Estimating Model Uncertainty when fitting Multiple b-value Diffusion Weighted Imaging

M. R. Orton[1], D. J. Collins[1], D-M. Koh[2], M. Germuska[1], and M. O. Leach[1]
[1]CR-UK and EPSRC Cancer Imaging Centre, Institute of Cancer Research, Sutton, Surrey, United Kingdom, [2]Department of Radiology, Royal Marsden NHS Foundation Trust, Sutton, Surrey, United Kingdom

Introduction A number of different models have been proposed in the literature for describing diffusion-weighted data acquired with a range of b-values. Due to data acquisition limitations these models have typically been applied in the brain, but more recent developments have enabled diffusion weighted imaging (DWI) to be obtained from other regions of the body, including the abdominal and pelvic organs. Based on the physiology of these organs a number of models have been suggested that link in with known organ physiology, especially in relation to anomalous diffusion, distributions of diffusion rates and perfusion effects. However, due to the complexity of the micro-environment being probed in a diffusion experiment, these links remain uncertain in human systems, and there is even greater uncertainty when considering tumours. There is little work that looks at determining the relative statistical appropriateness of these or other models for application to a given data set. In this abstract we present a Bayesian methodology for estimating the posterior model probability for a given selection of models. This probability is of interest to indicate statistical model uncertainty, and therefore uncertainty in the interpretation of the data. This methodology also has the potential to provide metrics that are more stable, and therefore more sensitive to a wider range of treatment effects than existing methods.

Signal Models Four diffusion signal models are considered here, though the statistical framework for model selection given later extends directly to other models.

$$M_1 : S(b) = S_0 \exp(-b\,D) \qquad M_2 : S(b) = S_0 \exp(-(b\,D)^{\alpha}) \qquad M_3 : S(b) = S_0 \int \pi(D) \exp(-b\,D)\,\mathrm{d}D \qquad M_4 : S(b) = S_0 \left(f \exp(-b\,D^*) + (1-f) \exp(-b\,D) \right)$$

M_1 is derived from a pure diffusion process; M_2 is derived from a sub-diffusion process [1]; M_3 assumes that the data arises from a population of particles with diffusion rates defined by the distribution $\pi(D)$, a normal distribution truncated below zero [2] with mode Δ and variance σ_D^2; M_4 is a bi-exponential form where the first term comes from an intra-voxel incoherent motion (IVIM) model [3], typically associated with a perfusion effect, and the second term describes a diffusion effect.

Bayesian Model Selection Bayesian methods can be used to derive posterior model probabilities that indicate how well each model is supported by the data. This approach requires the specification of prior distributions for all unknown parameters, and likelihood functions describing measurement errors. An important feature of using Bayesian Inference is that more complex models are appropriately penalised relative to simpler models that fit the data equally well, given the data SNR. From each of the above signal models we derive a likelihood function, the form of which reflects the measurement noise model – here we use a Gaussian model. Each model uses a different set of parameters: $\theta_1 = \{S_0, D\}$ for M_1, $\theta_2 = \{S_0, D, \alpha\}$ for M_2, $\theta_3 = \{S_0, \Delta, \sigma_D\}$ for M_3 and $\theta_4 = \{S_0, D, D^*, f\}$ for M_4. Prior distributions are needed for all unknown parameters for all models, and these are chosen to reflect known limits and plausible values, while being sufficiently dispersed to introduce minimal bias to the posterior parameter estimates. The following distributions and 95% *a priori* probability intervals are used: Gaussian distribution with $41.2 < S_0 < 158.8$ for all models; log-Gaussian distribution with $1.3\times10^{-4} < D < 6.5\times10^{-3}$ mm^2/sec for models 1, 2, 4, and also for Δ in model 3; Uniform distribution with $0.025 < \alpha < 0.975$ for model 2; Uniform distribution with $0.125\times10^{-3} < \sigma_D < 4.875\times10^{-3}$ mm^2/sec for model 3; Uniform distribution with $0.025 < f < 0.975$ for model 2; log-Gaussian distribution with $4.3\times10^{-3} < D^* < 0.21$ mm^2/sec for model 4. The prior probability for each model is also required and is set to ¼ for all models, indicating no prior model preference. Bayesian Inference combines the likelihood and prior distributions to give a posterior probability for each model. This inference process essentially involves averaging the data fit quality of each model over the whole parameter space, which amounts to a multi-dimensional integral, implemented with a grid-based approach here. In contrast, non-Bayesian methods typically compare the data fit quality of each model for the *best fitting* parameters only, which leads to less robust results.

