tasks. This is mainly because, like in many other medical imaging problems, the ground truth is not clearly stated. We chose to evaluate this framework

on patients diagnosed with tumour pathology because perfusion abnormalities have been widely studied in this context (Chalela et al., 2000; Sugahara et al., 2000; Warmuth et al., 2003; Weber et al., 2006; Chawla et al., 2007). Also, as perfusion is useful clinical information, DSC is part of the clinical routine for these patients.

-The normalisation in intensity used in this paper is not optimal and can lead to the detection of pseudo-decreases or pseudo-increases (for a related discussion cf. (Buchert et al., 2005)). In theory, based on a Q2TIPS ASL sequence (Wong et al., 1998) one should be able to derive quantitative CBF values and avoid intensity normalisation. However, in the context of this analysis, we observed a large variability in global CBF signal both in the control and patient groups that could not be explained by age or gender. To compensate for those large variations across subjects, using an intensity normalisation scheme appeared to be a requirement. We hypothesize that the large variations we observed could be explained by an imperfect timing of our sequence. Recent literature (e.g. (Villien et al., 2012), (Alsop et al., 2014)) recommends using a shorter bolus width than the one we typically used in this experiment. If our saturation pulses were indeed too late, it would mean that only pseudo-quantitative values could be derived from the acquisition hence calling for a normalisation in intensity. The a contrario approach can be applied in presence or absence of a normalisation procedure. When working on truly quantitative metrics, intensity normalisation can be avoided.

Subjects in our patient group were significantly older than the control subjects which is clearly not ideal as age is known to impact both CBF (Biagi et al., 2007) and grey matter/white matter structure leading to different observed level of perfusion. While the use of a normalisation scheme lessened the effect of change in global of perfusion, this is not enough to compensate for age-related variations. For more precise detections an age-matched template would be needed. Nevertheless the results observed here are valid and we expect that using an age-matched template would improve the detections.

Calculation of CBF using singular value decomposition on DSC can lead to under or over-estimation of the CBF if the arrival times are not the same in the arterial input function and in the tissue. This is a known drawback of SVD methods and might have biased our estimation of the CBF based on DSC. But, as the ground truth detections were not determined based on a thresholding on absolute values but rather using a rel-

ative comparison of left and right hemisphere and then manually updated by a clinician if needed, we do not expect this to be too much of a concern.

-The *a contrario* approach demonstrated a clear improvement over the GLM which is the most widely adopted method to analyse functional MRI data. Compared to other multivariate approaches, such as machine learning techniques, the *a contrario* approach has the advantage of remaining very simple and providing results in the same manner as a standard GLM analysis. Compared to a classic searchlight approach using Mahalanobis distance, the *a contrario* approaches has the possibility to intrinsically differentiate hypo- from hyper-perfusions which is critical in the context of brain tumours. While a comprehensive comparison with those techniques was beyond the scope of this paper, an interesting extension of this experiment could perform a direct comparison of the *a contrario* approach to machine learning (in particular one-class implementation as in (Mourão Miranda et al., 2011)) and searchlight procedures.

In our experiments on real data, the best performance for the *a contrario* approach were reached for a sphere radius of $rd = 3$ voxels and a threshold of $p_{PRE} = 0.001$. This parameter set was found to be best for 6 subjects out of 9 among all tested configurations. It would be interesting to further investigate how this extends to other analysis and in particular if even smaller p_{PRE} leads to more accurate results.

In this paper, region-based probabilities were computed by computing the joint probability of a multivariate normal distribution. While these can be easily pre-computed to reduce the computational time at execution, they can nevertheless be very computationally intensive to compute for a sphere radius greater than 2 or 3 voxels. Using a non-parametric approach, such as permutation-based or sign-flipping technique (Winkler et al., 2014) might be an interesting alternative.

5. Conclusion

We have presented a locally multivariate procedure using a *a contrario* metric for the detection of patient-specific perfusion abnormalities in ASL. Using ROC curves, we outlined that this locally multivariate procedure outperforms the massively univariate GLM, commonly employed in neuroimaging.

Furthermore, we introduced a non-independent version of this *a contrario* procedure in order to deal with the spatial autocorrelation of the noise observed in MRI.

We assessed our approach in the context of brain tumours since perfusion patterns have been widely studied in this clinical context. This allowed us to get an estimate of the ground truth and perform a quantitative validation. Our approach is however suitable to a broader range of applications and we plan in the future to study pathologies presenting more subtle patterns of abnormal perfusion.

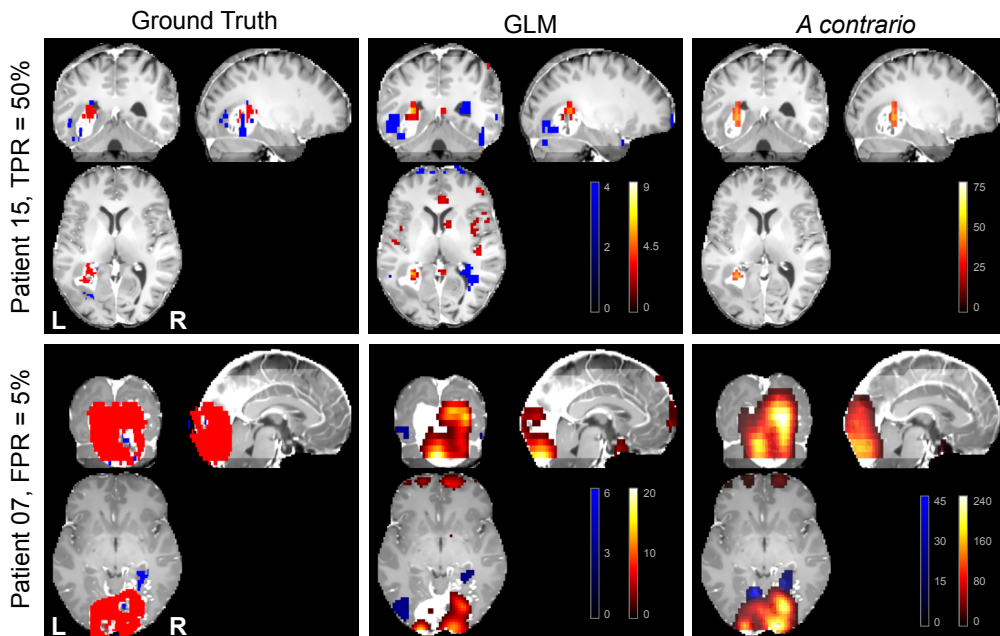


Figure 7: Perfusion abnormalities detections in 2 patients with the GLM and the *a contrario* approach using the best parameter set for each method. From left to right: ground truth hyper-perfusions (red) and hypo-perfusions (blue) overlaid on the T1w-Gd map; GLM, hyper-perfusions (red to yellow) and hypo-perfusions (dark to light blue); *a contrario* detections hyper-perfusions (red to yellow) and hypo-perfusions (dark to light blue). For the *a contrario* and GLM detections, minus logarithm of the p-values is displayed.

6. Acknowledgements

This work has been supported by Région Bretagne, Neurinfo Core Facilities and by the Wellcome Trust.

References

- Aguerrebere, C., Sprechmann, P., Musé, P., Ferrando, R., 2009. A-contrario localization of epileptogenic zones in SPECT images, in: Proceedings of the IEEE 6th International Symposium on Biomedical Imaging (ISBI 2009), pp. 570–573.
- Alsop, D.C., Detre, J.a., Golay, X., Günther, M., Hendrikse, J., Hernandez-Garcia, L., Lu, H., Macintosh, B.J., Parkes, L.M., Smits, M., van Osch, M.J.P., Wang, D.J.J., Wong, E.C., Zaharchuk, G., 2014. Recommended implementation of arterial spin-labeled perfusion MRI for clinical applications: A consensus of the ISMRM perfusion study group and the European consortium for ASL in dementia. *Magnetic resonance in medicine : official journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine* 00.
- Ashburner, J., Friston, K.J., 2005. Unified segmentation. *NeuroImage* 26, 839–851.
- Aslan, S., Lu, H., 2010. On the sensitivity of ASL MRI in detecting regional differences in cerebral blood flow. *Magnetic Resonance Imaging* 28, 928–935.
- Biagi, L., Abbruzzese, A., Bianchi, M.C., Alsop, D.C., Del Guerra, A., Tosetti, M., 2007. Age dependence of cerebral perfusion assessed by magnetic resonance continuous arterial spin labeling. *Journal of magnetic resonance imaging : JMRI* 25, 696–702.
- Boscolo Galazzo, I., Storti, S.F., Del Felice, A., Pizzini, F.B., Arcaro, C., Formaggio, E., Mai, R., Chappell, M., Beltramello, A., Manganotti, P., 2015. Patient-specific detection of cerebral blood flow alterations as assessed by arterial spin labeling in drug-resistant epileptic patients. *PLoS ONE* 10, e0123975.
- Buchert, R., Wilke, F., Chakrabarti, B., Martin, B., Brenner, W., Mester, J., Clausen, M., 2005. Adjusted scaling of FDG positron emission tomography images for statistical evaluation in patients with suspected Alzheimer’s disease. *Journal of neuroimaging : official journal of the American Society of Neuroimaging* 15, 348–55.
- Buxton, R.B., Frank, L.R., Wong, E.C., Siewert, B., Warach, S., Edelman, R.R., 1998. A general kinetic model for quantitative perfusion imaging with arterial spin labeling. *Magnetic Resonance in Medicine* 40, 383–396.
- Chalela, J., Alsop, D.C., Gonzalez-Atavales, J., Maldjian, J., Kasner, S., Detre, J.A., 2000. Magnetic resonance perfusion imaging in acute ischemic stroke using continuous arterial spin labeling. *Stroke* 31, 680–687.
- Chawla, S., Wang, S., Wolf, R.C., Woo, J.H., Wang, J., O’Rourke, D.M., Judy, K.D., Grady, M.S., Melhem, E.R., Poptani, H., 2007. Arterial spin-labeling and MR spectroscopy in the differentiation of gliomas. *American Journal of Neuroradiology* 28, 1683–1689.
- Chen, T.Y., Chiu, L., Wu, T.C., Wu, T.C., Lin, C.J., Wu, S.C., Tsui, Y.K., 2012. Arterial spin-labeling in routine clinical practice: a preliminary experience of 200 cases and correlation with MRI and clinical findings. *Clinical imaging* 36, 345–352.
- Chumbley, J.R., Friston, K.J., 2009. False discovery rate revisited: FDR and topological inference using Gaussian random fields. *NeuroImage* 44, 62–70.