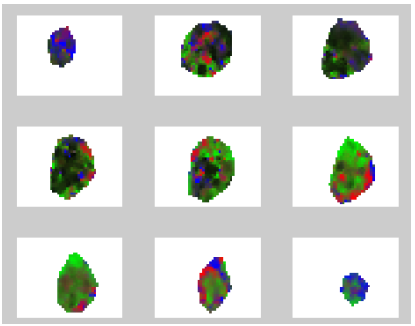

Figure 1: Nine slices of a liver metastasis, with head to foot slices arranged from left to right, top to bottom. The color indicates the posterior model probability with black being $p(M_1 \mid S_{1:N}) = 1$, red the same for M_2, green for M_3 and blue for M_4. Blends of colours indicate two or more models have similar probabilities, and so intense colours indicate high probability for the given model.

Data Acquisition Axial diffusion-weighted images were acquired under free-breathing using a 1.5T Siemens Avanto with a multi-slice EPI sequence with the following parameters: 20×5mm slices, FOV = 380mm, 128^2 matrix with 6/8 partial acquisition in PE direction, TE = 69ms, TR = 3500ms, NSA = 6, GRAPPA factor = 2, SPAIR fat suppression, b-values = 0, 50, 100, 300, 600, 900, 1050 s/mm^2, 3 orthogonal directions, 3-scan trace images, acquisition time 6 min 51 sec.

Case	Location	M_1	M_2	M_3	M_4
1	Liver	23.5	18.7	42.4	15.4
2	Pelvis	28.8	27.5	33.3	10.5
3	Liver	43.8	10.0	39.1	7.09
4	Liver	26.1	14.1	48.2	11.6
5	Liver	41.0	6.70	48.2	4.10
6	Pelvis	40.2	15.6	33.3	10.8
7	Pelvis	23.2	26.3	34.7	15.8

Table 1 : Percentage of voxels preferring each model.

Results & Discussion Figure 1 shows maps of the posterior model probability calculated from nine slices covering a liver metastasis. This example is typical of tumours we have analysed and indicates a large degree of model heterogeneity over the volume, with regions that have a strong preference for each of the four models. An interesting feature of these maps is the spatial smoothness – there are regions that have strong spatial correlations, but also edges where the model probability changes more rapidly. Table 1 shows results from four liver and three non-liver example cases where the proportion of voxels preferring each model is reported – the model preference for each voxel is the model with the largest posterior probability. Case 1 is the example in figure 1, and there is no obvious relationship between tumour location and model preference. The average model preferences over all seven cases are M_1 = 32.4%, M_2 = 17.0%, M_3 = 39.9%, M_4 = 10.8%, and as with the example in the figure, all cases have a significant proportion of pixels preferring each of the four models. It is possible that the estimated model preferences reflect underlying differences in the actual diffusion processes over the VOI, but without independent verification this interpretation is over-optimistic. This is because the "true" model is almost certainly *considerably* more complex than is possible to infer from such in-vivo data. Instead the preferred model should be treated as the most parsimonious description of the data chosen from the available models. Using over-parameterised models typically leads to better data fits, but the parameter estimates tend to be noisy due to large correlations between unnecessary parameters. Using the most parsimonious model is advantageous because the model complexity is matched to the data, and hence the parameter estimates tend to be more stable. The model preferences can be used as a metric describing the data, and we are currently investigating this in relation to treatment response of tumours. All four models have diffusion and amplitude parameters, so it is possible to compute model-averaged D and S_0 estimates for each voxel that may also be of use detecting treatment response. Similar metrics are needed for the other model parameters, but as these are specific to each model, further work is needed to determine the best method for combining estimates of these parameters with the rest of the analysis.

Conclusions A Bayesian methodology has been presented for analysing diffusion weighted images acquired with multiple b-values that gives posterior probabilities for a collection of signal models. These probabilities should only be interpreted as indicating the most parsimonious description of the data, and NOT as an indication of the true underlying diffusion processes. Examples have been presented indicating that the preferred model is highly heterogeneous within and between tumours.

Acknowledgements We acknowledge the support received from the CRUK and EPSRC Cancer Imaging Centre in association with the MRC and Department of Health (England) grant C1060/A10334, also NHS funding to the NIHR Biomedical Research Centre.

[1] Bennett et al., *Mag. Res. Med.* **50**:727-734 (2003). [2] Yablonskiy et al., *Mag. Res. Med.* **50**:664-669 (2003). [3] Le Bihan et al., *NMR Biomed.* **15**:431-434 (2002).