- Colliot, O., Bernasconi, N., Mansi, T., Naessens, V., Bernasconi, A., Klironomos, D., 2006. Segmentation of focal cortical dysplasia lesions on MRI using level set evolution. *NeuroImage* 32, 1621–30.
- Crinion, J., Ashburner, J., Leff, A., Brett, M., Price, C., Friston, K., 2007. Spatial normalization of lesioned brains: performance evaluation and impact on fMRI analyses. *NeuroImage* 37, 866–875.
- Desolneux, A., Moisan, L., Morel, J., 2003. A grouping principle and four applications. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 25, 508–513.
- Detre, J., Leigh, J., et al., 1992. Perfusion imaging. *Magnetic Resonance in Medicine* 23, 37–45.
- Ferré, J.C., Petr, J., Bannier, E., Barillot, C., Gauvrit, J.Y., 2012. Improving quality of arterial spin labeling MR imaging at 3 tesla with a 32-channel coil and parallel imaging. *Journal of Magnetic Resonance Imaging* 1239, 1233–1239.
- Huck, S., Kerl, H.U., Al-Zghloul, M., Groden, C., Nölte, I., 2012. Arterial spin labeling at 3.0 Tesla in subacute ischemia: comparison to dynamic susceptibility perfusion. *Clinical neuroradiology* 22, 29–37.
- Kiebel, S.J., Poline, J.B., Friston, K.J., Holmes, a.P., Worsley, K.J., 1999. Robust smoothness estimation in statistical parametric maps using standardized residuals from the general linear model. *NeuroImage* 10, 756–66.
- Klöppel, S., Abdulkadir, A., Jack, C.R., Koutsouleris, N., Mourão Miranda, J., Vemuri, P., 2012. Diagnostic neuroimaging across diseases. *NeuroImage* 61, 457–63.
- Kriegeskorte, N., Bandettini, P., 2007a. Analyzing for information, not activation, to exploit high-resolution fMRI. *NeuroImage* 38, 649–62.
- Kriegeskorte, N., Bandettini, P., 2007b. Combining the tools: activation- and information-based fMRI analysis. *NeuroImage* 38, 666–8.
- Kriegeskorte, N., Goebel, R., Bandettini, P., 2006. Information-based functional brain mapping. *Proceedings of the National Academy of Sciences of the United States of America* 103, 3863–8.
- Maumet, C., Maurel, P., Ferré, J.C., Barillot, C., 2012a. A contrario detection of focal brain perfusion abnormalities based on an ASL template, in: *Proceedings of the IEEE 9th International Symposium on Biomedical Imaging (ISBI 2012)*, pp. 1176–1179.
- Maumet, C., Maurel, P., Ferré, J.C., Barillot, C., 2012b. A comprehensive framework for the detection of individual brain perfusion abnormalities using arterial spin labeling, in: *Proceedings of the 15th International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI 2012)*, pp. 542–549.
- Maumet, C., Maurel, P., Ferré, J.C., Carsin, B., Barillot, C., 2013. Patient-specific detection of perfusion abnormalities combining within-subject and between-subject variances in Arterial Spin Labeling. *NeuroImage* 81C, 121–130.
- Mazziotta, J., Toga, A., Evans, A., 2001. A four-dimensional probabilistic atlas of the human brain. *Journal of the American Medical Informatics Association* 8, 401–430.
- Mourão Miranda, J., Hardoon, D.R., Hahn, T., Marquand, A.F., Williams, S.C.R., Shawe-Taylor, J., Brammer, M., 2011. Patient classification as an outlier detection problem: an application of the One-Class Support Vector Machine. *NeuroImage* 58, 793–804.

- Myaskovskiy, A., Gousseau, Y., Lindenbaum, M., 2012. Beyond Independence: An Extension of the A Contrario Decision Procedure. *International Journal of Computer Vision* .
- Noguchi, T., Yoshiura, T., Hiwatashi, A., Togao, O., Yamashita, K., Nagao, E., Shono, T., Mizoguchi, M., Nagata, S., Sasaki, T., Suzuki, S.O., Iwaki, T., Kobayashi, K., Mihara, F., Honda, H., 2008. Perfusion imaging of brain tumors using arterial spin-labeling: correlation with histopathologic vascular density. *American Journal of Neuroradiology* 29, 688–93.
- Ostergaard, L., Weisskoff, R.M., Chesler, D.A., Gyldensted, C., Rosen, B.R., 1996. High resolution measurement of cerebral blood flow using intravascular tracer bolus passages. Part I: Mathematical approach and statistical analysis. *Magnetic Resonance in Medicine* 36, 715–725.
- Paloyelis, Y., Doyle, O.M., Zelaya, F.O., Maltezos, S., Williams, S.C., Fotopoulou, A., Howard, M.A., 2014. A Spatiotemporal Profile of In Vivo Cerebral Blood Flow Changes Following Intranasal Oxytocin in Humans. *Biological Psychiatry* , 1–13.
- Pendse, N., Wissmeyer, M., Altrichter, S., Vargas, M., Delavelle, J., Viallon, M., Federspiel, A., Seeck, M., Schaller, K., Lövlblad, K.O., 2010. Interictal arterial spin-labeling MRI perfusion in intractable epilepsy. *Journal of neuroradiology* 37, 60–3.
- Petr, J., Ferré, J.C., Raoult, H., Bannier, E., Gauvrit, J.Y., Barillot, C., 2013. Template-based approach for detecting motor task activation-related hyperperfusion in pulsed asl data. *Human Brain Mapping* .
- Pinkham, A., Loughead, J., Ruparel, K., Wu, W.C., Overton, E., Gur, R., Gur, R., 2011. Resting quantitative cerebral blood flow in schizophrenia measured by pulsed arterial spin labeling perfusion MRI. *Psychiatry research* 194, 64–72.
- Poldrack, R.A., Mumford, J., Nichols, T., 2011. *Handbook of functional MRI data analysis*. Cambridge University Press, Cambridge.
- Rasmussen, C.E., Williams, C.K.I., 2006. Covariance Functions, in: *Gaussian Processes for Machine Learning*. MIT Press. chapter 4.
- Rousseau, F., Faisan, S., Heitz, F., Armspach, J.P., Chevalier, Y., Blanc, F., de Seze, J., Rumbach, L., 2007. An a contrario approach for change detection in 3D multimodal images: application to multiple sclerosis in MRI. *Proceedings of the 29th International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS 2007)* , 2069–2072.
- Scarpazza, C., Sartori, G., De Simone, M.S., Mechelli, a., 2013. When the single matters more than the group: Very high false positive rates in single case Voxel Based Morphometry. *NeuroImage* 70, 175–88.
- Skudlarski, P., Constable, R.T., Gore, J.C., 1999. ROC analysis of statistical methods used in functional MRI: individual subjects. *NeuroImage* 9, 311–329.
- Sugahara, T., Korogi, Y., Tomiguchi, S., Shigematsu, Y., Ikushima, I., Kira, T., Liang, L., Ushio, Y., Takahashi, M., 2000. Posttherapeutic intraaxial brain tumor: the value of perfusion-sensitive contrast-enhanced MR imaging for differentiating tumor recurrence from nonneoplastic contrast-enhancing tissue. *American Journal of Neuroradiology* 21, 901–909.
- Uddin, L.Q., Menon, V., Young, C.B., Ryali, S., Chen, T., Khouzam, A., Minshew, N.J., Hardan, A.Y.,

2011. Multivariate searchlight classification of structural magnetic resonance imaging in children and adolescents with autism. *Biological psychiatry* 70, 833–41.
- Villien, M., Chipon, E., Troprès, I., Bouvier, J., Cantin, S., Chechin, D., Le Bas, J.F., Krainik, a., Warnking, J.M., 2012. Per-subject characterization of bolus width in pulsed arterial spin labeling using bolus turbo sampling. *Magnetic resonance in medicine : official journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine* 000, 1–6.
- Viviani, R., Beschoner, P., Ehrhard, K., Schmitz, B., Thöne, J., 2007. Non-normality and transformations of random fields, with an application to voxel-based morphometry. *NeuroImage* 35, 121–30.
- Wang, J., Aguirre, G.K., Kimberg, D.Y., Detre, J.a., 2003a. Empirical analyses of null-hypothesis perfusion fMRI data at 1.5 and 4 T. *NeuroImage* 19, 1449–1462.
- Wang, J., Licht, D.J., Jahng, G.H., Liu, C.S., Rubin, J.T., Haselgrove, J., Zimmerman, R.a., Detre, J.a., 2003b. Pediatric perfusion imaging using pulsed arterial spin labeling. *Journal of Magnetic Resonance Imaging* 18, 404–13.
- Wang, Y., Saykin, A.J., Pfeuffer, J., Lin, C., Mosier, K.M., Shen, L., Kim, S., Hutchins, G.D., 2011. Regional reproducibility of pulsed arterial spin labeling perfusion imaging at 3T. *NeuroImage* 54, 1188–1195.
- Warmuth, C., Günther, M., Zimmer, C., 2003. Quantification of Blood Flow in Brain Tumors: Comparison of Arterial Spin Labeling and Dynamic Susceptibility-weighted Contrast-enhanced MR Imaging. *Radiology* 228, 523–532.
- Weber, M.A., Zoubaa, S., Schlieter, M., Jüttler, E., Huttner, H.B., Geletneky, K., Ittrich, C., Lichy, M.P., Kroll, A., Debus, J., Giesel, F.L., Hartmann, M., Essig, M., 2006. Diagnostic performance of spectroscopic and perfusion MRI for distinction of brain tumors. *Neurology* 66, 1899–1906.
- Wheeler, H.M., Mlynash, M., Inoue, M., Tipirneni, A., Liggins, J., Zaharchuk, G., Straka, M., Kemp, S., Bammer, R., Lansberg, M.G., Albers, G.W., 2013. Early Diffusion-Weighted Imaging and Perfusion-Weighted Imaging Lesion Volumes Forecast Final Infarct Size in DEFUSE 2. *Stroke; a journal of cerebral circulation* 44, 681–5.
- Winkler, A.M., Ridgway, G.R., Webster, M.a., Smith, S.M., Nichols, T.E., 2014. Permutation inference for the general linear model. *NeuroImage* 92, 381–397.
- Wintermark, M., Sesay, M., Barbier, E., Borbély, K., Dillon, W.P., Eastwood, J.D., Glenn, T.C., Grandin, C.B., Pedraza, S., Soustiel, J.F., Nariai, T., Zaharchuk, G., Caillé, J.M., Dousset, V., Yonas, H., 2005. Comparative overview of brain perfusion imaging techniques. *Stroke* 36, e83–e99.
- Wong, E.C., Buxton, R.B., Frank, L.R., 1998. Quantitative imaging of perfusion using a single subtraction (QUIPSS and QUIPSS II). *Magnetic Resonance in Medicine* 39, 702–708.
- Worsley, K.J., Marrett, S., Neelin, P., Vandal, a.C., Friston, K.J., Evans, a.C., 1996. A unified statistical approach for determining significant signals in images of cerebral activation. *Human brain mapping* 4, 58–73.
- Zaharchuk, G., Bammer, R., Straka, M., Shankaranarayan, A., Alsop, D.C., Fischbein, N.J., Atlas, S.W., Moseley, M.E., 2009. Arterial spin-label imaging in patients with normal bolus perfusion-weighted MR imaging findings: pilot identification of the borderzone sign. *Radiology* 252, 797–807.

- Zaharchuk, G., El Mogy, I.S., Fischbein, N.J., Albers, G.W., 2012. Comparison of arterial spin labeling and bolus perfusion-weighted imaging for detecting mismatch in acute stroke. *Stroke* 43, 1843–1848.
- Zhang, H., Nichols, T.E., Johnson, T.D., 2009. Cluster mass inference via random field theory. *NeuroImage* 44, 51–61.